Advances in Intelligent Systems and Computing 331

Somnuk Phon-Amnuaisuk
Thien-Wan Au *Editors*

# Computational Intelligence in Information Systems

## Proceedings of the Fourth INNS Symposia Series on Computational Intelligence in Information Systems (INNS-CIIS 2014)

Springer

# Advances in Intelligent Systems and Computing

Volume 331

**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

Somnuk Phon-Amnuaisuk · Thien-Wan Au
Editors

# Computational Intelligence in Information Systems

## Proceedings of the Fourth INNS Symposia Series on Computational Intelligence in Information Systems (INNS-CIIS 2014)

Springer

*Editors*
Somnuk Phon-Amnuaisuk
Institut Teknologi Brunei
Gadong
Brunei Darussalam

Thien-Wan Au
Institut Teknologi Brunei
Gadong
Brunei Darussalam

ISSN 2194-5357 ISSN 2194-5365 (electronic)
ISBN 978-3-319-13152-8 ISBN 978-3-319-13153-5 (eBook)
DOI 10.1007/978-3-319-13153-5

Library of Congress Control Number: 2014954027

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

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

# Preface

The International Neural Network Society Symposia Series (INNS Symposia Series) is a biennial event. The first country that hosted the INNS Symposia Series was New Zealand back in 2008; followed by Peru and Thailand for the second and third Series, respectively. The fourth INNS symposia series was held in Bandar Seri Begawan, Brunei from November 7 to November 9, 2014 at Institut Teknologi Brunei. The event provides a platform for researchers to exchange the latest ideas and to present new research advances in general areas related to computational intelligence and their applications in various domains. The theme for the fourth Symposia Series was “Computational Intelligence in Information Systems”.

This year, the international program committees were formed by 75 researchers from 15 different countries. A total of 72 submissions from 16 countries were submitted to the event for double-blind peer-review. Of those, 34 submissions (47.2%) were selected to be included in this book.

First and foremost, we would like to thank the keynote speakers, the invited speakers and all the authors who had spent time and effort to make important contributions to this event. We would also like to thank the symposium chairs and all their members who had provided expert evaluation of the submitted papers; members of the local organizing committee for their contribution, despite their busy schedule; Pg Hj Mohd. Esa Al-Islam Bin Pg HjMd. Yunus and all members of the steering committee for their useful advice; Irwin King, Jonathan H. Chan, and Hussein A. Abbass for their useful suggestions and discussions; Janusz Kacprzyk, the series editor of the “Advances in Intelligent Systems and Computing” book series; Hjh Zohrah binti Haji Sulaiman, our Acting Vice Chancellor; and last but not least, Hjh Noor Maya binti Haji Md Salleh, who was our immediate supervisor when we decided to host the INNS symposia series in Brunei.

We would also like to acknowledge the following organizations: Institut Teknologi Brunei for its institutional and financial support, and for providing premises and administrative assistance; Bank Islam Brunei Darussalam (BIBD) and DST for their generous financial support; International Neural Network Society (INNS); King Mongkut Institute of Technology, Ladkrabang; King Mongkut University of Technology, Thonburi; ICONIP 2014 organizing committee; and Springer for their technical support.

We thank all of you and we hope that you will consider contributing to the future INNS symposia series event.

September 2014

Somnuk Phon-Amnuaisuk
Thien-Wan Au

# Organization

## Honorary Chair/Advisor

| | |
|---|---|
| Hjh Zohrah binti Haji Sulaiman | Acting Vice Chancellor, Institut Teknologi Brunei |

## Steering Committee

| | |
|---|---|
| Hj Mohd. Esa Al-Islam Bin Pg Hj Md. Yunus | Institut Teknologi Brunei |
| Saiful Baharin Pg Duraman | Institut Teknologi Brunei |
| Jennifer Hiew Lim | Institut Teknologi Brunei |
| Hamdani Bin Hj Ibrahim | Institut Teknologi Brunei |

## Advisory Committee

| | |
|---|---|
| Danil Prokhorov | TRINA, Toyota Technical Center, USA - President, INNS |
| Irwin King | Chinese University of Hong Kong, China |
| Ali Minai | University of Cincinnati, USA |
| David Casasent | Carnegie Mellon University, USA |
| Hava Siegelmann | University of Massachusetts, USA |
| Ron Sun | Rensselaer Polytechnic Institute, USA |
| Nikola Kasabov | Auckland University of Technology, New Zealand |
| Geraint Wiggins | Queen Mary University of London, UK |
| Hussein A. Abbass | University of New South Wales, Australia |
| Monai Krairiksh | King Mongkut Institute of Technology, Ladkrabang, Thailand |
| Jonathan H. Chan | King Mongkut University of Technology, Thonburi, Thailand |

## General Chair/Co-chair

| | |
|---|---|
| Somnuk Phon-Amnuaisuk | Institut Teknologi Brunei |
| Au Thien Wan | Institut Teknologi Brunei |

## Secretariat

| | |
|---|---|
| Mohammad Saiful Bin Hj Omar | Institut Teknologi Brunei |
| Hj Mohd Noah bin Hj A. Rahman | Institut Teknologi Brunei |

## International Journal Publication Chairs

| | |
|---|---|
| Lakhmi C. Jain | University of South Australia |
| Chee-Peng Lim | Deakin University, Australia |
| Srikanta Patnaik | SOA University, India |

## International Publicity Chairs

| | |
|---|---|
| Hj Ady Syarmin bin Hj Md Taib | Institut Teknologi Brunei |
| Wlodzislaw Duch | Nicolaus Copernicus University, Poland |
| Penousal Machado | University of Coimbra, Portugal |
| Simon Colton | Imperial College, UK |
| Soo-Young Lee | KAIST, Korea |

## Scientific Program Committee

| | |
|---|---|
| Ah-Hwee Tan | Nanyang Technological University, Singapore |
| Andrew Teoh Beng Jin | Yonsei University, Korea |
| Asim Roy | Arizona State University, USA |
| Fredric M. Ham | College of Engineering USA |
| Geraint Wiggins | Queen Mary University of London, UK |
| Irwin King | Chinese University of Hong Kong, China |
| Jonathan H. Chan | King Mongkut University of Technology, Thonburi, Thailand |
| Juyang Weng | Michigan State University, USA |
| Kevin Wong | Murdoch University, Australia |
| Loo Chu Kiong | University of Malaya, Malaysia |
| Marley Vellasco | PUC-Rio, Brazil |
| Paul Pang | Unitec Institute of Technology, New Zealand |
| Penousal Machado | University of Coimbra, Portugal |
| Simon Colton | Imperial College, UK |
| Soo-Young Lee | KAIST, Korea |
| Sung Bae Cho | Yonsei University, Korea |
| Somnuk Phon-Amnuaisuk | Institut Teknologi Brunei, Brunei |
| Ting Chuan Kang | National Chung Cheng University, Taiwan |
| Zhi-Hua Zhou | Nanjing Univeristy, China |

## Local Organising Committee

### Sponsorship Committee

| | |
|---|---|
| Hj Rudy Erwan bin Hj Ramlie | Institut Teknologi Brunei |
| Hj Idham Maswadi bin Hj Mashud | Institut Teknologi Brunei |

### Technical Committee

| | |
|---|---|
| Franky Looi Hong Cheong | Institut Teknologi Brunei |
| Suresh Sankaranarayanan | Institut Teknologi Brunei |
| Hj Afzaal H. Seyal | Institut Teknologi Brunei |
| Bruno Starrs | Institut Teknologi Brunei |

### Finance Committee

| | |
|---|---|
| Siti Hanisah binti Hj Mohd Salleh | Institut Teknologi Brunei |
| Noor Deenina Hj Mohd Salleh | Institut Teknologi Brunei |

### Fair, Exhibition & workshop Committee

| | |
|---|---|
| Armanadurni binti Abdul Rahman | Institut Teknologi Brunei |

### Ceremony Committee

| | |
|---|---|
| Ak Hj Azhan bin Pg Hj Ahmad | Institut Teknologi Brunei |

### Welfare & Accommodation Committee

| | |
|---|---|
| Wong Pong Kit | Institut Teknologi Brunei |

### Publication Committee

| | |
|---|---|
| Peter David Shannon | Institut Teknologi Brunei |

### Invitation & Protocol committee

| | |
|---|---|
| Norhuraizah binti Hj Md. Jaafar | Institut Teknologi Brunei |

## Refreshment Committee

| | |
|---|---|
| Mardiyah binti Hj Matusin | Institut Teknologi Brunei |

## Photography Committee

| | |
|---|---|
| Ibrahim Edris | Institut Teknologi Brunei |

## Logistics Committee

| | |
|---|---|
| Pg Siti Awa Bte Pg Hj Mahmud | Institut Teknologi Brunei |
| Hj Irwan Mashadi bin Hj Mashud | Institut Teknologi Brunei |

## Publicity Committee

| | |
|---|---|
| Hj Ady Syarmin bin Hj Md Taib | Institut Teknologi Brunei |
| Siti Asmahlati Bolkini binti Ahmad | Institut Teknologi Brunei |

## Souvenir & Certificates Committee

| | |
|---|---|
| Hjh Nor Zainah binti Hj Siau | Institut Teknologi Brunei |
| Siti Noorfatimah binti Hj Awg Safar @ Hj Sapar | Institut Teknologi Brunei |

## Symposium Chairs

### SoDAC Chair

| | |
|---|---|
| Khor Kok Chin | Multimedia University, Malaysia |

### SoEDM Chair

| | |
|---|---|
| Mohd Shahrizal Sunar | Universiti Teknologi Malaysia, Malaysia |

### SoSHI Chair

| | |
|---|---|
| Paul Chung | Loughborough University, UK |

### SoUC Chair

| | |
|---|---|
| M. I. Seyed Mohamed Buhari | King Abdulaziz University, Saudi Arabia |

### SoVIP Chair

| | |
|---|---|
| Chee Seng Chan | University of Malaya, Malaysia |

## Members

| | |
|---|---|
| Abdelrahman Osman Elfaki | Tabuk University, Kingdom of Saudi Arabia |
| Abdollah Dehzangi | Griffith University, Australia |
| Adham Atyabi | Salford University, UK |
| Ahmad Rafi | Multimedia University, Malaysia |
| Ali Selamat | Universiti Teknologi Malaysia, Malaysia |
| Ang Tan Fong | University of Malaya, Malaysia |
| Arosha Senanayake | Universiti Brunei Darussalam, Brunei |
| Atikom Ruekbutra | Mahanakorn University of Technology, Thailand |
| Beibei Cheng | Microsoft, USA |
| Chaiwat Nuthong | King Mongkut Institute of Technology, Ladkrabang, Thailand |
| Chin Wen Cheong | Multimedia University, Malaysia |
| Chikkannan Eswaran | Multimedia University, Malaysia |
| Eran Edirisinghe | Loughborough University, UK |
| Fatih Alagoz | Bogazici University |
| Farnan Mohamed | Universiti Teknologi Malaysia, Malaysia |
| Foo Lee Kien | Multimedia University, Malaysia |
| Franky Looi Hong Cheong | Institut Teknologi Brunei, Brunei |
| Gerald Schaefer | Loughborough University, UK |
| Goi Bok Min | University Tunku Abdul Rahman, Malaysia |
| Goh Chien Le | Multimedia University, Malaysia |
| Goh Hock Guan | University Tunku Abdul Rahman, Malaysia |
| Hj Afzaal H. Seyal | Institut Teknologi Brunei |
| Hoshang Kolivand | Universiti Teknologi Malaysia, Malaysia |
| Indra Adji Sulistijono | Polytechnic Institute of Surabaya, Indonesia |
| Kannan Ramakrishnan | Multimedia University, Malaysia |
| Karunakar A. Kotegar | Manipal lnstitute of Technology, lndia |
| Karthik Subburathinum | SNS College of Technology |
| Khaled Salah | Khalifah University |
| Janos Botzheim | Tokyo Metropolitan University, Japan |
| Junaidi Abdullah | Multimedia University, Malaysia |
| Linda Chua Sook Ling | Multimedia University, Malaysia |
| Masaki Oshita | Kyushu lnstitute of Technology, Japan |
| Mei Kuan Lim | University of Malaya, Malaysia |
| Mohammad Saiful bin Hj Omar | Institut Teknologi Brunei, Brunei |
| Mohd Fadzil Hassan | Universiti Teknologi Petronas, Malaysia |

| | |
|---|---|
| Muhammad Qaiser Saleem | Al Baha University |
| Ng Keng Hoong | Multimedia University, Malaysia |
| Naeimeh Delavari | Georgia Institute of Technology, USA |
| Nantha Kumar | Open University Malaysia, Malaysia |
| Satria Mandala | Universiti Teknologi Malaysia, Malaysia |
| Shaheen Fatima | Loughborough University, UK |
| Suresh Sankaranarayanan | Institut Teknologi Brunei, Brunei |
| Ting Choo Yee | Multimedia University, Malaysia |
| Ukrit Watchareeruetai | King Mongkut Institute of Technology, Ladkrabang, Thailand |
| Voon Nyuk Hiong | Institut Teknologi Brunei, Brunei |
| Waqur ul Quonain | Punjab University College of lnformation Technology, Pakistan |
| Wai Lam Hoo | University of Malaya, Malaysia |
| Werasak Kurutach | Mahanakorn University of Technology, Thailand |
| Yun-Li Lee | Sunway University, Malaysia |
| Zhaojie Ju | University of Portsmouth, U.K. |

# Contents

## Invited Paper

## Applied Intelligent Computing

## Data Analytics

## Emerging Digital Media

## Sports and Health Informatics

## Ubiquitous Computing

# Advances in Workflow Systems

Paul Wai Hing Chung

Department of Computer Science, Loughborough University LE11 3TU
`p.w.h.chung@lboro.ac.uk`

**Abstract.** In modern enterprises, workflow technology is commonly used for business process automation. Established business processes represent successful business practice and become a crucial part of corporate assets. Workflow systems allow processes to be defined and the associated information flow to be specified to ensure that tasks are done in the right order and necessary information is being passed to the appropriate tasks. Such systems are typically very rigid to ensure compliance to agreed processes. However, the disadvantages are that they are unresponsive to changes in circumstances or are not suitable for supporting interactions between organisations. This paper highlights some of the advances that have been made in creating cross-organisational workflows and dynamic workflow systems, addressing research issues such as cross-organisation workflow compatibility and workflow compliance in adaptive workflow systems.

**Keywords:** Workflow Systems, Adaptive Workflow, Compliance Checks, Cross-organisational Workflow, Workflow System Architecture, Workflow Generation.

## 1 Introduction

Back in 1996 the workflow management system market was already estimated to be worth US$2 billion. Recently the forecast is that the market will reach US$5.53 billion by 2019 [1]. Market leaders provide workflow solution with integrated enterprise information systems to meet the needs of large organisations.

Workflow process definitions specify the order that a sequence of activities should be carried out, the people responsible for carrying out those activities and the output expected from each activity. Workflow management systems have two major components. The build-time component is responsible for specifying workflow processes that capture the practice of an organisation ready for execution. The runtime component is for supporting the execution of workflow instances of the specified workflow to enable an organisation to get on with its business [2], see figures 1 and 2.

The above models work for processes that are stable and do not require interactions with other processes. However, to support B2B (business to business) transactions workflow instances need to interact with one another. Organisations normally have their own workflow processes and they can be quite different among organisations. To create compatible workflows so that organisations can carry out mutual transactions in an integrated fashion can be a very complex and time consuming process.

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_1

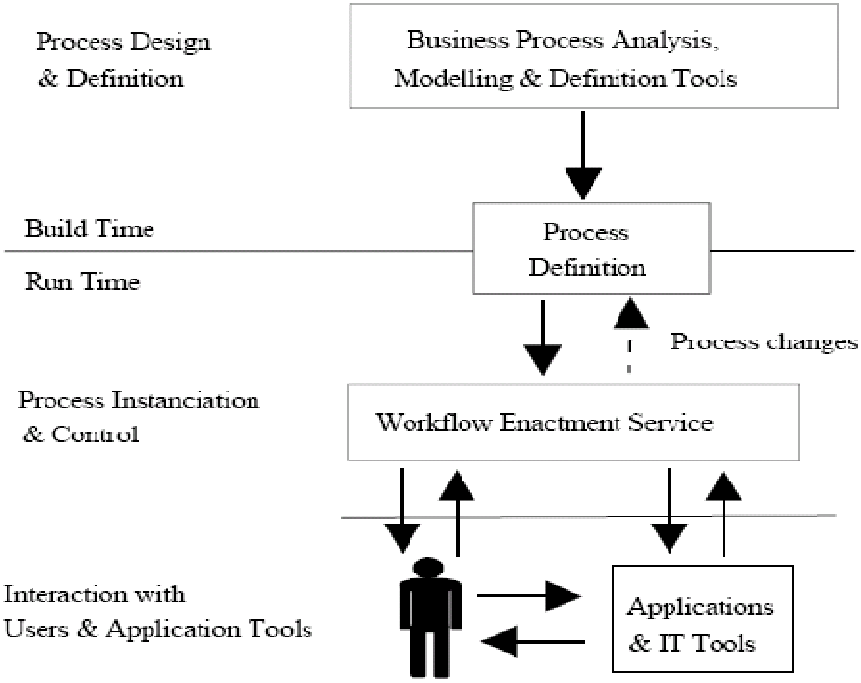


**Fig. 1.** Key workflow components (adapted from [2])

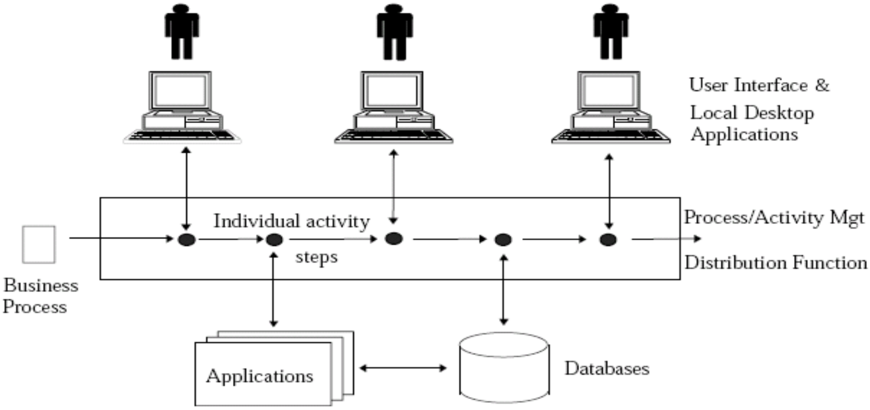


**Fig. 2.** Activity-based workflow (adapted from [2])

Section 2 of this paper considers two different approaches for creating cross-organisation workflows, i.e. workflows that include interactions across organisation boundaries. Another emerging area of research is adaptive workflow systems. These are required to support dynamic processes where tasks may be added to, or deleted from, workflow instance during runtime. Section 3 of this paper considers the support for checking the compliance of adaptive workflows at build time and at run time. Conclusions and ideas for future work are given in section 4.

## 2 Cross-Organisation Workflow

The concept of cross-organisational workflow collaboration is emerging as there is an increasing demand for IT systems to work together to support B2B (business to business) transactions. An example is the interactions between vendors and customers as found in the manufacturing sector. A workflow instance of the customer and a workflow instance of the vendor will have a number of interaction tasks. An example of such a task is that the customer sends an acceptance note to the vendor after receiving a quotation. Therefore, the workflow process definitions of both the vendor and the customer need to include interaction tasks between the two workflows as well as tasks local to the vendor and the customer. This is illustrated in figure 3 where an interaction point between two collaborating processes is shown.

The work by [3] highlighted the need that interaction points between collaborative processes need to be modelled and that they need to be in the right places in the processes. Each interaction point is represented as a paired interface activities, i.e. one for each process. Control of the execution of the workflow is passed from one process to another through the interface activities.

This emerging field of cross-organisation workflow opens up new challenges due to workflow incompatibility. Processes from different organizations often follow different standards and work on different systems so compatibility is hard to achieve and there is difficulty in directly linking and automating cross-organization processes to work in an integrated fashion. Therefore, there is a significant requirement for creating compatible workflows between collaborating organisations before linking and automation can take place. The obvious bottom-up approach of creating compatible workflows for B2B collaboration is very labour intensive and time consuming and novel ways of supporting the creation process is required. Two approaches that provide different levels of automation are discussed. The first approach provides IT-support for reconciling mismatches between workflows [3]. The second approach deals with the generation of compatible workflows automatically to meet the stated goals of the collaborating organisations [4].

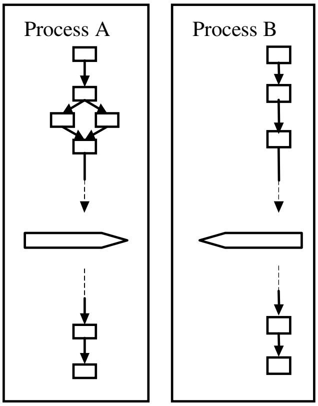


**Fig. 3.** Collaborative processes with an interaction point

If the interaction activities of two interaction processes are properly modelled and in exactly the same order on both sides then the two processes can certainly collaborate with each other given a proper communication channel. However, two sets of interaction processes can be made to collaborate even if not all the interaction activities are in the same order. An example is shown in Figure 4. In the example, each interaction process has activities labelled ‘a’ to ‘h’. A ‘sending’ activity is denoted as ‘[s]’ and a ‘receiving’ activity is denoted as ‘[r]’. The two processes can be executed by passing control from one to another without reaching a deadlock or a dead end. The drawback is that the execution of the two processes may experience a delay because activity g[r] in process A has to wait for the control to be passed to it until after activity g[s] in process B is completed.

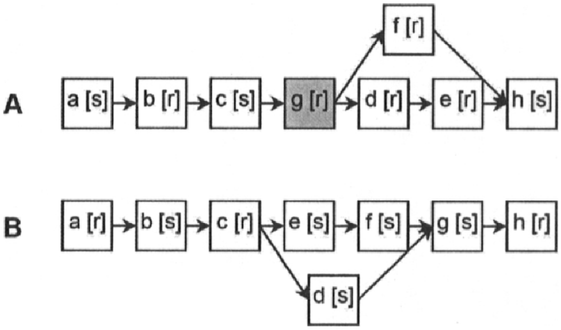


**Fig. 4.** Example of interaction processes with potential delay

### 2.1 Cross-Organisation Workflow Reconciliation through Negotiation

The word 'negotiation' is derived from the Latin word 'negociare' - to conduct business. Most negotiation was first understood in the context of business transactions although it has been in a wider context for quite some time. As an approach to alternative dispute resolution, negotiation has been studied and defined in many ways [5, 6, 7, 8, 9]. The process of negotiation includes the steps of identifying conflicts, reconciling differences and reaching agreement.

Therefore, IT-support for cross-organisation workflow reconciliation through negotiation needs to provide functionalities for the following steps:

1) Process matching to identify conflicts or causes of delay;
2) Generation of options for activity re-ordering to remove conflicts or delays;
3) Present options to business partners alternately to accept or reject proposed changes until agreement is reached or one partner withdraw from the negotiation.

Algorithms for process matching are described in [8, 9]. A prototype for supporting steps 2 and 3 above are described in [3].

The provision of IT support for reconciliation through negotiation takes away some of the labour intensive effort for creating compatible workflows from existing workflows. However, it still requires human involvement. The problem is further compounded as an organization may have to work with many trading partners. Therefore, negotiation will have to be done individually and different processes will have to be created for each partner.

### 2.2 Automatic Cross-Organisation Workflow Generation

To overcome the limitations identified above, another approach to solve the problem is by dynamically synthesising workflow processes that satisfy the stated business rules of collaborating partners without the need of human intervention. This approach utilises state-of-the-art research in service-oriented architecture, agent based technologies and planning technologies. A service-oriented architecture allows participating organisations to specify the services that they offer and the services that they wish to employ. The business practices/rules in an organisation that governs the sequence of activities and data requirements of the services are specified in a declarative manner. Taking this information, rather than strictly defined workflow processes, an intelligent planning agent can synthesize compatible workflow processes to allow potential partners to collaborate. The service-oriented architecture can also be extended to handle the run time aspect of workflow so that the whole process from partner matching, workflow creation and workflow execution can be fully integrated and supported. This approach proposed in [4] offers considerable time saving for collaborating partners. The system architecture is shown in figure 5.

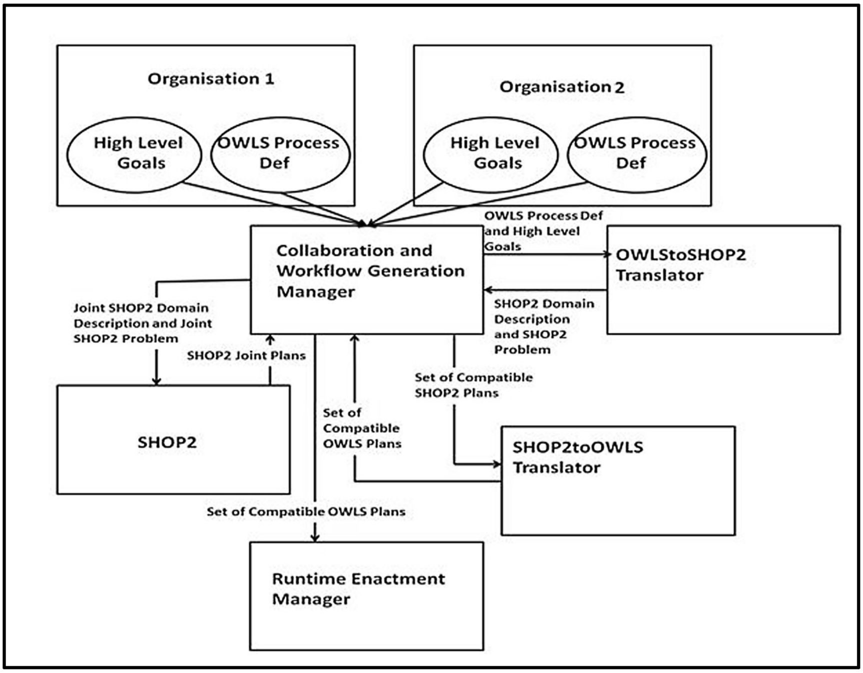


**Fig. 5.** Architecture for Compatible Collaborative Workflow Generation (adapted from [4])

In this architecture Web Ontology Language for Services (OWLS) [14] is used to represent atomic activities with the pre- and post-conditions for executing those activities. For example, an atomic activity for a vendor to receive advance payment from a customer would have the following information:

*Name: receive_AdvancePayment (receive advance payment)*
*Inputs/Preconditions: send_AdvancePayment (advance payment sent by customer)*
*Outputs/Effects: AdvancePaymentReceived (advance payment received from customer)*

SHOP2 is a hierarchical task network planner [13] and it is available as open source software. [4] modified SHOP2 for the purpose of generating joint plans by taking the required input from the collaborating organisations. In this architecture, workflow specifications are represented as plans. The interface activities between the two plans are clearly labelled so that runtime workflow systems will pass control from one system to another. This research demonstrated that compatible workflows for collaborating organisations can be generated automatically. Given a set of atomic activities that each organisation needs to carry out, the system can produce different plan options for achieving the stated goals for all the participating organisations. Human intervention is kept to a minimum through the use of web services.

# 3 Adaptive Workflow

Another area that presents research challenges is adaptive workflow [15]. To develop a reliable system or product, the current best practice for the development process is typically embodied in standards and guidelines, such as IEC61508 for safety. The purpose of such a standard is to provide a systematic framework for all the necessary activities that need to be carried out to meet a specified quality level. However, no two projects are the same with all the details in exactly the same order. In safety-related engineering projects significant amounts of resource are spent on their management and in demonstrating standards compliance; much of the time of developers, managers and quality assurance teams is occupied with tracking and managing the compliance of the project. A workflow system with compliance management ability can considerably shorten the development time and reduce costs. Current workflow systems lacks the ability to automatically check for compliance when workflow specifications are being produced and when workflow instances are being modified during runtime due to the addition or deletion of activities.

## 3.1 Compliance Checks

Compliance Flow [16] is one of the first major efforts to introduce compliance checks in workflow systems. The prototype implemented four different checks:

- Completeness Check – to check that all the activities and expected outputs specified in the standard are included in the user-defined process.
- Correctness Check – to check that the sequence of activities and expected outputs specified in the user-defined process is in accordance with the selected standard.
- Capability Check – to check that the required capabilities of an agent allocated to carry out an activity in the user-defined process match those that are specified in the selected standard.
- Recommendation Check – to check that the recommended techniques, measures, tools or methods for performing a particular activity as specified in the selected standard are fully considered in the user-defined process.

These checks are used for assisting in process specification, identifying compliance errors, and preventing non-compliant tasks in a process from being performed.

To implement the above checks, the required standard is represented as a reference model – the model of standards – for matching and checking. Figure 6 provides an illustrative example of matching that particular documents are included in the user-defined process and they are produced in the correct order.

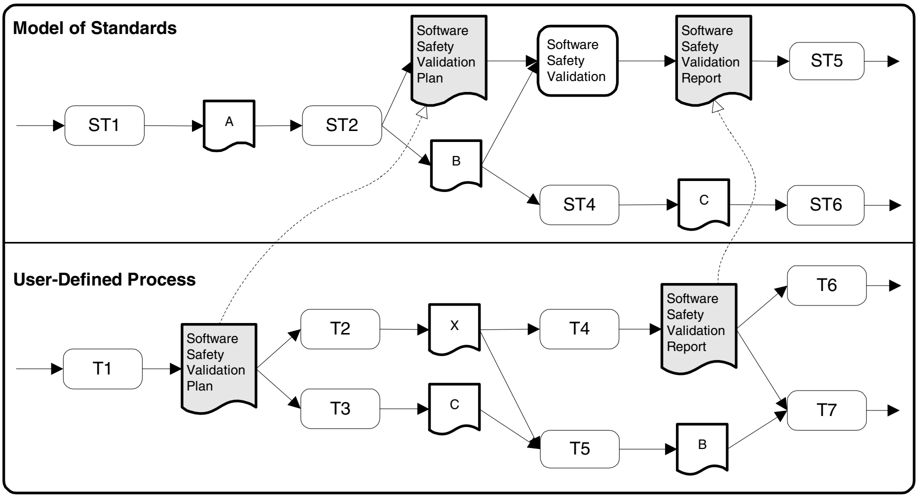


**Fig. 6.** An example illustrating the matching of required outputs are generated in the correct order

### 3.2 Structural Checks and Data Driven Checks

The work described in [17] points out that the different checks described above could be categorised into two types: structural checks and data driven checks. Completeness and correctness checks belong to the first type; capability and recommendation checks belong to the latter. It is suggested a generic rule-based component is integrated with a workflow system. Rules can be defined in a declarative manner for checking the corresponding data items in the model of standards and user-defined process. The rules will be triggered when data items are changed. Rules can also be specified to monitor the progress of the user-defined process as it evolves. For example, if a project is meant to have monthly project meetings then a reminder rule could be written so that it is triggered when the number of progress meetings held falls behind the number of months since the start of the project by a specified threshold, e.g. 2:

*if (number of months since start of project - number of progress meetings held –) > 2 then issue reminder.*

## 4 Conclusions and Future Work

This paper described some of the advances that have been made in providing support features for cross-organisational workflow systems and adaptive workflow systems. The aim is to reduce the effort required to create compatible workflow processes for collaboration in the former case, and the aim of the latter case is to reduce the effort required to ensure the compliance of adaptive workflows. Much work still needs to be done to develop these ideas so that they are sufficiently mature and accepted for use in deployed systems.

The potential to develop workflow systems to meet other business needs is huge. Workflow systems capture a wealth of information for an organisation. For example: Who did what and what activity was delayed? Treating a workflow system as a management information system would be an interesting avenue for exploration.

## References

1. `http://www.marketsandmarkets.com/PressReleases/workflow-management-system.asp` (accessed on July 31, 2014)
2. WfMC. The Workflow Reference Model – Technical Report WFMC-TC-1003. Workflow Management Coalition (1995)
3. Chen, X., Chung, P.W.H.: Facilitating B2B E-business by IT-supported business process negotiation services. In: Proceedings of IEEE International Conference on Service Operations and Logistics, and Informatics, pp. 2800–2805 (2008)
4. Saleem, M., Chung, P.W.H., Fatima, S., Dai, W.: A cross organization compatible workflows generation and execution framework. Knowledge-Based Systems 56, 1–14 (2014)
5. Longman Dictionary of Contemporary English. Pearson Education Limited, Essex England (2002)
6. Zartman, I.W.: Negotiation as a Joint Decision-Making Process. In: Zartman, I.W. (ed.) The Negotiation Process: Theories and Applications. Sage Publications, London (1977)
7. Hamner, W.C.: G A. YukI: The Effectiveness of Different Offer Strategies in Bargaining. In: Druckman, D. (ed.) Negotiations: Social Psychological Perspectives. Sage Publications, Beverly Hills (1977)
8. Gulliver, P.: Disputes and Negotiation: A Cross-Culture Perspective. Academic Press Inc., San Diego (1979)
9. Rosenschein, J.S., Zlotkin: Rules of Encounter. MIT Press (1994)
10. Hiltrop, J., Udall, S.: The Essence of Negotiation. Prentice Hall, London (1995)
11. Krukkert, D.: Matchmaking of ebXML business processes, Technical Report IST-28584-0X_D2.3_v.2.0, openXchange Project (2003)
12. Chen, X., Chung, P.W.H.: A simulation-based difference detection technique for bottom-up process reconciliation. In: Proceedings of 9th International Conference on Enterprise Information Systems (ICEIS 2007), Funchal, Madeira-Portugal (2007)
13. `http://www.w3.org/TR/owl-ref/` (accessed on September 2, 2014)
14. Sirin, E., Parsia, B., Wu, D., Hendler, J., Nau, D.: HTN planning for web service composition using SHOP2. Journal of Web Semantics 1(4), 377–396 (2004)
15. Chung, P.W.H., Cheung, L., Stader, J., Jarvis, P., Moore, J., Macintosh, A.: Knowledge-based process management – an approach to handling adaptive workflow. Knowledge-Based Systems 16(3), 149–160 (2003)
16. Chung, P.W.H., Cheung, L.Y.C., Machin, C.H.C.: Compliance Flow – Managing the compliance of dynamic and complex processes. Knowledge-Based Systems 21(4), 332–354 (2008)
17. Haji Omar, M.S., Chung, P.W.H., Dawson, C.W.: A Novel Workflow Management System for Handling Process Adaptation and Compliance. In: Proceedings of the 27th International Conference on Advanced Information Networking and Applications Workshops (WAINA), pp. 1174–1179. IEEE (2013)

# A Novel Algorithm to Detect a QPSK Signal with the Minimum Number of Samples

Mohammed Saleh[1], Hilmi Sanusi[1], Adamu I. Abubakar[2], Haruna Chiroma[3],
Edi Sutoyo[3], Mungad[3], and Tutut Herawan[3]

[1] Department of Electrical Electronic and System Engineering
Universiti Kebangsaan Malaysia
[2] Department of Information Systems
International Islamic University Malaysia
Kuala Lumpur, Malaysia
[3] Faculty of Computer Science and Information Technology
University of Malaya
50603 Pantai Valley, Kuala Lumpur, Malaysia
`mohamed.function@gmail.com, hilmi@eng.ukm.my, adamu@iium.edu.my,`
`hchiroma@acm.org, {edisutoyo,mungad,tutut}@um.edu.my`

**Abstract.** In digital communication the QPSK (quadrature phase-shift keying) signal occupies a special position, from satellite systems through to Local Area Networks (LAN). The QPSK signal can be detected or received using analog or digital receivers. The detection capability of digital receivers relies mainly on the number of samples per signal element and signal-to-noise ratio. The number of samples required by a digital receiver is the bottleneck for detection speed. This paper presents a novel algorithm to detect a QPSK signal with the minimum number of samples, or as few as four samples per symbol. The algorithm relies only on the samples' polarities, which leads to superior immunity to amplitude noise. The algorithm was simulated and tested using ANSI C++. The results showed that the algorithm is capable of detecting a QPSK signal with minimum Bit Error Rate (BER) at signal-to-noise ratio of 7.57 dB and maximum phase distortion of PI/8.

**Keywords:** Bit Error Rate, Quadrature Phase-Shift Keying, Signal-to-noise ratio, Maximum phase distortion.

## 1 Introduction

Throughout the past few decades, demand for wireless applications has increased tremendously. Accordingly, standards have had to move toward Wireless Personal Area Networks (WPAN) and Wireless Local Area Networks (WLAN) to fulfill market demands and lower the cost of manufacturing. Signaling schemes with higher order including the QPSK can be a particularly good alternative for their counterpart binary signaling schemes. The QPSK modulation scheme has the advantage of in-phase and quadrature channels to achieve the same error performance as Binary Phase

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_2

Shift Keying (BPSK) when the signal is coherently detected at the receiver side. Coherent QPSK signal detection is a costly technique because it relies on carrier phase and frequency synchronization circuits, such as the Phase Locked Loop Circuit (PLL) to track the signal phase in real time [1]. For the time being, we should bear in mind that a QPSK signal can be detected at the receiver side using either analog or digital techniques.

Demodulation is the opposite process, whereby a message or information signal is recovered. Demodulators can be analog or digital. Analog demodulators for QPSK signals abstractly consist of receiving antennae, a tuning circuit, filters to separate carrier from message signals, and a decision circuit to determine the received pattern. These components vary according to the manufacturer's design and implementation. However, a digital QPSK demodulator consists of antennae, an analog-to-digital converter, and a decision unit. The two systems may have some similarities, but when it comes to the decision circuit they are completely different.

QPSK signals provide more efficient throughput and preserve high immunity to amplitude noise. However, their immunity to phase noise is low, which makes QPSK signal detection using digital techniques very difficult due to the mandatory large number of samples required to detect the phase. Subsequently, the detection process would be slow and inaccurate, besides the massive processing power needed to accomplish the task. If it is possible to find an algorithm with the benefits of the characteristics of QPSK signal noise amplitude immunity while keeping the number of samples to a minimum, then it would be possible to obtain a system that can detect QPSK signals with high accuracy, less processing power and high immunity to phase distortion.

This paper introduces an algorithm that can detect QPSK signals very efficiently and accurately while using a minimum number of samples, as low as four samples, and yet has superior immunity to amplitude and phase distortion. Detecting a QPSK signal using a digital demodulator is very complicated and requires very sophisticated techniques for three main reasons. One reason is that the number of samples per signal element required to analyze and detect the signal pattern is relatively high. This is actually a serious problem in digital receiver design in general. In some applications like modulation scheme identification, 200 samples per signal element are required to detect the modulation scheme automatically [2]. The second reason is related to the processing speed of the Digital Signal Processor that is responsible for analysis and decision making. If the received signal frequency is high, then the sampling rate of the ADC, or ADC frequency, will be multiples of that frequency, meaning that the DSP should be able to detect the input signal before the next signal elements come –something that may not be feasible owing to hardware limitations. The third reason is the complexity of the applied algorithms. Although detection methods that rely on domain transformation, such as Wavelet and Discrete Cosine (DCT) transforms as in [3,4] have proven acceptable detection accuracy and require less computational time than methods that uses digital PLL for example, they still need high processing power.

In this paper, there are three main objectives that will help solve the previously emphasized problems. The first objective is to study and analyze the QPSK signal samples' polarities and find out if only these can help with detecting QPSK patterns efficiently. This first objective is also related to the algorithm complexity objective.

The second objective is to design an algorithm that uses the magnitude polarity to detect a QPSK signal with the least number of required sampled data. The third objective is to simulate the algorithm and identify the limitations, robustness and detection capability. The second and third objectives are related to algorithm implementation.

# 2 QPSK Signal

## 2.1 QPSK Signal Description

QPSK stands for Quadratic Phase Shift Keying or Quaternary Phase Shift Keying. It is a digital modulation technique used to modulate digital data into an analog signal that can be easily transmitted over a medium. QPSK combines the in-phase and out-phase of a sinusoidal signal in such a way to output a sinusoidal signal with a specific phase shift. Essentially, the phase of the signal changes according to the input bits. QPSK can carry two bits per signal element; therefore the modulated signal has one of four possible phases: PI*1/4, PI*3/4, PI*5/4 and PI*7/4 (see Figure 1). Figure 2 shows a typical QPSK modulator.

In [1], a square wave was used as a detection method to enhance the overall performance by 60%. This method does not require multiplication, which in turn means fewer components than the Conventional Phase Frequency Shift Keying (CPFSK).

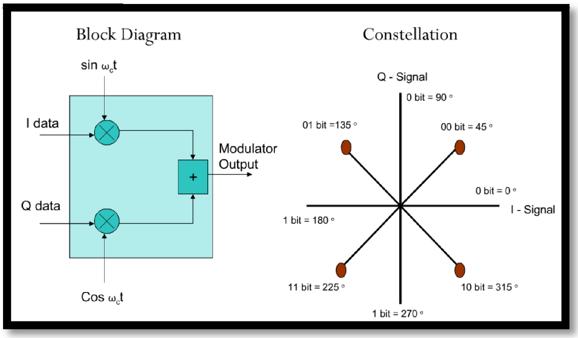


**Fig. 1.** I data and Q data represent a pair of bits

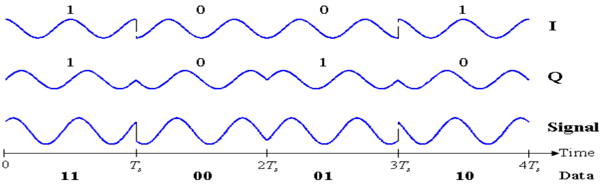


**Fig. 2.** Typical example of a QPSK modulator output

The Rayleigh Fading Channels in Mobile Digital Communication Systems is described as follows. Part 1: Characterization, and Part 2: Mitigation [5,6]. This paper elaborates each element that contributes to fading. It also explains the effect of Rayleigh fading on the 9 communication system where the fading will create delay and cause Inter-symbol Interference (ISI).

The four-sample QPSK Detector (4SQD) algorithm that will be used to detect a QPSK signal employs only four samples to identify the received pattern. This means less processing time, less power consumption and superior detection speed. According to the results, if the system requirements and conditions are fulfilled the system can attain a relatively high percentage of accuracy.

The 4SQD algorithm will be explained comprehensively in the next section together with the simulation environment, restrictions and limiting conditions, determining the beginning of the information signal and the relation between the rumble signal and information signal, sample position selection and finally, the rules to be used in identifying the QPSK pattern.

## 2.2 Restrictions and Limitation Conditions

In order for the designed algorithm to succeed in identifying the QPSK pattern five conditions must be fulfilled. The first condition is that the rumble signal frequency must be at least twice the information signal frequency. The second condition has two cases: the sampling rate is either constant or variable. In the case of constant sampling rate, the ADC frequency should be at least 16 times the information signal frequency (e.g., if the information signal frequency is 1 MHz then the ADC frequency should be 16 MHz or 16 samples per symbol). The optimum number of samples could be 32 samples per symbol. In case of using the variable sampling rate with a parallel ADC system to sample the information signal the sample width should be 1/16th of the signal period. The third condition assumes that the analog-to-digital converter employed can be replaced by a sample and hold circuit because the algorithm relies on the sample's polarity and not its value. The fourth and last condition is that the signal-to-noise ratio must be above 7.57 dB and the phase distortion must not exceed PI*0.125 (i.e. PI/8), otherwise symbol identification may not be accurate. In digital communication an information signal is divided and sent in frames. Each frame has rumble bits at the beginning. The rumble bits do not carry any information, but their main task is to help detect the beginning of the information signal.

## 2.3 The Beginning of an Information Signal

When a frame is sent the receiver catches the rumble signal somewhere at the beginning or before the information signal begins. The detection circuit initially locates the first ZCP and stores it in T0 then continues searching for the second ZCP and stores it in T1. If the time interval between T0 and T1 is equal to that of the rumble signal then the information signal has not reached yet and therefore T0 will store the value of T1 and T1 will store the value of the next ZCP. This process continues until the time interval between T0 and T1 is not equal to the rumble signal period. In this case T0 is considered the beginning of the signal and the information signal frequency is calculated by finding the next ZCP and storing it in T2. The frequency of the information signal is as follows:

$$\text{Information Signal Frequency} = 1 / (2 * (T2\text{-}T1)) \tag{1}$$

See the flow chart for the algorithm in Figure 3.

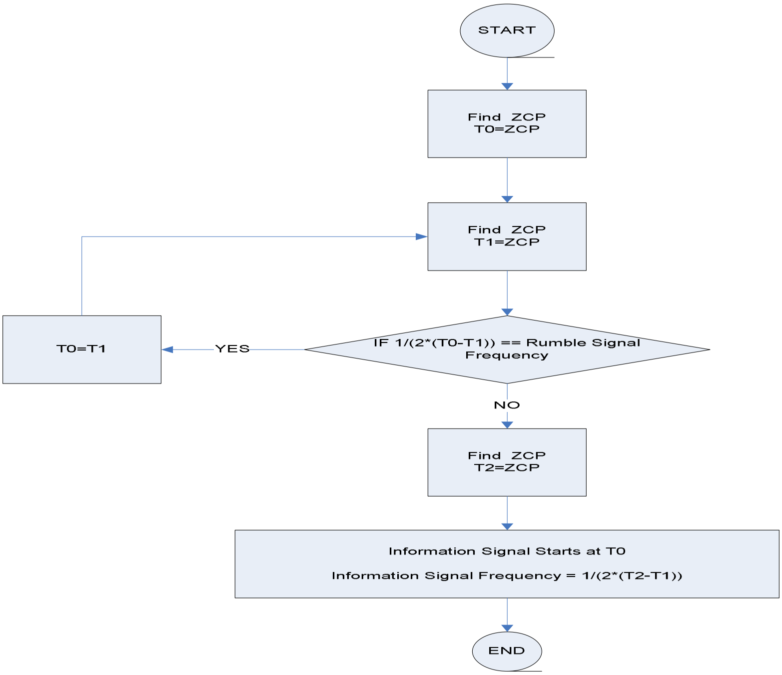


**Fig. 3.** Flow chart of the detection of the beginning of the information sample algorithm

In the 4SQD four samples are sufficient to identify the signal pattern or phase; yet these four samples must be in specific locations. One of two scenarios can be used to locate the samples. The first scenario is when a constant sampling rate ADC is used. The 4 samples' locations are defined as: S1= 0.025*(T), S2= 0.25*(T), S3= 0.75*(T), S4= 0.975*(T) where S1, S2, S3 and S4 are the samples' positions rounded to the smallest integer; T is the ratio between the information Signal Time Interval and the ADC Time Interval T = (1 / InformationSignalFrequency) / (1 / ADCFrequency). Because S1, S2, S3 and S4 are rounded to the smallest integer the samples' positions may not be accurate. Therefore, the number of samples per symbol plays an important role in locating the four samples.

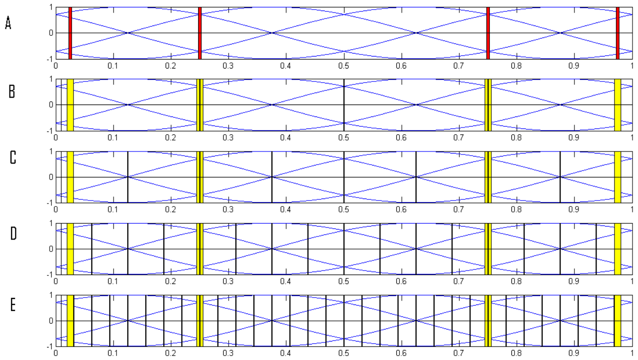


**Fig. 4.** Sample positions

Figure 4 is explained as follows. The yellow bars show the optimum samples' positions while the black bars show the actual ADC samples' positions. (A) signifies the optimum sample positions. (B) shows the difference between the optimum samples and actual samples in case of using 4 samples per symbol ADC. (C) shows the difference between the optimum and actual samples in case of using 8 samples per symbol ADC. (D) shows the difference between the optimum and actual samples in case of using 16 samples per symbol ADC. (E) shows the difference between the optimum and actual samples in case of using 32 samples per symbol ADC.

The minimum acceptable number of samples per symbol is 16 while 32 samples per symbol are considered optimum for practical implementation

In the second scenario, if samples can only be taken at specific times then only four samples can be obtained. This can be accomplished in two ways. The first way applies in case the information signal frequency is low and the ADC frequency is at least 8 times the signal info signal frequency. Then the ADC can be triggered at defined times. The second way is applicable for a high-frequency information signal. In this case four ADCs will be triggered, but each one has a relative delay with respect to the previous one. The signal positions in the second scenario are as follows:

S1=0.025*(Information Signal Time Interval)
S2=0.25*(Information Signal Time Interval)
S3=0.75*(Information Signal Time Interval)
S4=0.795*(Information Signal Time Interval)

## 2.4 Decision Algorithm

After acquiring the four samples, the transfer function in equation (2) is applied:

$$F(S) = 1 \text{ if } f(S) >= 0,\ 0 \text{ if } f(S) < 0 \quad (2)$$

where S is the sample position and f(S) is the amplitude of the sample at location S. The decision and pattern identification are done based on the following criteria:

PI*(1/4) pattern: F(S1)=1, F(S2)=1, F(S3)= -1, F(S4)=1
PI*(3/4) pattern: F(S1)=1, F(S2)=-1, F(S3)= 1, F(S4)=1
PI*(5/4) pattern: F(S1)=-1, F(S2)=-1, F(S3)= 1, F(S4)=-1
PI*(7/4) pattern: F(S1)=-1, F(S2)=1, F(S3)= -1, F(S4)=-1.

Figure 5 presents the identification pattern visually.

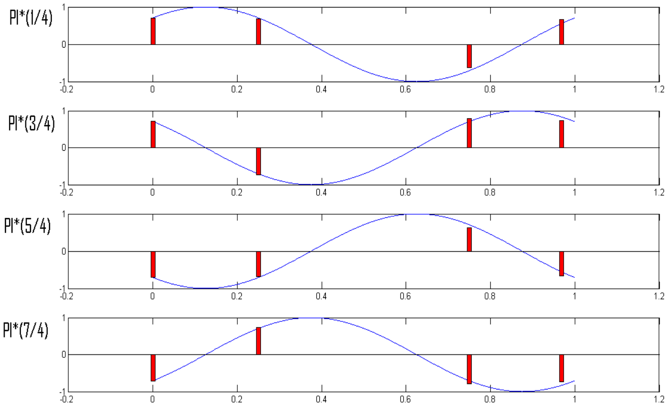


**Fig. 5.** The four optimum samples the detection algorithm relies on

The algorithm assumes that noise does not affect the polarity of the four samples; otherwise the longest match is considered the pattern of choice, and error in detection may occur.

# 3 Simulation Results

## 3.1 Description

The 4SQD algorithm was tested and simulated using ANSI C++. The Integrated Development Environment of choice is Microsoft VC++.NET 2010 Express Edition. Simulation software consists mainly of two classes. The first class is called Signal. It represents a generic class for dealing with digital signals. It stores the digitized signal in a buffer and applies basic signal processing operations such as addition, subtraction, multiplication and noise addition. The second class is the QPSK Signal class, which is inherited form the Signal class. The QPSK Signal class is designed to produce a QPSK signal with a specific phase and frequency.

Simulation Procedure:

The simulation procedure begins when the program generates a random number with a value of 0, 1, 2, or 3. Then a QPSK pattern is generated and stored in a buffer according to the random number generated. The next step is the addition of phase distortion. The last step in signal generation is noise addition. Additive White Gaussian Noise is added to the signal to simulate amplitude noise that a transmission channel may have.

The moment the QPSK signal is generated the detection stage starts. In this simulation the signal is already stored in a buffer, for which reason the analog-to-digital stage is bypassed. The subsequent step is sample selection. Finally, pattern identification and bit error rate calculation are done. The complete simulation process flow is shown in Figure 6.

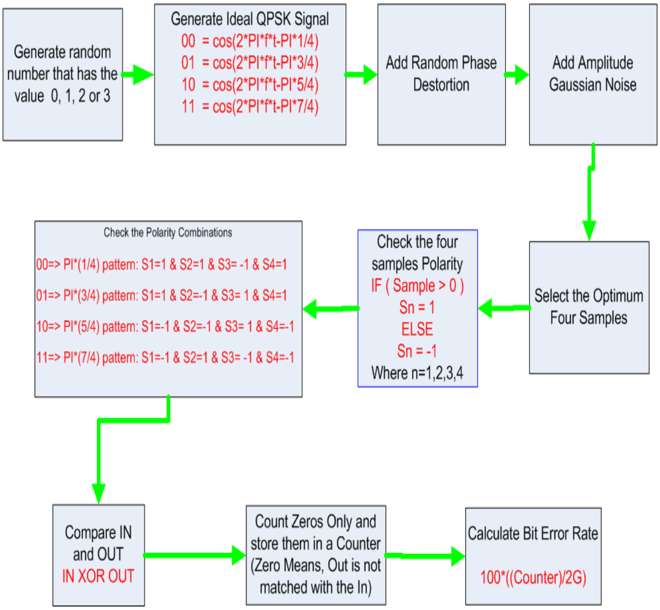


**Fig. 6.** Block diagram of the simulation steps

The simulation procedure begins with creating a QPSK signal, followed by adding a phase shift and white Gaussian noise, obtaining the optimum samples' positions and finally, identifying the QPSK pattern.

The *QPSKSignal* class is responsible for creating a QPSK signal with the desired frequency, phase shift and number of samples, and storing it in a buffer. The following line of code will create a QPSK signal of 1 GHz frequency, phase shift of PI/4 and 32 samples per symbol.

QPSKSignal * MySymbol = new QPSKSignal (PI/4.0, 1000 000 000, 32). To access MySymbol samples the following line of code is used: MySymbol -> SignalData [I], where I can be any integer less than the number of samples per symbol, which is 32 in this case.

The QPSK signal is already generated with a known phase shift, but adding a phase shift to the original phase shift may be needed to simulate noise[7-10]. To solve this issue the *ShiftPahse* method is used in the following line of code:

MySymbol -> ShiftPhase (PhaseShiftAmount)

To simulate a noisy channel we apply the *GaussianNoise* method as follows:

MySymbol -> GaussianNoise (NoisePowerdB)

As mentioned in the methodology section, our decision will be based on the polarity of four samples at specific positions. The *FindQPSKOptimumSamples* static method is used to determine the four optimum samples. The function takes the signal frequency and ADC frequency and returns indexes of sample1, sample2, sample3, and sample4 in S1, S2, S3 and S4 respectively. Notice that S1, S2, S3 and S4 are sent by reference.

QPSKSignal::findQPSKOptimumSamples(Frequency,ADC_Frequency&S1&S2&S3&S4)

This function of the simulation process is the core of the detection algorithm and cannot be modified, unlike other functions that can be modified to serve the system requirements and needs. The detectQPSKSignal method takes as its input the polarities of the four samples and returns the detected pattern as follows:

QPSKSignal::detectQPSKSignal(F1, F2, F3, F4)

Where:

F1=1 if MySYmbol.SignalData[S1]>=0 , F1=-1 if MySYmbol.SignalData[S1]<0
F2=1 if MySYmbol.SignalData[S2]>=0 , F2=-1 if MySYmbol.SignalData[S2]<0
F3=1 if MySYmbol.SignalData[S3]>=0 , F3=-1 if MySYmbol.SignalData[S3]<0
F4=1 if MySYmbol.SignalData[S4]>=0 , F4=-1 if MySYmbol.SignalData[S4]<0

## 3.2 Results

After simulating 2 GB of random data at different signal-to-noise ratios we obtained the results shown in Table 1.

**Table 1.** SQD simulation results using different SNR

| SNR dB | 5.9 | 6.4 | 6.9 | 7.6 | 8.3 | 9.1 | 10.2 | 11.5 | 13.3 | 16.3 |
|---|---|---|---|---|---|---|---|---|---|---|
| Error% | 49.9 | 25 | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

It is clear from Table 1 that if the signal-to-noise ratio is greater than or equal to 7.57 dB, then the detection algorithm capability is significant and the bit error rate becomes zero. However, if the signal-to-noise ratio is less than the threshold (i.e. 7.57 dB) the algorithm is still able to detect the patterns but with relatively high bit error rates compared to other algorithms. The proposed algorithm has proven high levels of efficiency with the least number of samples. The algorithm has also proven its immunity to noise since the detection decision is not based on the samples' values rather than the samples' polarities. In other words, whatever the noise level if the four samples' polarities are not affected the algorithm can still detect the pattern efficiently.

## 4 Conclusion

A QPSK signal can be detected using digital techniques. In this work a novel algorithm was proposed for this purpose. The 4SQD algorithm is able to detect QPSK patterns efficiently if the signal-to-noise ratio is above 7.57 dB and the Phase Shift Error does not exceed PI/8. If the implementation does not satisfy these two conditions, the algorithm is still capable of detecting, but with less accuracy and higher Bit Error Rate (BER). The algorithm only needs the polarity of four samples at specific positions within the signal element to be able to identify the pattern. The algorithm was simulated and tested using ANSI C++ on Microsoft Visual Studio Express Edition 2010.

It is recommended to test the 4SQD on a real environment using Field Programmable Gate Array (FPGA) and applying different types of amplitude noise to validate the algorithm and show its robustness. It is also recommended to combine the QPSK signal with an AM modulation scheme to increase system throughput. In other words, if 2 or more QPSK signals are AM-modulated at different carrier frequencies so that no QPSK signals intersect with the other, the receiver will be able to detect every QPSK signal envelope separately and apply the 4SQD algorithm on each one separately.

**Acknowledgement.** This work is supported by University of Malaya High Impact Research Grant no vote UM.C/625/HIR/MOHE/SC/13/2 from Ministry of Higher Education Malaysia.

## References

1. Proakis, J.G.: Digital Communication, 4th edn. McGraw Hill, New York (2001)
2. Kubankova, A., Kubanek, D.: Methods of digital modulation recognition and their testing. Latest Trends on Communication (2010)
3. Ho, K.M., Vaz, C., Daut, D.G.: Improved Demodulation of Phase Shift Keyed Signals Using Wavelet Thresholding. IEEE (2008)
4. Barsanti, R.J., Smith, T., Lee, R.: Performance of a Wavelet-Based Receiver for BPSK and QPSK Signals in Additive White Gaussian Noise Channels. IEEE (2007)
5. Sklar, B.: Rayleigh Fading Channel in Mobile Digital Communication System Part 1: Characterization. IEEE Communication Magazine, 90–100 (1997)
6. Theodore Rappaport, S.: Wireless Communication: Principle and Practice, 2nd edn. Pearson Educational International (2002)
7. Nikolaos, S.: Designing and Implementation of a 900MHz Radio QPSK Receiver. University of IOANNINA; Siva Ram Murthy, C., Manoj, B.S. Ad hoc wireless networks: Architectures and protocols. Pearson Education (2005)
8. Saunders, S., Aragón-Zavala, A.: Antennas and propagation for wireless communication systems. John Wiley & Sons (2007)
9. Dardari, D., Luise, M., Falletti, E. (eds.): Satellite and terrestrial radio positioning techniques: a signal processing perspective. Academic Press (2011)
10. Hayes, J., Weinstein, S.B.: Data communications principles. Springer (1992)

# An Ensemble CRT, RVFLN, SVM Method for Estimating Propane Spot Price

Haruna Chiroma[1,4], Sameem Abdul-kareem[1], Abdulsalam Ya'u Gital[2],
Sanah Abdullahi Muaz[1], Adamu I. Abubakar[3], Mungad[1], and Tutut Herawan[1]

[1] Faculty of Computer Science and Information Technology
University of Malaya
50603 Pantai Valley, Kuala Lumpur, Malaysia
[2] Department of Computer Science
Universiti Teknologi Malaysia
Kampus UTM Skudai, Johor Bahru, malaysia
[3] Department of Information Science
International Islamic University Kuala Lumpur, Malaysia
[4] Department of Computer Science,
Federal College of Education (Technical), Gombe, Nigeria
`hchiroma@acm.org, asgital@yahoo.com, 100adamu@gmail.com,`
`samaaz.csc@buk.edu.ng, {sameem,mungad,tutut}@um.edu.my`

**Abstract.** In this paper, we propose an ensemble of the CRT-RVFLN-SVM (Classification and Regression Tree (CRT), Random Variable Functional Link Neural Network (RVFLN), and Support Vector Machine (SVM)) to improve robustness and effectiveness in estimating propane spot price. The propane spot price data which are collected from the Energy Information Administration of the US Department of Energy and Barchart were used to build an ensemble CRT-RVFLN-SVM model for the estimating of propane spot price. For the purpose of evaluation, the constituted intelligent computing technologies of the proposed ensemble methodology in addition to Multilayer Back-Propagation Neural Network (MBPNN) were also applied to estimate the propane spot price. Experimental results show that the proposed ensemble CRT-RVFLN-SVM model has improved the performance of CRT, RVFLN, SVM, and MBPNN. The can help to reduce the level of future uncertainty of the propane spot price. Propane investors can use our model as an alternative investment tool for generating more revenue because accurate estimations of future propane price implies generating more profits.

**Keywords:** Propane Spot Price, CRT, RVFLN, SVM, Estimation.

## 1 Introduction

The propane is an energy product obtained through the refining of crude oil, and further refinement of natural gas, although in small proportion compared to the amount obtained from crude oil. Yet, the propane is considered as a refined product of crude oil, and natural gas [1]. Propane has various uses in the society including its use

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_3

as a quenching gas. Experimental investigation reveals that propane is better than other quench gases such as ethane, isobutene, among others in terms of gains and energy resolutions [2]. Propane is found to be effective in trapping adult mosquitos and sand flies which makes it a suitable substitute for carbon dioxide [3]. The prices of energy products including propane spot prices in the future are uncertain. Therefore, the data on spot prices of propane were collected from barchart (www.barchart.com) and used to build a multilayer backpropagation neural network (MBPNN) model for the prediction of the propane spot price among other energy product prices. It was found the MBPNN prediction model outperformed multilayer regression in prediction accuracy [4]. However, MBPNN has difficulties in determining the required number of hidden layer neurons. In addition, further increases in the number of neurons in the hidden layer increase the complexity of the MBPNN structure, thereby undermine its performance. These limitations, triggered [5] to propose a structure of Random Variable Functional Link Neural Network (RVFLN) which do not use hidden layers, as a result, it eliminates the need for hidden neurons, therefore, the RVFLN outwit the cumbersome task of determining the number of hidden layer neurons. The structure of the network complexity is also reduced. Yet, the RVFLN can solve non-linear functions with a high level of accuracy and in less computational time complexity as experimentally proved in [6]. The RVFLN has the possibility of being stuck in local minima which undermine its robustness in real life applications [7]. In contrast, support vector machine (SVM) proposes by [8] is immune to local minima. In classification and regression tree (CRT), local models are easy to be determined once the tree is built [9]. Thus, the CRT does not require much parameter settings compared to the RVFLN, and SVM.

Therefore, we propose to ensemble these algorithms to create a synergistic computational methodology. Hence, to improve the estimation accuracy of the propane spot price, we ensemble C & R Tree, FLN, and SVM and we code name it as an ensemble CRT-RVFLN-SVM. This methodology as suggested by experimental evidence has efficient performance estimation accuracy.

The rest of this paper is organized as follow. Section 2 presents the theoretical concept of the intelligent algorithms adopted in the methodology and the proposed conceptual framework. In section 3 results and discussion are reported. Finally, the conclusion of this work is described in Section 4.

## 2 Framework and Proposed Method

### 2.1 Theoretical Framework

#### 2.1.1 Random Variable Functional Link Neural Network

The RVFLN was originally proposed by [5] inspired by the theorem of [10-11]. Consider the Definition 1 and Theorem 1given as follow.

**Definition 1.** *For each measurable function* $\chi(\cdot): R^r \rightarrow R$ *and* $r \in N, \sum\prod^r(\chi)$ *be the class of functions,*

$$\left\{\begin{matrix} f : R^r \to R : f(x) = \sum_{j=1}^{q} \beta_j \cdot \prod_{k=1}^{l_j} \chi(A_{jk}(x)), \\ x \in R^r, \beta_j \in R, A_{jk} \in A^r, l_j \in N, q = 1,2,\cdots \end{matrix}\right\},$$

*where $A^r$ is the set of all functions from $R^r$ to R. For the special case of $l_j = 1$, we have the $\sum$ networks.*

**Theorem 1.** *For every continuous non constant function $\chi$ every r and every probability measure μ on $(R^r, \beta^r)$, $\sum\prod^r(x)$ is $\rho\mu$ dense in $M^r$, where $\mu$ is a probability measure taken convenience to describe the relative frequency of occurrence of inputs $\mu$, $\beta^r$ is the bored field of $R^r$ and $M^r$ is the set of all bored measurable functions from $R^r$ to R.*

Theorem 1 theoretically proves that a single hidden layer neuron is sufficient to approximate any arbitrary function regardless of the non-constant function $\chi$ used, dimensionareof thethe environmentors *r,* and environment of the input space $\mu$ [11].

**Theorem 2.** *For every squashing function $\Psi$ every r , and μ on ($R^r$, $\beta^r$ ), $\sum^r(\Psi)$ is uniformly dense on compacta in $C^r$ and $\rho\mu$ dense in $M^r$ (where $C^r$ is the set of continuous functions from $R^r$ to R and $C^r$ is a subset of $M^r$).*

Theorem 2 shows that if a function *(f)* is map from $R^2$ to $R^2$ , then the *(f)* can be arbitrary approximated by Eq. (1)

$$f = \sum_{j=1}^{N} \beta_j \chi(A_j x + b_j). \tag{1}$$

Eq. (1) indicates that it can be implemented as a one hidden layer FFNN with all the weights $A_j$ and $\beta_j$ a BP learning algorithmg the thresholds using BP learning algorithm or it can be implemented with functional link neural network with flat architecture. The RVFLN randomly generate $A_j$ and $b_j$, and learn only $\beta_j$. This creates a flat network architecture through which only weights $\beta_j$ must be learned. Quadratic optimization provides learning and is rapidly performed. For supervise learning of the RVFLN, enhancement of inputs to $x_p$ with elements $(x_{p1}, x_{p2}, \cdots, x_{pj})$ and assign the target output of the pattern $x_p$ is $t_p$ . The algorithm initially assigns random weight $\beta_j$ values. The output $o_p$ is computed linearly as

$$o_p = \sum \beta_j x_{pj}. \tag{2}$$

For each input vector change in weights is computed as

$$\Delta\beta_{pj} = \eta(t_p - o_p)x_{pj}. \tag{3}$$

Changes are computed for each input vector in the training set. The weights are updated after each presentation accordingly based on Eq. (4)

$$\beta_j(k+1) = \beta_j(k) + \sum\nolimits_p \Delta\beta_{pj}. \tag{4}$$

The update continued until it reaches a position in which the values of the weights $\beta_j$ do not significantly change. The decrease of $t_p - o_p$ increases the value of $\eta$ parameter.

### 2.1.2 Support Vector Machine

In SVM, the unknown function is approximated by mapping *x* to higher dimensional space through a function $\phi$. The SVM determines linear maximum – margin hyper plan [8]:

$$f(x) = \omega'\phi(x) + b = 0. \tag{5}$$

$$\frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m} L_\varepsilon(y_i). \tag{6}$$

Eq. 5 is the hyper plane of SVM, where $\omega$ and $b$ are estimated by minimizing Eq.2. $\|\omega\| =$ Regularize term

$$\sum_{i=1}^{m} L_\varepsilon(y_i). \tag{7}$$

Eq. 7 is empirical error and *C* = regularization constant and *C* > 0. The SVM penalizes if it deviae from $y_i$ by means of an $\varepsilon$-insensitive loss function

$$L_\varepsilon(y_i) = \begin{cases} 0 \quad if \left| f(x_i) - y_i < \varepsilon \right| \\ \left| f(x_i) - y \right| - \varepsilon \quad otherwise \end{cases} \quad . \tag{8}$$

$$y_i - (\omega'\phi(X_i) + b) \le \varepsilon + \xi_i^- \ , \ \forall_i \tag{9}$$

$$(\omega'\phi(X_i) + b) - y_i \le \varepsilon + \xi_i^+ \ . \tag{10}$$

$$\xi_i^- \ge 0, \xi_i^+ \ge 0 \ . \tag{11}$$

Introducing $\xi_i^-$ and $\xi_i^+$ as slack variables for minimizing Eq. 6

$$\min_{\omega,b,\xi_i^-,\xi_i^+} \frac{1}{2}\|\omega\| + C\sum_{i=1}^{n}(\xi_i^- + \xi_i^+)\ . \tag{12}$$

Subject to Eqs. 5 – 11, Eq. 13 is a solution to the minimization problem

$$f(x) = \sum_{i=1}^{m}(\lambda_i - \lambda_i^*)K(X_i, X) + b. \tag{13}$$

$$y_i - (\omega'\phi(X_i) + b) \le \varepsilon + \xi_i^-. \tag{14}$$

$$(\omega'\phi(X_i) + b) - y_i) \le \varepsilon + \xi_i^+. \tag{15}$$

where $\lambda_i$ and $\lambda_i^*$ are Lagrange multipliers related to constrains Eqs. 14 – 15 [12].

Radial basis function, Sigmoid function, Polynomial function, and Linear function. are widely use kernel functions in literature for computation in the hidden layer of the SVM.

### 2.1.3 Classification and Regression Tree

The CRT can be used for classification as well as for regression. In this study, the CRT we adopted is the regression tree since our focus is on regression problem. The regression tree typically starts with a single node comprising of the entire points. The algorithm stops when all input variables have the same value. On the contrary, the binary splits of the variables are searched to find the variable that reduces error as much as possible. The search continues until the error is lower than a certain threshold otherwise takes a split to create another two nodes. In every new node start the process from the beginning. At every stage the tree uses a single input variable whereas for multiple good features, selection is done on the basis of chance. Training and test data sets are required for the successful building of the CRT model [13].

## 2.2 Datasets and Conceptual Framework

The propane daily spot prices in dollars/gallon were collected from 9 July, 1992 to 16 October, 2012 sourced from the Energy Information Administration (EIA) of the US Department of Energy. The data were freely available, published by EIA of the US Department of Energy. The observations of the propane spot price are 5090 comprised of independent variables including the closing value of daily propane spot price, percentage change of the daily closing spot price based on the previous day, standard deviation (SD) for over five (5) trading days (previous days), and SD for over twenty one (21) trading days (previous days) as suggested in [4]. The independent variables were collected from www.Barchart.com. The descriptive statistics show that the minimum, maximum, mean, and SD of the propane spot price are 0.20, 1.98, 0.6969 and 0.40872 (indicate the statistical dispassion of the propane spot price), respectively. In summary, the relational database has a total of five columns and 5090 rows. The data in our research were not normalized as argued in [14] that normalization can destroy the original pattern in the historical data. The reason why the data was collected between 1992 – 2012 is because of the availability of the data which limit the collection period. Typically, collection period are determined by data availability [15-17].

The data for our research were partition into 50% for training and 50% for testing the performance and generalization ability of the estimation model propose the research. The data were partion into equal proportion (see Fig. 1) to avoid bias that might possibly introduce when the training data is larger than the test dataset.

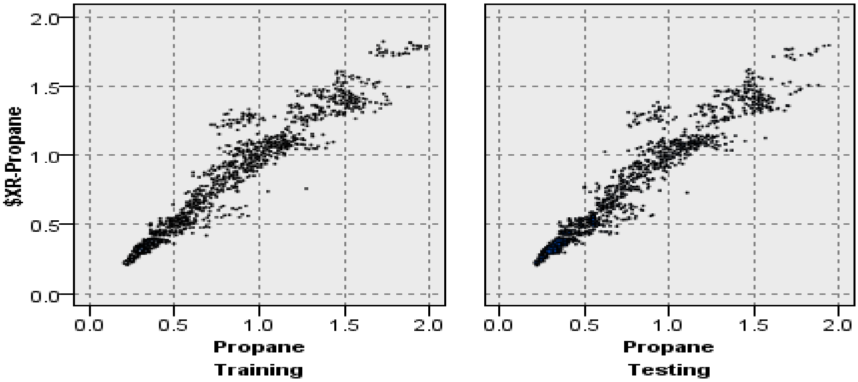


**Fig. 1.** The propane spot price data partition ratio (50% - 50%)

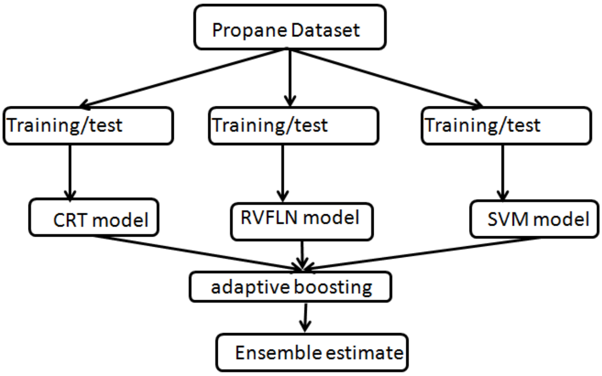


**Fig. 2.** The propose framework of the ensemble CRT-RVFLN-SVM

The intelligent computing technologies (CRT, RVFLN, and SVM) were ensemble through adaptive boosting algorithm. Each of models is then applied for the estimation. Subsequently, ensemble the results to boost the accuracy of the propane spot price estimation. The entire process is depicted in Fig. 2. in which the major components can be summarized as: Data collection and preprocessing, modeling and ensemble.

For comparison, CRT, RVFLN, SVM, and MBPNN were applied to estimate the spot price of propane. The results were compared with the accuracy of the ensemble CRT-RVFLN-SVM to evaluate the effectiveness of the methodology. The parameter settings of the RVFLN, CRT and SVM were realize through experimentation including learning rate, momentum, µ, C, support vectors (SVs), kernel function etc.

## 3 Results and Discussion

The optimal parameters of the propose ensemble CRT-RVFLN-SVM and another compared methods are reported in Table 1. These are the parameters realized after several trials of the experiments in building the models. The simulation experiments were conducted on HP L1750 model, 4Gb RAM, 232.4 GB HDD, 32- bit OS, Intel (R) Core (TM)2 Duo CPU @ 3.00 GHz.

**Table 1.** Optimal parameters of the ensemble CRT-RVFLN-SVM, SVM, RVFLN, CRT, and MBPNN

| Parameter | Configuration |
|---|---|
| Learning rate | 0.2 |
| Momentum | 0.5 |
| MBPNN Structure | 4-7-1 |
| RVFNN Structure | 2th Order |
| µ | 1.4 |
| C | 15 |
| SVs | 3112 |
| Kernel Function | Sigmoid |
| C & R Tree | Regression |
| Maxi. Iteration | 1000 |

The results generated by the propose ensemble CRT-RVFLN-SVM is compared with the results of CRT, SVM, MBPNN, and RVFLN. The simulated comparison results are presented in Table 1.

**Table 2.** Comparing the performances of the intelligent computing technologies with the propose ensemble CRT-RVFLN-SVM as well as showing improvement made by the ensemble CRT-RVFLN-SVM over the results of the constituted algorithms

| Methodology | MAE | $R^2$ | Impr. MAE (%) | Impr. $R^2$ (%) |
|---|---|---|---|---|
| CRT | 0.051 | 0.76 | 99.63 | 21.1 |
| SVM | 0.068 | 0.79 | 99.72 | 16.46 |
| RVFLN | 0.040 | 0.69 | 99.52 | 33.33 |
| MBPNN | 0.059 | 0.72 | 99.68 | 27.78 |
| Ensemble CRT-RVFLN-SVM | 0.00019 | 0.92 | N/A | N/A |

The "N/A" in Table 2 indicated that the computation of the % improvement is not applicable to the proposed method.

The estimations performances suggested from the results obtained by the ensemble CRT-FLN-SVM for the propane spot price datasets outperforms MBPNN, CRT, RVFLN, and SVM in terms of Mean Absolute Error (MAE) and Regression ($R^2$). Table 2 clearly verified that the propose ensemble CRT-FLN-SVM methodology has produced more accurate results compared to the other methods. The fourth (4th) column of Table 2 shows the percentage of improvement achieved by the propose methodology in terms of MAE computed using Eq. (16) and the average improvement is 99.64% (the average is computed by summing all Impro. MAE (%) values divid by 4).

$$PI_{MAE} = \frac{MAE_i - MAE_{ECRS}}{MAE_i} \times 100 \tag{16}$$

$$PI_R = \frac{R_{ECRS}^2 - R_i^2}{R_i^2} \times 100 \tag{17}$$

where $PI_{MAE}$, $PI_{MAE}$, and $MAE_{ECRS}$ are percentage improvement of *MAE* by the ensemble CRT-FLN-SVM, percentage improvement of $R^2$ by the ensemble CRT-FLN-SVM and *MAE* of the ensemble CRT-FLN-SVM. $MAE_i$ and $R_i^2$ (*i*= CRT, SVM, and MBPNN, RVFLN). Here, $R_{ECRS}^2$ is $R^2$ of the ensemble CRT-FLN-SVM. The fifth ($5^{th}$) column of the Table indicated improvement based on $R^2$ computed using Eq. (17) and the average improvement is 24.67% (the average is computed by summing all the impr. $R^2$ values divid by 4). Therefore, the propose ensemble CRT-RVFLN-SVM shows performance advancement over the state of the art methods. Thus, the ensemble CRT-RVFLN-SVM can be considered to be effective and robust. The performance of the ensemble CRT-FLN-SVM can likely be caused by the capability of the methodology to eliminate the possibility of being stuck in local minima, and the reduction of its structural complexity. The results generated by the ensemble CRT-RVFLN-SVM can reduce the level of propane future spot price uncertainty with an acceptable level of accuracy. In this way, investors can use our propose model as an advisory tool for the investment of propane which can yield profits since accurate prediction implies more generation of revenue.

## 4 Conclusion

In this paper, we have proposed an ensemble of CRT, RVFLN, and SVM to eliminate their weaknesses and capitalized on their strengths to improve performance of the state of the art methodology of estimating the propane spot price. For the purpose of performance evaluation of our proposal, we applied CRT, RVFLN, SVM, and MBPNN to estimate the spot price of propane. The performance of the ensemble CRT-RVFLN-SVM proposed in this research was compared to that of CRT, RVFLN, SVM, and MBPNN. Simulation results show that there is a performance improvement over the state of the art methodologies. The level of uncertainty of the future spot price of the propane can be reduced using the model we proposed in this research. Propane investors can use our model as an alternative investment tool for generating more revenue because accurate estimations of future propane price implies generating more profits.

**Acknowledgement.** This research is supported by High Impact Research grant University of Malaya no vote UM.C/628/HIR/MOHE/SC/042 from ministry of Higher Education Malaysia.

## References

1. US Department of Energy: Annual Energy Outlook 2004 with Projections to 2025 (2004), `http://www.eia.doe.gov/oiaf/archive/aeo04/index.html`
2. Agrawal, P.C., Ramsey, B.D.: Use of propane as a quench gas in argon-filled proportional counters and comparison with other quench gases. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 273(1), 331–337 (1988)
3. Kline, D.L., Müller, G.C., Jerome, A.: Hogsette, Evaluation of propane combustion traps for the collection of Phlebotomus papatasi (Scopoli) in southern Israel. J. Vector Ecol. 36(suppl. 1), 166–171 (2011)
4. Malliaris, M.E., Malliaris, S.G.: Forecasting energy product prices. Eur. J. Financ. 14(6), 453–468 (2008)
5. Pao, Y.-H., Phillips, S.M., Sobajic, D.J.: Neural-net computing and intelligent control systems. Int. J. Control 56(2), 263–289 (1992)
6. Hassim, Y.M.M., Ghazali, R.: A Modified Artificial Bee Colony Optimization for Functional Link Neural Network Training. In: Herawan, T., Deris, M.M., Abawajy, J. (eds.) Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013). LNEE, vol. 285, pp. 69–78. Springer, Heidelberg (2014)
7. Sexton, R.S., Gupta, J.N.D.: Comparative evaluation of genetic algorithm and backpropagation for training neural. Inf. Sci. 129, 45–59 (2000)
8. Cortes, C., Vapnik, V.: Support vector networks. Mach. Learn. 20, 273–297 (1995)

9. Garg, M., Singh, U.: C & R Tree based Air Target Classification Using Kinematics. Int. J. Commun. Comp. Tech. 3(1), 15–21 (2012)
10. Hornik, K., et al.: Multi-layer feed-forward networks are universal approximators. Neural Netw. 2, 359–366 (1998)
11. Hornik, K.: Approximation capabilities of multilayer feedforward networks. Neural Netw. 4, 251–257 (1991)
12. Tipping, M.E.: Sparse Bayesian learning and the relevance vector machine. J. Mach. Learn. Res. 1, 211–244 (2001)
13. Richard, B.A.: statistical learning from a regression perspective. Springer Series in Statistics. Springer, New York (2008)
14. Jammazi, R., Aloui, C.: Crude oil forecasting: experimental evidence from wavelet decomposition and neural network modeling. Energy Econ. 34, 828–841 (2012)
15. Ekonomou, L.: Greek long-term energy consumption prediction using artificial neural networks. Energy 35(2), 512–517 (2010)
16. Chiroma, H., Abdulkareem, S., Abubakar, A., Mohammed, J.U.: Computational intelligence techniques with application to crude oil price projection: a literature survey from 2001-2012. Neural Network World 23(6), 523–551 (2013)
17. Abubakar, A., Chiroma, H., Zeki, A., Uddin, M.: Utilising key climate element variability for the prediction of future climate change using a support vector machine Model. International Journal of Global Warming (in Press, 2014)

# Hybrid Artificial Neural Network with Artificial Bee Colony Algorithm for Crime Classification

Syahid Anuar, Ali Selamat, and Roselina Sallehuddin

Faculty of Computing, Universiti Teknologi Malaysia, Malaysia
syah2105@yahoo.com, {aselamat,roselina}@utm.my

**Abstract.** Crime prevention is an important roles in police system for any country. Crime classification is one of the components in crime prevention. In this study, we proposed a hybrid crime classification model by combining Artificial Neural Network (ANN) and Artificial Bee Colony (ABC) algorithm (codename ANN-ABC). The idea is by using ABC as a learning mechanism for ANN to overcome the ANN's local optima problem thus produce more significant results. The ANN-ABC is applied to Communities and Crime dataset to predict 'Crime Categories'. The dataset was collected from UCI machine learning repository. The result of ANN-ABC will be compare with other classification algorithms. The experiment results show that ANN-ABC outperform other algorithms and achieved 86.48% accuracy with average 7% improvement compare to other algorithms.

## 1 Introduction

Increasing volumes of crime had brought serious problems for community in any country. However, the increase in realization of information technology has opened up new doors for government to include crime prevention component as a strategies to reduce crime. Wilpen Gorrs and Richard Harries have introduced the crime forecasting due to the fact that forecasting crime is still new and not widely practiced by police. The main concern of crime forecasting is to study the crime patterns, to analyzing the modus operandi of serial criminals and to allocate useful resources (for tactical purpose) in the places where there are lots of crimes. With accurate forecasts, police would be able to take tactical actions such as targeting patrols to hot spots, conducting surveillance for deployment of special units, scheduling vacations and training for police officer and making crime alerts available to neighborhood watch groups [2].

Classification is one of the data mining technique which been use to analyze crime patterns. The data mining approach can help to detect the crimes patterns and speed up the process of solving crime [11]. Because of this, research on crime classification has increased because of the potential and effectiveness of classification in crime prevention programs. Some researches in crime classification have been done by several researchers [4,5,10]. Crime can be divided into

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems*,
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_4

several types and the most common findings at the city level are crimes against property (burglary, robbery and theft) and crime of aggression such as assaults, homicides and rape [3]. For every crime's types, the dataset can be classify into several categories such as low, medium and high. Thus, the objective of this paper is to propose a new classifier to classify the crime categories.

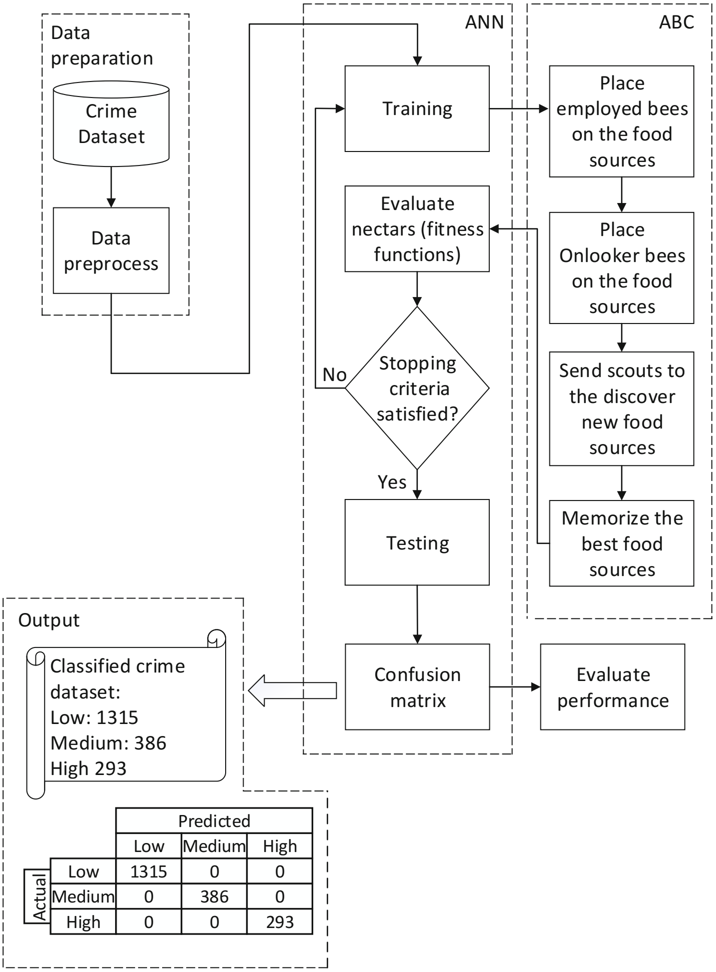
The artificial neural networks (ANN) is most popular computational model and usually use as a benchmark model for any classifier. Learning algorithm is an important aspect that can influence ANN model in producing better results [12]. The ANN with back-propagation (BP) algorithm is a common learning algorithm to minimize mean-squared error between the desired outputs and the actual outputs for the particular inputs to the networks [13]. However, BP has some limitations. First, the result may converge to local optima, meaning that the convergence is not guaranteed [9]. Second, the convergence of back propagation may require many iterations. The shortcomings of BP have resulted on slow learning process or sometimes produce less significance results [12, 13].

To improve the accuracy and the convergence rate of BP algorithm, a new hybrid model called as ANN-ABC is proposed. The goal of the hybridization of artificial bee colony (ABC) and back-propagation algorithm is to optimize the ANN training by avoiding the local minima solution. To validate the performance of the proposed hybrid model, ANN-ABC will be tested on crime classification problem with dataset obtained from UCI machine learning repository [1]. This data has also been used by Iqbal et al. [4] to classify crime using decision tree and Naive Bayes. This work employ machine learning technique, in contrast with work done by Iqbal et al. which was using statistical techniques. The performance of the ANN-ABC algorithm is compared with ANN using standard BP algorithm, Naive Bayes and decision tree algorithm

The rest of the paper is organize as follows. Section 2 discusses the proposed crime classification model, containing the implementation of ABC to train ANN. In section 3, the experiments and results produced by ANN-ABC, ANN, decision tree and Naive Bayes are presented and discussed. In section 4, the statistical test is presented to validate the experiment results.

## 2 Proposed Crime Classification Model

The implementations of ABC algorithm is relatively simple but being able to produce a very good result [8] and has the advantage of not requiring a lot of parameters to be tuned [7]. Figure 1 illustrate the overall process involved in implementation of the proposed hybrid model. Detailed explanation on each process will be described in the next subsection.



| | | Predicted | | |
|---|---|---|---|---|
| | | Low | Medium | High |
| Actual | Low | 1315 | 0 | 0 |
| | Medium | 0 | 386 | 0 |
| | High | 0 | 0 | 293 |

**Fig. 1.** Implementation of Proposed hybrid model ANN-ABC

## 2.1 Data Preparation

The Communities and Crime data set is obtained from the UCI Machine Learning Repository. This dataset focuses on communities in United States of America (USA). The data comprises of socio-economic data from the 90 Census, law enforcement data from the 1990 Law Enforcement Management and Admin Stats survey and crime data from the 1995 FBI UCR. The dataset consists of 128 number of attributes and 1994 number of instances with missing values [1].

Data preparation is essential for successful data classification. Poor quality data typically result in incorrect and unreliable data classification results. Thus, the following data preparation mechanisms were carried out to obtain the final set of attribute in an appropriate form for further analysis and processing. The data preparation is based on works done by Iqbal et. al [4].

- All the attributes with large number of missing values were removed
- The newly added nominal attribute named Categories is created based on attribute named 'ViolentCrimesPerPop' - if the value is less than 25% than the Categories is Low. If the value is equal to or greater than 25% than the Categories is Medium. If the value is equal to or greater than 40% than the Categories is High. The final count of crime categories are low: 1315, medium: 386 and high: 293.
- All attributes are set to numeric except Categories which is set as nominal.
- The final number of attributes after data preparation are implemented is 104. This number of attribute will undergo feature selection process.
- All the data will be normalize into [0, 1] using min-max method by using Eq. (1), where $X_i$ is the normalized input data and $x_i$ is the unnormalized input data during $i$th iteration

$$X_i = \frac{x_i - min(x_i)}{max(x_i) - min(x_i)} \tag{1}$$

The dataset is divided into training and testing using 10-fold cross validation method. In this method, for each of the 10 times, 9 portion of the dataset is divided into training set and a portion of dataset is divided into testing set. The training set is used to train the algorithm for good learning capability while the testing set is applied to evaluate the generalization capability of the proposed algorithm.

### 2.2 Artificial Neural Network

Artificial neural network (ANN) consists of a set of processing elements known as neurons or nodes and connected with each other. Each node is multiplied with separate weight value and receive signal from the nodes in the previous layer. Output of the $i$th neuron can be described by (Eq. 1)

$$y_i = f_i(\sum_{j=1}^{n} w_{ij}x_j + \theta_i) \tag{2}$$

where $y_i$ is the output of the node, $x_j$ is the $j$th input to the node, $w_{ij}$ is the connection weight between the node and input $x_j$, $\theta_i$ is the threshold (or bias) of the node, and $f_i$ is the node transfer function. The goal is to minimize the mean square error (MSE) function given by (Eq. 2)

$$E(w(t)) = \frac{1}{n}\sum_{j=1}^{n}(d_k - o_k)^2 \tag{3}$$

where $E(w(t))$ is the error at the $t$th iteration; $w(t)$, the weight in the connections at the $t$th iteration; $d_k$ is the desired output node; $0_k$ is the target value for the $k$th output node.

For crime classification, the output $(y_1, ..., y_3)$, consist of three crime categories which are low, medium and high. The outputs are represent as three bit binary, for example, {1,0,0} as low, {0,1,0} as medium and {0,0,1} as high.

### 2.3 Artificial Bee Colony Algorithm

Artificial Bee Colony (ABC) algorithm consists of three groups of bees which are employed bees, onlooker bees and scouts bees. Employed bees go to the food sources and come back to hive and dance on this area. Onlookers bees watch the dances and choose food sources depending on the dances. The employed bee whose food sources has been abandoned becomes scout and starts searching a new food source. Algorithm 1 describe the detail implementation of ABC algorithm.

The position of the food source represents a possible solution to the optimization problem and the nectar amount of a food source corresponds to the quality (fitness) of the associated solution. The number of employed bees or the onlooker bees is equal to the number of solutions in the population. For the first step, the ABC generates a randomly distributed initial population $P(C = 0)$ of $SN$ solutions (food source positions), where $SN$ represents the size of employed bees or onlooker bees. Each solution $x_i (i = 1, 2, ..., SN)$ is a $D$-dimensional vector where $D$ is the number of parameters to be optimize. The population of the positions (search process of the employed bees, onlooker bees and the scout bees) is repeated until reach maximum cycle number (MCN), $C = 1, 2, ...MCN$.

**Algorithm 1.** Artificial bee colony algorithm [6]

```
Initialize the population of solutions $x_i, i = 1...SN$
Evaluate the population
cycle = 1
repeat
    Produce new solution $v_i$ for the employed bees by using (Eq. 4) and
    evaluate them
    Apply greedy selection process
    Calculate the probability values $p_i$ for the solutions $x_i$ by (Eq. 3)
    Produce the new solutions $v_i$ for the onlookers from the solutions $x_i$
    selected depending on $p_i$ and evaluate them
    Apply greedy selection process
    Determine the abandoned solution for the scout, if exists, and replace
    it with a new randomly produce solution $x_i$ by (Eq. 5)
    Memorize the best solution achieved so far
    cycle = cycle+1
until Until cycle = MCN;
```

An employed bee produces a modification on the position using (Eq 3). If the nectar amount of the new position is higher than previous, the bee memorizes the new position and discard the old one. Otherwise, the bee keeps the position of the previous on in memory.

$$v_{ij} = x_{ij} + \emptyset_{ij}(x_{ij} - x_{kj}) \tag{4}$$

where $k \in \{1, 2, ..., SN\}$ and $j \in \{1, 2, ..., D\}$ are randomly chosen indexes. $k$ is determined randomly and should be differ from $i$. $\emptyset_{ij}$ us a random generated number between [-1,1].

After all employed bees complete the search process, the sharing information begins where the food sources and their position information shared with the onlooker bees. An onlooker bee evaluates the nectar information and choose a food source with a probability, $p_i$ related to its nectar amount following (Eq. 4).

$$p_i = \frac{fit_i}{\sum\limits_{n=1}^{SN} fit_n} \tag{5}$$

where $fit_i$ is the fitness value of the solution $i$ and $SN$ is the number of food sources. The employed bee produces a modification of the position and checks the nectar amount of the candidate source. If the nectar is higher than previous one, the onlooker bee memorizes the new position and discards the old one. The food source of which the nectar is abandoned by the bees is replaced with a new food source by the scouts by (Eq. 10) in case if the position cannot be improved further. The parameter "$l$imit" is the control parameter to determine the abandonment of the food sources within predetermined number of cycles.

$$x_i^j = x_{min}^j + rand(0,1)(x_{max}^j - x_{min}^j) \tag{6}$$

## 3 Experiment Results and Discussions

The performance of ANN-ABC is compare with ANN using standard BP algorithm. The result from other two algorithm, decision tree and Naive Bayes are taken from the work done by Iqbal et. al [4]. The aim of this comparison is to evaluate the performance capability of the proposed ANN-ABC in crime classification. Or in other words, to investigate whether the application of ABC has successfully optimize ANN and increase the classification accuracy.

Figure 2 shows the experiment result of correctly classified (or accuracy) and incorrectly classified crime dataset. The blue column indicate the dataset is correctly classified and the red column indicate the dataset is incorrectly classified by tested algorithm. The result show that ANN-ABC have correctly classified 86.48% of the dataset outperformed other. ANN itself score 83.49%, decision tree score 83.95% and Naive Bayes score 70.81%. Moreover, ANN-ABC reduced the amount of incorrectly classified and this shows that ANN-ABC has improved in term of accuracy compared with others.

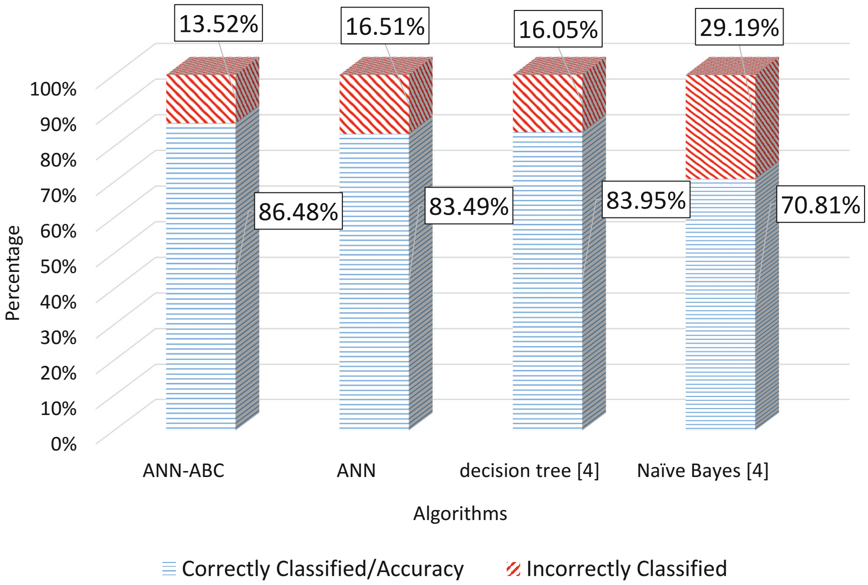


**Fig. 2.** Comparison of correctly classified and incorrectly classified

Figure 3 illustrate the experiment result for classification. The performance measurements consists of precision, recall and F-measure. The result for precision shows that ANN-ABC outperforms the other algorithms with 86.48%. Meanwhile, decision tree is the second algorithm with 83.50% followed by ANN with 83.49%. The last algorithm is Naive Bayes with 66.40%. ANN-ABC shows 8.6% of improvement over other algorithms for precision.

Result for recall performance measurement shows that ANN-ABC outperforms other with 86.60%. The second algorithm is decision tree with 84%. The third algorithm is Naive Bayes followed by ANN both with 70.80% and 69.34% respectively. Result for recall indicates that ANN-ABC 11.9% improvement over other algorithms.

From F-measure results, ANN-ABC is the best classification algorithms for crime classification with 86.53%. The second best algorithm is decision tree with 82.6% followed by ANN with 72.09% and Naive Bayes produces the worst performance with 67.5%. This result indicate 12.5% improvement of ANN-ABC in term of F-measure performance measurement.

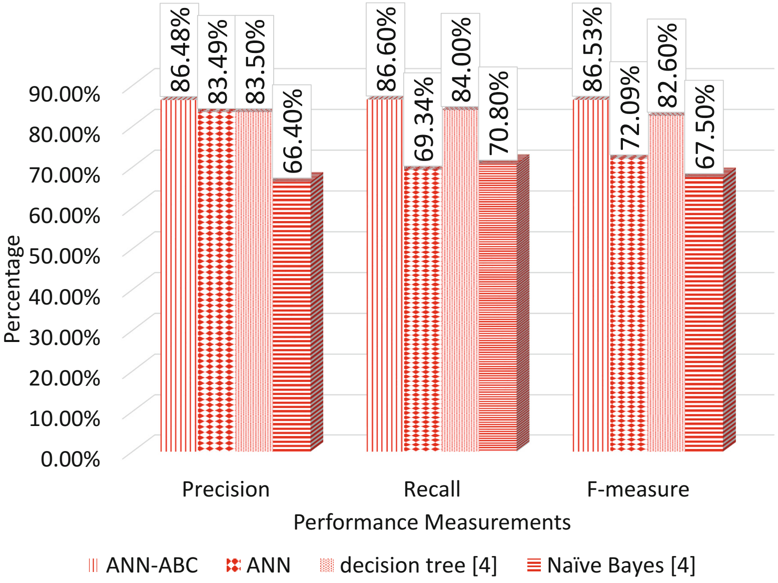


**Fig. 3.** Comparison of performance measurements

The experiment results indicate that the ANN-ABC outperforms other algorithms on classification of crime dataset. The exploration and exploitation process of ABC to optimize ANNs weights has succesfully minimize the learning error for better convergence results. Despite of using the gradient stepest descent that prone to trap in local minima, the ANN-ABC used search strategy that analyze a wider region of search space has efficiently optimize the search strategies

## 4 Statistical Test

The experiment results shown in previous section indicate that ANN-ABC outperformed other algorithms. To further validate whether the result is significant, independent sample t-test are carried out. This test is performs to examine whether the ANN-ABC is performs better the ANN at the level of confidence 0.05. This test is perform on test dataset and the test variable are the performance metrics such as accuracy, precision, recall and F-measure.

The null hypothesis ($H_0$) is define as there is no statistically significant difference between the performance of ANN-ABC and ANN, whereas the alternative hypothesis ($H_1$) has been defined as, the performance of ANN-ABC is better than the ANN. Table 1 shows the result of statistical test.

**Table 1.** Statistical test for ANN-ABC versus ANN

| | df | t | Sig. (2-tailed) |
|---|---|---|---|
| Accuracy | 18 | 12.81 | .000 |
| Presicion | 18 | 2.23 | .039 |
| Recall | 18 | 8.00 | .000 |
| F-measure | 18 | 7.29 | .000 |

By observing Table 1, it can be seen that the $p$-value of the performance metrics are smaller than 0.05. This indicate that there is a significant difference for the performance of both algorithms. Hence, the evidence support the claim on $H_1$ which conclude that the performance of ANN-ABC is better than the ANN in classification of crime dataset.

## 5 Conclusions

This paper present the usage of ABC algorithm in learning ANN for hybrid crime classification model. The hybrid model, named ANN-ABC was compared with other classification algorithm such as ANN with back propagation algorithm, Naive Bayes and decision tree algorithm. The performance of the classification algorithms was tested on real world Communities and Crime dataset, collected from UCI machine learning repository. Overall specific performance measurement for classification such as accuracy, precision, recall and F-measure are compared with each other. The results from the experiment show that ANN-ABC has produce higher accuracy and improve average 7% compare to other algorithms. Furthermore, the statistical test has been conducted and the results show that the performance of ANN-ABC is significantly better than the ANN. Overall results indicate that ANN-ABC is better than the other algorithms in learning and classifying crime dataset. For crime classification purpose, the improvement is important for ANN-ABC to establish reliable classification model thus provide constant result for crime analyst to analyze the crime data. On another point of view, suggestion of using machine learning techniques to analyze crime data efficiently can be highlighted after the success of this experiment. The future direction include the consideration of using feature selection to select the significant attributes of the crime dataset.

**Acknowledgement.** The Ministry of Higher Education Malaysia and Universiti Teknologi Malaysia (UTM) under research grant 00M19 and ERGS Grant Vote number :R.J130000.7828.4L107, are hereby acknowledged for the research funding support and some of the facilities utilized during the course of this research work.

## References

1. Bache, K., Lichman, M.: UCI machine learning repository (2013), `http://archive.ics.uci.edu/ml`
2. Gorr, W., Harries, R.: Introduction to crime forecasting. International Journal of Forecasting 19(4), 551–555 (2003)
3. Gorr, W., Olligschlaeger, A., Thompson, Y.: Assessment of crime forecasting accuracy for deployment of police. International Journal of Forecasting (2000)
4. Iqbal, R., Azmi Murad, M.A., Mustapha, A., Panahy, S., Hassany, P., Khanahmadliravi, N.: An experimental study of classification algorithms for crime prediction. Indian Journal of Science & Technology 6(3) (2013)
5. Janeela Theresa, M., Joseph Raj, V.: Fuzzy based genetic neural networks for the classification of murder cases using trapezoidal and lagrange interpolation membership functions. Applied Soft Computing 13(1), 743–754 (2013)
6. Karaboga, D.: An idea based on honey bee swarm for numerical optimization. Techn. Rep. TR06, Erciyes Univ. Press, Erciyes (2005)
7. Karaboga, D., Akay, B., Ozturk, C.: Artificial bee colony (ABC) optimization algorithm for training feed-forward neural networks. In: Torra, V., Narukawa, Y., Yoshida, Y. (eds.) MDAI 2007. LNCS (LNAI), vol. 4617, pp. 318–329. Springer, Heidelberg (2007)
8. Karaboga, D., Gorkemli, B., Ozturk, C., Karaboga, N.: A comprehensive survey: artificial bee colony (abc) algorithm and applications. Artificial Intelligence Review, 1–37 (2012)
9. Kattan, A., Abdullah, R., Salam, R.A.: Harmony search based supervised training of artificial neural networks. In: 2010 International Conference on Intelligent Systems, Modelling and Simulation (ISMS), pp. 105–110. IEEE (2010)
10. Nasridinov, A., Ihm, S.-Y., Park, Y.-H.: A decision tree-based classification model for crime prediction. In: Park, J.J., Barolli, L., Xhafa, F., Jeong, H.-Y. (eds.) Information Technology Convergence. LNEE, vol. 253, pp. 531–538. Springer, Heidelberg (2012)
11. Nath, S.V.: Crime pattern detection using data mining. In: 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops, WI-IAT 2006 Workshops, pp. 41–44. IEEE (2006)
12. Ozturk, C., Karaboga, D.: Hybrid artificial bee colony algorithm for neural network training. In: 2011 IEEE Congress on Evolutionary Computation (CEC), pp. 84–88. IEEE (2011)
13. Yaghini, M., Khoshraftar, M.M., Fallahi, M.: A hybrid algorithm for artificial neural network training. Engineering Applications of Artificial Intelligence 26(1), 293–301 (2013)

# PSO Based Throughput Optimization for Mobile Sensor Networks

Adnan Fida, Nor Jaidi Tuah, and Trung Dung Ngo

Universiti Brunei Darussalam, Brunei Darussalam
The More Than One Robotics Laboratory
`www.morelab.org`

**Abstract.** This paper addresses the problem of optimal route selection and throughput optimization using heuristic computation in realistic communication environments for mobile sensor networks. The strategy is developed in two phases: communication aware optimization that selects the optimal route by using communication quality as the metric for routing decision; and position aware optimization that identifies the bottleneck link on the selected route, then enables route reconfiguration by relocating the bottleneck involved relay to the optimal position with higher throughput. The optimal position is found using particle swarm optimization and mobility of the relay is governed by a feedback control. We demonstrate through simulations that proposed strategy provides better routes in terms of end-to-end throughput as compared to other routing strategies based on ideal communication models.

## 1 Introduction

Mobile sensor networks (MSNs) consisting of wireless mobile sensor nodes can play significant role in several applications such as rescue operations[1], monitoring systems [2], and habitat monitoring [3]. The sensor nodes automatically organize a multihop ad hoc network for information exchange among the nodes where all the other nodes except source or destinations act as relay[4]. Once the sensor nodes have collected environment exploited data, a major challenge is the selection of optimal route for transmitting that data from source node to peer nodes or base stations through the self-organized ad hoc network.

We focus on the problem of transferring the collected data to peer nodes or base station, and are interested in maximizing the throughput of data transmission routes in presence of Rician fading channels. The conventional routing strategies for MSN are based on simplified communication models neglecting the effects of multipath fading. In a realistic scenario, the transmitted waves are attenuated due to fading and results in performance degradation of communication links. This paper proposes a framework for optimal route selection and throughput optimization in presence of noise, path loss and fading.

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems*,
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_5

## 2 Related Work

MSNs borrow the routing literature available for mobile ad hoc networks. The mainstream route selection schemes in ad hoc networks are based on simplified communication link models. Hop Count (HC) is a widely used route selection metric in ad hoc networks [5,6], based on a binary communication model where nodes can perfectly communicate within a transmission radius. HC aims to find routes with minimum number of hops and is popular because of its simplicity in implementation. Euclidean Distance (ED) is a route selection metric which utilizes the geographic position of nodes and selects routes with the shortest physical distance between source and destination [7]. ED is often based on disk graph communication models where signal strength decays according to distance [8,9,10] disregarding the fact that signal strength can vary at smaller distances due to fading. Lately, there is a growing awareness regarding the integration of realistic communication models into route selection process. A dynamic window concept to record link information and use it in routing decision is proposed in [11]. Sequential Assigment Routing (SAR) presented in [12] incorporates QoS requirements into routing process. We developed a framework for optimal route selection in wireless sensor networks considering Rayleigh fading channels due to NLOS signal propagation [13]. However, Rayleigh fading model is not sufficient to portray many channels found in real world. It is important to consider other models such as Rician fading [14] applicable in scenarios when the wireless link between transmitter and receiver has a direct component in addition to diffused Rayleigh components.

## 3 Communication Link Model

We assume a quasi-static narrowband Rician fading wireless communication link with an additive white Gaussian noise (AWGN) process and large scale path loss exponent $\alpha$ [15].Each transmitted signal reaches the destination via random number of multiple paths with a dominant line of sight (LOS) signal. The Signal-to-Noise-Ratio (SNR) denoted by $\zeta$, is a discrete random process given by $\zeta = \frac{R}{N_o}$. $R$ is the desired signal received power, $N_o$ is the noise power. Due to the difficulty of predicting Rician fading, it is usually modelled as stochastic effect. The probability density function (PDF) and cummulative density function (CDf) of $\zeta$ due to Rician fading is given as [16],

$$f_\zeta(\zeta) = \frac{(1+K)^{-K}}{\overline{\zeta}} \exp\left(\frac{(1-k)\zeta}{\overline{\zeta}}\right) I_o\left(\sqrt{\frac{4K(1+K)\zeta}{\overline{\zeta}}}\right), \tag{1}$$

$$F_\zeta(\zeta) = 1 - Q_m\left(\sqrt{2K}, \sqrt{\frac{2(1+K)\zeta}{\overline{\zeta}}}\right). \tag{2}$$

where $\overline{\zeta}$ is the average of $\zeta$ and $I_o()$ is the zero order modified Bessel function of the first kind. $Q_m(,)$ is the Marcum Q function [17]. $K$ is called Rician factor and is defined as ratio between direct signal power ($K_d^2$) and power of indirect scattered signals ($2\sigma^2$). It is expressed in decibels as

$$K(dB) = 10 \log_{10} \left( \frac{K_d^2}{2\sigma^2} \right). \tag{3}$$

### 3.1 Communication Quality Estimation Using Reception Probability

The quality of wireless link between any two nodes can be determined using instantaneous SNR ($\zeta$). A packet is successfully received if instantaneous SNR ($\zeta$) is above a certain threshold $\zeta_t$. Reception probability is defined as probability of $\zeta$ between two nodes is above $\zeta_t$ i.e. $RRP := \mathbb{P}[\zeta \geq \zeta_t]$. The value of $\zeta_t$ depends upon the modulation and coding scheme [18]. Rician Reception probability (RRP) under Rician fading channel for any generic link between two nodes $i$ and $j$ is calculated using Eq.(2) and given as,

$$\begin{aligned} RRP := \mathbb{P}[\zeta_{i,j} \geq \zeta_t] = 1 - F(\zeta) &= 1 - \left(1 - Q_m\left(\sqrt{2K}, \sqrt{\frac{2(1+K)\zeta_t}{\zeta_{i,j}}}\right)\right) \\ &= Q_m\left(\sqrt{2K}, \sqrt{\frac{2(1+K)\zeta_t}{\zeta_{i,j}}}\right). \end{aligned} \tag{4}$$

## 4 Two Phase Throughput Optimization Strategy

This section describes the proposed two phase optimization strategy comprised of communication aware and position aware optimization. The flow diagram of proposed strategy is shown in Fig.1

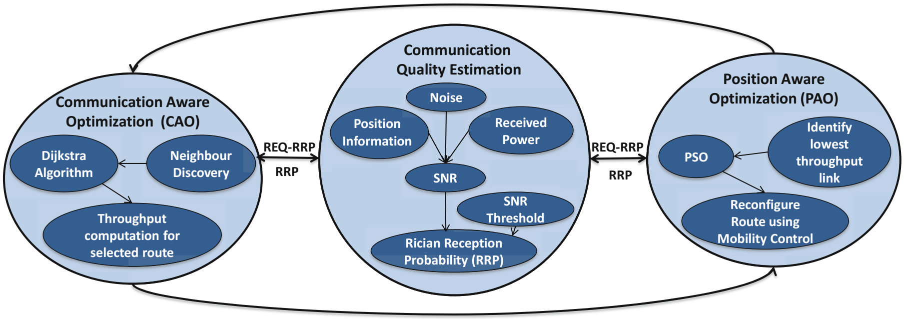


**Fig. 1.** Flow Diagram for Two Phase Throughput Optimization

### 4.1 Communication Aware Optimization(CAO)

The route selection schemes rely on shortest path algorithms such as Dijkstra algorithm [19] for efficient route computation. This phase of the proposed strategy incorporates route selection with knowledge of communication quality including noise, path loss and fading. To include the affects of noise, path loss and rician fading, CAO uses RRP in Eq.4 as route selection metric. As presented in Fig.1, whenever a route is required for data transmission, all the nodes discover their one hop neighbor. The receiving node on each link computes the independent reception probability of its respective link using Eq.4. RRP for each link is independent of other link's quality. Dijkstra algorithm is provided with negative logarithms of all computed RRPs, the negative logarithm is taken to convert the multiplication of probabilities into addition of non-negative weights. Dijkstra algorithm utilizes these weights to find the minimum weight routes which corresponds to routes with maximum RRP. The end-to-end throughput of the route is equal to the minimum of throughput values for links comprising the route. We compare the end-to-end throughput of routes computed using minimum hop count (HC), minimum euclidean distance (ED) and RRP metric.

### 4.2 Position Aware Optimization(PAO)

This phase exploits the available a priori channel and position information of transmitting, receiving nodes to improve the end-to-end throughput of the route selected by CAO. PAO examines the selected route and identifies the bottleneck link with lowest throughput limiting the overall route performance. The route is reconfigured by positioning the relay node with bottleneck link such that link quality and overall throughput is improved. The wireless communication links are unpredictable due to noise and fading, which makes it challenging for iterative methods to find the global optimal positions. Heuristic computation techniques can provide efficient performance in avoiding the bad local optima.

**PSO-Based Optimal Position Search:** Particle Swarm Optimization (PSO) [20] is a heuristic optimization method inspired by observation of animal foraging behavior. The set of potential problem solutions are modeled as swarm moving in a virtual space. Each solution is termed as particle and the group of particles is called population. As the particles search for better positions in the space, they are influenced by their own previous behavior and by the successes of their neighboring particles.

We integrate the RRP in Eq.4 as the fitness function of PSO. The flow diagram of PSO-based optimal position search is presented in Fig.2 The algorithm finds the optimal position in three steps:

- Initialization: random population of 20 particles is generated with 150 number of cycles. Each particle in population begins the search with a random position ($X_{i,j}$) and random velocity ($V_{i,j}$) in n-dimensional search space. $'i'$ represents the particle index and $'j'$ represents the dimension in search

space. Particles are attracted to positions yielding personal the best results and remembers the position of the highest performance;
- Evaluation: at each cycle algorithm reevaluates the previous best position positions and compare the new fitness value based on RRP with previous ones to get a more accurate measure of actual fitness;
- Adaptation: particles are updated by assigning velocity to each element using Eq.(5). The particle achieving the highest performance is termed as global particle. The problem requires smooth changeover from one position to another, so the hard limit for maximum and minimum velocity is replaced with *tanh* function [21].

The equations used by the algorithm are given as,

$$V_{ij} = \underbrace{V_{ij}}_{Inertia} + \underbrace{rand()\alpha_i.(P_{ij} - X_{ij})}_{Personal-Influence} + \underbrace{rand().(pg(i) - X_{ij})}_{Social-Influence}, \qquad (5)$$

$$X_{ij}(n) = X_{ij}(n-1) + \alpha tanh(V_{ij}). \qquad (6)$$

Where $P$ is the personal best and $pg$ is the global best in neighborhood. $rand()$ is a uniformly distributed random number between $[0, 1]$.The inertia part in Eq.5 enables the particle to move in same direction with same velocity; personal influence improves the individual best by moving the particle to its personal best position; social influence helps the particle follow the best neighbor direction. PSO provides the optimal position with improved RRP and throughput. This position is provided as input to feedback mobility control law to move the relay to desired optimal position.

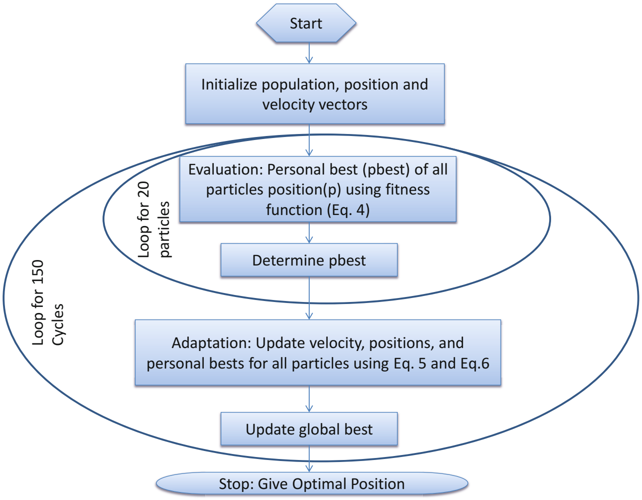


**Fig. 2.** Flow Diagram for PSO-based optimal position search

**Feedback Mobility Control:** In order to move the relay node to optimal position computed by PSO, we need to design a control law such that relay reaches exactly at desired position. The differential equations for direct control of acceleration for the relay are,

$$\dot{x_1} = x_2, \tag{7}$$

$$\dot{x_2} = u. \tag{8}$$

Here, $x_1$, $x_2$ and $u$ represents, position, velocity and control input respectively. State space representation of the above equation is given as:

$$\dot{x} = Ax + Bu \tag{9}$$

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T, B = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$$

$x$ represent the state vector of the relay node, $A$ is system matrix, $B$ is actuator matrix and $u$ is the control law:

$$u = -Kx \tag{10}$$

Since the system has two states, so $K = \begin{bmatrix} k_1 & k_2 \end{bmatrix}$ is $1 \times 2$ row vector. Combining eq.9 and eq.10 results in:

$$\dot{x} = (A - BK)x \tag{11}$$

The appropriate selection of $k_1$ and $k_2$ moves the relay node to optimal position computed by PSO.

## 5 Performance Evaluation

This section exhibits the effectiveness of proposed heuristic based throughput optimization for mobile sensor network in presence of Rician fading scenarios. Throughput is a conventional performance measure for estimating the amount of traffic delivered by the network [10]. We use normalized end-to-end throughput to evaluate the performance of our proposed scheme. Normalized throughput is defined as expected number of successful packet transmission for a given node per time slot [22]. The end-to-end throughput is minimum of throughput values of individual links comprising the route. First, we use Rician Reception Probability (RRP) as route selection metric in realistic Rician fading scenarios and compare the performance with Hop Count (HC) and Euclidean Distance (ED) based conventional route selection schemes. Second, we show how throughput of route selected using RRP can be further optimized using PSO and mobility control (MC). The throughput performance of reconfigured route is compared with original route computed by RRP metric.

### 5.1 Simulation Setup

The sensor network is represented as directed graph $G(N, E)$, where $N$ is the set of nodes and $E$ is the set of links between nodes, called edges. We equate the weight of each edge to its respective RRP. Following assumptions lie beneath the proposed framework. a) Each node is distinguished with a unique identification (ID) b) Each node has the information about relative distance to the neighbors. c) Transmission power is same for all nodes. The values assigned to the parameters are in conformity with real world low power wireless networks [23]. Table.1 describes the fixed values used in simulation.

**Table 1.** Values for parameters used in simulation

| Parameter | Description | Value |
|---|---|---|
| $P$ | Transmit Power | 0 dBm |
| $N_o$ | Noise Variance | -75 dBm |
| $\zeta_t$ | SNR Threshold | 15 dBm |
| $\lambda$ | Wavelength | 0.12 m |
| $\alpha$ | Path loss exponent | 2.5 |

### 5.2 Results from Illustrative Scenarios

**Simulation 1. One-to-One:** Here, we investigate the scenario when a source node requires to share information with a specific far away destination node. Node-1 and node-49 are randomly chosen as source and destination nodes respectively in a network of 100 nodes as shown in Fig.3. Initially, routes are selected using three different metric i.e. HC (magneta), ED(cyan) and RRP(red). The throughput for RRP is further optimized during PAO phase and the bottleneck relay is moved to new position using mobility control, the yellow edges describes the connectivity of relay node at optimal position computed by PSO. Fig.4(a) shows the end-to-end throughput achieved by HC, ED, RRP and after applying PAO. It is inferred that initially RRP based routes achieves an approximate gain of 185% and 40% over HC and ED based routes respectively in this particular case. The PAO phase examines the RRP based routes and identify the relay-39→relay-89 as bottleneck link of the route with RRP of 0.506 and respective throughput of 0.126. An optimal position is computed for relay-89 using PSO such that throughput and RRP is improved. The RRP-based route is reconfigured using feedback mobility control. The initial position of relay-89 in RRP-based route is (32.28, 25.28), after applying PAO phase the relay is positioned at (30.96, 23.51). Thus moving the relay node-89 by a small distance of 221 cm at an angle 0.930 provides a throughput gain of 30% as compared to previous RRP-based route, PSO metric in Fig.4(a) represents the reconfigured route after applying PAO phase.

The HC and ED based schemes selects routes with less number of hops for this case as shown in Fig.4(b). However, in realistic communication scenarios, minimizing the hop count does not necessarily increases the end-to-end throughput.

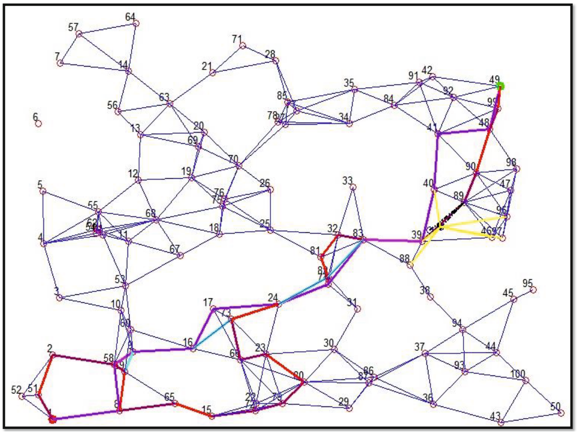
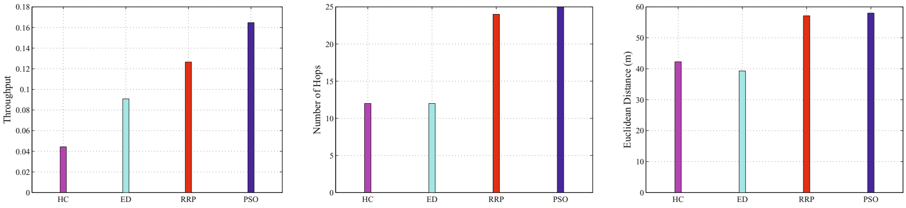


**Fig. 3.** Wireless network of 100 node with source node-1 (red) and destination node-49 (green), routes are selected using RRP (red), ED (cyan), and HC (magneta) metrics



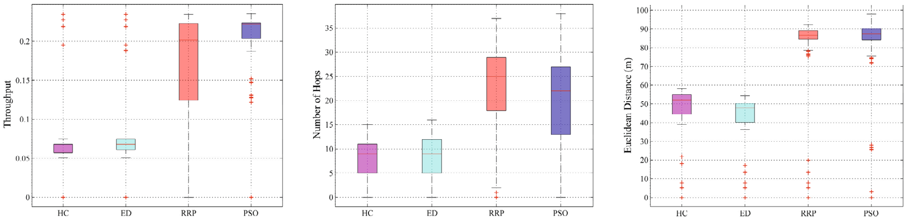
(a) End-to-end Throughput (b) Number of Hops (c) Euclidean Distance

**Fig. 4.** Throughput, Number of Hops and Euclidean Distance for a single source-destination pair

RRP and PSO based routes take extra hops to reach destination, but it regards the stochastic nature of wireless communication links. Whereas, HC and ED based schemes selects the route regardless of quality of communication links. Intuitively, ED based routes travels the smallest physical distance between source and destination as shown in Fig.4(c).

**Simulation 2. One-to-All:** The proposed strategy is further validated by investigating the case when a source node requires sharing information with all other destinations in network. We randomly choose node-1 as source in Fig.3 and 99 routes to all other destinations in network are computed using HC, ED, RRP. Furthermore, RRP based routes are further reconfigured by PAO phase using PSO and mobility control. Fig.5(a) presents the median end-to-end throughput for 99 source-destination pairs. The effectiveness of proposed scheme is established again as it provides a median throughput gain of approximatly 10% as compared to RRP based routes without position aware optimization. We select median instead of mean for statistical data analysis because the data distribution is quite oblique. The difference between edges of each box in Fig.5 presents the inter-quartile range (IQR). IQR is defined as difference between 25th and 75th percentile of data and is used to measure the data dispersion.

Fig.5(b) shows that reconfiguration of RRP-based routes using PAO also minimizes the number of hops. The minimization has a positive impact on the performance of network as it helps in increasing robustness and decreasing delay. The euclidean distance is lowest for ED-based routes and is approximately same for RRP and reconfigured RRP routes as shown in Fig.5(c).



(a) Median Throughput (b) Median Number of Hops (c) Median Euclidean Distance

**Fig. 5.** Throughput, Number of hops, and Euclidean distance for 99 source-destination routes

## 6 Conclusions

Attaining high throughput in mobile sensor networks is crucial for data transmission between far away nodes connected through relay nodes. We proposed two phase throughput optimization strategy to select high throughput routes in presence of path loss, noise and Rician fading. The performance is compared with classic route selection schemes based on simplified communication models. We also show that a priori channel and position information can be exploited to heuristically determine optimal position for relay nodes with considerably better performance.

## References

1. Kibler, S., Raskovic, D.: Coordinated multi-robot exploration of a building for search and rescue situations. In: 2012 44th Southeastern Symposium on System Theory (SSST), pp. 159–163 (2012)
2. Werner-Allen, G., Lorincz, K., Ruiz, M., Marcillo, O., Johnson, J., Lees, J., Welsh, M.: Deploying a wireless sensor network on an active volcano. IEEE Internet Computing 10(2), 18–25 (2006)
3. Mainwaring, A., Culler, D., Polastre, J., Szewczyk, R., Anderson, J.: Wireless sensor networks for habitat monitoring (2002)
4. Lloyd, E.L., Xue, G.: Relay node placement in wireless sensor networks. IEEE Transactions on Computers 56(1), 134–138 (2007)

5. Johnson, D.B., Maltz, D.A.: Dynamic source routing in ad-hoc wireless networks. In: Mobile Computing, pp. 153–181. Kluwer Academic Publishers (1996)
6. Park, V.D., Corson, M.S.: A highly adaptive distributed routing algorithm for mobile wireless networks. In: Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Driving the Information (INFOCOM 1997). IEEE Computer Society (1997)
7. Mauve, M., Widmer, J., Hartenstein, H.: A survey on position-based routing in mobile ad hoc networks. IEEE Network Magazine 15(6), 30–39 (2001)
8. Nmeth, G., Turnyi, Z.R., Valk, A.: Throughput of ideally routed wireless ad hoc networks. ACM Mobile Computing and Communications Review 5(4), 40–46 (2001)
9. Grossglauser, M., Tse, D.: Mobility increases the capacity of ad-hoc wireless networks. In: IEEE Conference on Information Communications (INFOCOM), pp. 477–486 (2001)
10. Gupta, P., Kumar, P.R.: The capacity of wireless networks. IEEE Transaction on Information Theory 46(2), 388–404 (2000)
11. Chen, J., Linemail, R., Liemail, Y., Sunemail, Y.: Lqer: A link quality estimation based routing for wireless sensor networks. Sensors 8(2), 1025–1038 (2008)
12. Sohrabi, K., Gao, J., Ailawadhi, V., Pottie, G.: Protocols for self-organization of a wireless sensor network. IEEE Personal Communications 7(5), 16–27 (2000)
13. Fida, A., Pham, D.H., Tuah, N.J., Ngo, T.: Communication aware route selection in wireless sensor networks. To Appear in Proceedings of the 13th International Conference on Intelligent and Autonomous Systems, IAS 2013, Padova, Italy. Springer (2014)
14. Goldsmith, A.: Wireless Communications. Cambridge University Press (2005)
15. Rappaport, T.S.: Wireless Communications: Principles and Practice, 2nd edn. Prentice Hall (2001)
16. Papoulis, A.: Probability, Random Variables, And Stochastic Processes. McGraw-Hill, New York (1984)
17. Marcum, J.I.: Table of Q Functions. Rand Corporation (1950)
18. Ephremides, A.: Energy concerns in wireless networks. IEEE Transaction on Wireless Communications 9(4), 45–59 (2002)
19. Dijkstra, E.W.: A note on two problems in connexion with graphs. Numerische Mathematik, 269–271 (1959)
20. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: International Conference on Neural Networks, vol. 4, Perth, WA, vol. 4, pp. 1942–1948. IEEE (1995)
21. Zubair, M., Choudry, M.A.S., Naveed, A., Qureshi, I.M.: Particle swarm with soft decision for multiuser detection of synchronous multicarrier cdma. IEICE Transactions on Communications E91-B(5), 1640–1643 (2008)
22. DeCouto, D.S.J., Aguayo, D., Bicket, J., Morris, R.: A high-throughput path metric for multi-hop wireless routing. In: 9th Annual International Conference on Mobile Computing and Networking (MobiCom), MobiCom 2003, San Diego, CA, USA, pp. 134–146. ACM (2003)
23. SIG. Sig bluetooth, specification of the bluetooth system. core version 1.2 (2003)

# Clustering Natural Language Morphemes from EEG Signals Using the Artificial Bee Colony Algorithm

Suriani Sulaiman[1], Saba Ahmed Yahya[1], Nur Sakinah Mohd Shukor[1],
Amelia Ritahani Ismail[1], Qazi Zaahirah[1], Hamwira Yaacob[1],
Abdul Wahab Abdul Rahman[1], and Mariam Adawiah Dzulkifli[2]

[1] Department of Computer Science, Kulliyyah of Information and Communication Technology,
International Islamic University Malaysia
[2] Department of Psychology, Kulliyyah of Islamic Revealed Knowledge and Human Sciences,
International Islamic University Malaysia
`{ssuriani,amelia,hyaacob,abdulwahab,m.adawiah}@iium.edu.my,`
`{aljalalsaba,skinah91}@gmail.com, qazizaahirah@ymail.com`

**Abstract.** We present a preliminary study on the use of a Brain Computer Interface (BCI) device to investigate the feasibility of recognizing patterns of natural language morphemes from EEG signals. This study aims at analyzing EEG signals for the purpose of clustering natural language morphemes using the Artificial Bee Colony (ABC) algorithm. Using as input the features extracted from EEG signals during morphological priming tasks, our experimental results indicate that applying the ABC algorithm on EEG datasets to cluster Malay morphemes produces promising results.

**Keywords:** Clustering,·Artificial Bee Colony algorithm, EEG signals, Natural language morphemes, Morphological priming tasks, BCI.

## 1 Introduction

The importance of the role played by the morphological structure of words in lexical access and representation has been discussed by various studies in psycholinguistics and cognitive neuropsychology with evidence showing that the mental lexicon is organized morphologically to prevent redundant representation of related words [1],[4]. Moreover, results from psycholinguistic experiments commonly show a faster recognition of a word when it is preceded by a morphologically related prime word [2],[6].

In understanding how human brain decomposes words and classify its morphological units (i.e., morphemes), we can analyze the patterns of the EEG signals produced by the brain during language learning tasks with the assistance of supervised (i.e., classification) and unsupervised (i.e., clustering) machine learning techniques. Clustering is one of the most popular unsupervised learning methods aims at discovering a structure in a collection of unlabeled data by organizing data or objects into groups (i.e., clusters) in such a way that each group shares a high degree of similarity while being dissimilar from the data belonging to other clusters. The most commonly used clustering algorithm is the *k*-means algorithm, a center-based algorithm that tends to create clusters of similar sizes. The main disadvantage of this algorithm however, is its tendency to

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_6

converge to the nearest local optimum, leading to the search for more efficient clustering and classification methods such as the swarm intelligence algorithms [1],[3],[9,10],[16,17],[20]. While clustering has been an important tool for applications involving data in high dimensional space such as data mining, image processing, pattern recognition and data compression, this machine learning method has not been widely used in handling complex EEG data.

In this paper, we propose the use of Artificial Bee Colony (ABC) algorithm on EEG datasets to identify and cluster patterns of natural language morphemes. Results from our experiments are then compared with the *k*-means algorithm. The rest of the paper is organized as follows. Section 2 summarizes related work in ABC algorithm, EEG analysis and morphological priming. Section 3 explains the clustering methods applied on our EEG datasets and describe our experimental procedures. In Section 4, we present our experimental results. Section 5 concludes our research findings with suggested improvements and future work.

## 2 Related Work

### 2.1 Morphology and Internal Structure of Words

Words are complex symbols that can have multiple meanings and forms, including related sound patterns and visual structures. A word can be decomposed into separate components that hold its internal structure [5]. In natural language processing, the act of decomposing a word into its morphological constituents (i.e., morphemes), namely roots and affixes, is known as morphological analysis.

Researchers have demonstrated that the morphological structure of complex words plays a role in the way they are represented in the mental lexicon [4,5]. As further noted by [5], complex words recognition in the brain depends on several aspects, in that they may share the same structure (orthography) but not necessarily the same meaning (e.g., CORN/CORNER) or share the same meaning (semantic) but not necessarily the same form (e.g., COPY/IMITATE) or share both structure and meaning (morphology) (e.g., HUNT/HUNTER).

### 2.2 Artificial Bee Colony

The Artificial Bee Colony (ABC) is a swarm intelligence based algorithm for problem solving that simulates the food foraging behavior of swarm honey bees in finding the best food source. The ABC algorithm has recently gained popularity and widely used to solved data clustering [3],[9],[10],[13],[16] and feature selection problems involving complex datasets such as images and EEG signals [17],[20]. This algorithm is favored over similar population-based algorithms namely the Genetic Algorithm (GA), Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) due to its faster convergence speed as well as better performance with fewer control parameters [9,10].

In an actual bee colony, swarm bees amplify the quality difference between nest sites (i.e., food sources) through group communication and decision making in the form of *waggle dances* that serve as signals for the scout bees to attract the swarm towards potential nest sites. 'Lively' *waggle dances* thus represents the quality of a nest site deemed superior. Likewise, the nerve cells in the human brain can produce action potentials (AP) during information processing whose amplitude does not attenuate in the presence of a strong stimulus [15].

### 2.3 Electroencephalogram (EEG) and Event Related Potentials (ERP)

Electroencephalogram (EEG) measures the flow of ionic currents and the voltage fluctuations of neurons within the human brain [21]. EEG signals are commonly analyzed in frequency bands analogous to different mental states. It is however possible to analyze specific brain responses towards certain mental events by recording small potential changes in EEG, immediately after the presentation of a sensory stimulus, a method known as Event-Related Potentials (ERPs).

ERPs are those EEGs that directly measure the electrical response of the cortex in the form of voltage fluctuations, induced within the brain as a sum of numerous action potentials (APs) that are time locked to sensory, motor, or cognitive events [19],[21]. The ERP parameters (e.g., amplitude and latency) are typically generated as a result of peripheral or external stimulations signifying the presence of cognitive processes and dysfunctions not accessible through behavioral testing [19]. The ERP signals can either be positive, such as P300, or negative, such as N400, with digits indicative of the time in milliseconds after a stimulus' presentation. Some researchers focus on N400 [6,12,18], a part of the normal brain response to words and other meaningful stimuli including visual and auditory words, sign languages, pictures and smells.

Priming experiments are effective ways of analyzing the EEG signals for morphologically-related tasks that observe the degree to which a prior presentation of a morphologically related word form (e.g., "walked") promotes the recognition of its target form (e.g. "walk") [18]. The decision making process in the human brain is however not straightforward because of the distraction in the stimulus of relatively the same importance. Due to the limitations on our BCI device and EEG recordings, we have only performed EEG analysis instead of ERP analysis on our morphological priming tasks. We simplify the analysis process by considering the whole EEG recordings, assuming fixed equal segmentations of the time window for each stimulus. Details of the feature extraction process are discussed in Section 3.3.

## 3 Methodology

The basic steps in EEG data analysis are shown in the flowchart in Figure 1.

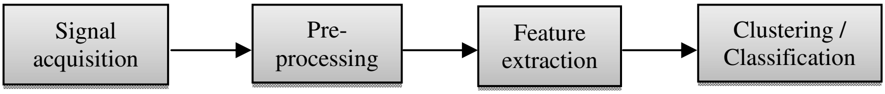


**Fig. 1.** Steps in EEG data analysis

### 3.1 Signal Acquisition

At this stage, the EEG signals were extracted from human subjects during morphological priming tasks. Subjects were required to respond towards six sets of morphologically-related stimuli. EEG signal recording from each subject for each set of stimuli are separately stored in the European Data Format (EDF) format. The EEG signals are collected using the *Emotiv* [7] BCI device while all stimuli are developed and presented to subjects using the *DMDX* [8] software.

**Participants.** Twenty healthy Malay native speakers (10 males and 10 females) within the age range of 21-23 years old participated in this study. All subjects are strongly right handed. Subjects were briefed on the experiment protocol and gave written consents before taking part in the experiment.

**Stimuli.** Our stimuli were grouped under 3 different conditions (i.e., prefix, root and suffix) and divided into 6 training sets (i.e., 2 sets for each condition). Each set consists of 20 randomly ordered Malay affixed words chosen from hand-labelled data. Under each condition, 50% of the stimuli contain invalid affixes. Each affixed word serves as a prime word, succeeded either by a target prefix, a root or a suffix. In each morphological priming task, subjects were required to respond with a correct or an incorrect answer to whether a presented prime word is succeeded by a corresponding target prefix, root or suffix. Each frame containing a word stimulus was displayed at high contrast with white letters on a black background using font size 36 for approximately 1500 ms, followed by a fixation point of the same duration. Subjects were required to respond towards each stimulus within a 2500 ms timeframe. Figure 2 depicts a schematic structure of our stimuli.

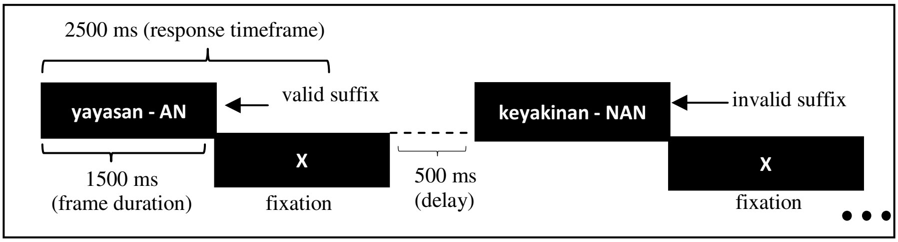


**Fig. 2.** Schematic structure of stimuli in morphological priming (suffix) task

**Procedure.** The experiment was conducted in a quiet room. Subjects were instructed to sit at their own comfort at a distance of about 80 cm from the computer running the stimuli. Each testing session began with a short training block, followed by the experimental block. Participants were briefed that they were expected to view a list of words on the screen and were required to identify and respond whether each word is followed by a valid prefix, root or suffix. Subjects were instructed to respond as quickly and as accurately as possible by pressing one of the dedicated response keys.

*Recording procedure.* The electroencephalogram (EEG) was recorded from the scalp using the *Emotiv* [7] headset with 14 available channels plus 2 reference channels. As illustrated in Figure 3, electrode placements were restricted to six channels; F7, F3, F4, F8, O1 and O2 associated with some of the commonly used channels in morphological priming studies [2],[6],[12],[14]. The EEG signals were continuously recorded at a frequency sampling rate of 128 Hz throughout the experiment.

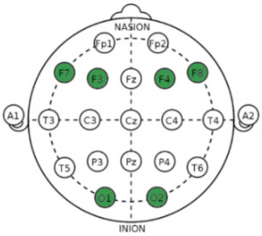


**Fig. 3.** EEG electrodes placement based on the International 10-20 locations

## 3.2 Preprocessing

We normalized and filtered our EEG signals to eliminate noises and irrelevant artefacts. The EEG signals were *band pass* filtered between 0.01-30 Hz. Figure 4 shows an example of pre-processed EEG signals for a single female subject for one morphological priming (suffix) task. Since EEG signals tend to vary across subjects, each EEG recording stated various durations ranging between 60 – 70 s. Each recording is further divided into 20 equal window segments corresponding to each word stimulus.

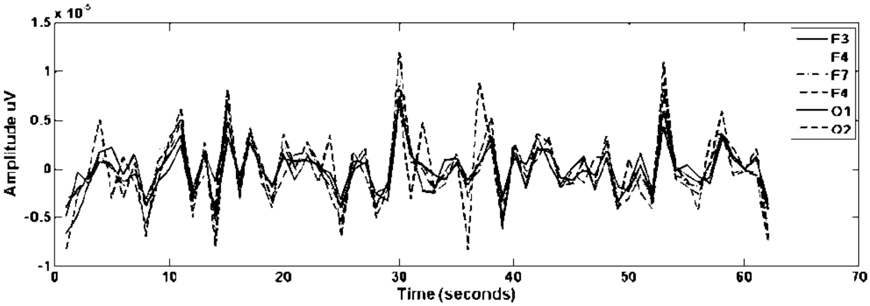


**Fig. 4.** Sample of EEG filtered signals captured from six electrode channels (F3, F4, F7, F8, O1 and O2) for one female subject from the morphological priming (suffix) tasks. The *y* axis represents the amplitude of the signal and the *x* axis represents the time domain for the whole EEG recording. Signals from the six channels were observed to have almost similar fluctuation.

## 3.3 Feature Extraction

The EEG features were calculated based on the mean of discrete time instances corresponding to each word stimulus. For a total of 120 word stimulus distributed across

6 tasks, 20 instances (i.e., amplitude values) are extracted from each channel for each task. The resulting feature vector for each task contains 20 x 6 features corresponding to the number of channels. The size of the feature vector was further reduced into a single value to represent the mean amplitude from all six channels for each word stimulus. Furthermore, the mean reaction time recorded from each subject for each word stimulus served as the second feature of our dataset.

### 3.4 Clustering Algorithms

In the clustering stage, two unsupervised learning methods, the *k*-means and the ABC algorithm were employed, both developed using C++. Data object with similar features will be clustered together based on the minimum calculated *Euclidean* distance between each centroid and data object, computed using Eq. (1),

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^{P} \left(x_{ik} - c_{jk}\right)^2} \tag{1}$$

where $P$ is the total number clusters in the problem, $x_{ik}$ is the *i*th data object belonging to the *k*th cluster and $c_{jk}$ is the center of the *k*th cluster to be found by Eq. (2):

$$c_{jk} = \frac{1}{N_j} \sum_{i=1}^{N} x_i \tag{2}$$

We used the sum of squared *Euclidean* distances between data objects and their closest cluster centers as the objective function and evaluation criterion, $E$, calculated as:

$$E = \sum_{j=1}^{K} \sum_{i=1}^{N} \left\| x_i - c_j \right\|^2 \tag{3}$$

***k*-Means Algorithm.** The *k*-means algorithm is based on the foundation of analysis of variances. It clusters a group of data vectors into a predefined number of *k* clusters. The *k*-means algorithm searches for the local optimal solution in the vicinity of the initial solution and refine the partitioned results to minimize the sum of squared distances between data objects and their closest cluster centers [3]. The clustering process starts with random initialization of cluster centroids (i.e., centroids) to construct initial cluster structures. Data objects are repeatedly assigned to cluster centers based on the distance between each object and the cluster centers.

**K-Means Pseudo code**

```
Initialize cluster centroids $c_i$
REPEAT
    Calculate the Euclidean distance, $dist\_c_i$ between each
    point and each centroid
    Group the data according to the nearest centroids
    Update centroid values
UNTIL threshold is met or no more member movements
```

**Artificial Bee Colony (ABC) Algorithm.** In the ABC algorithm, three groups of bees makes up a colony; 1) the employed bees that find initial food sources, 2) the onlooker bees waiting for information in the dance area to decide on the best food source and 3) the scout bee that randomly searches for new food sources. The location of a food source signifies a possible solution while the nectar amount corresponds to fitness of solution. The honey bees iteratively evaluate the fitness of a food source, memorize their locations and share the information with other bees. We define our cost function $f_i$ as the normalized sum on all Euclidean distances in *N*-dimensional space between each data object $x_i$ and its cluster center $c_j$ as described by [10]:

$$f_i = \frac{1}{D_{sample}} \sum_{i=1}^{D_{sample}} d(x_i, c_j) \tag{4}$$

where $D_{sample}$ represents the number of samples in our dataset. The fitness or quality of our clusters, $fit_i$, can be further calculated as:

$$fit_i = \frac{1}{1 + f_i} \tag{5}$$

In ABC, a food source is selected by an artificial onlooker bee based on the probability value associated with that food source, $p_i$, as a result of the following expression,

$$p_i = \frac{fit_i}{\sum_{n=1}^{SN} fit_n} \tag{6}$$

where $SN$ is the number of food sources (i.e., clusters), equivalent to the number of employed bees while $fit_i$ is the fitness of the solution evaluated by the employed bee. A simplified pseudo code for the ABC algorithm is described as follows:

**ABC Pseudo code**

```
Initialize population size
REPEAT
    Send employed bees to food source positions
    Send onlooker bees to evaluate fitness of neighboring food source
    Memorize best food source
UNTIL maximum cycle is reached or desired error is produced.
```

# 4 Experimental Setup

## 4.1 Datasets

We collected the data from twenty subjects consisting of 120 samples. Each sample is represented by two input features averaged across all subjects; 1) the mean amplitude (Amp) from six EEG channels and 2) the mean reaction time (RT) recorded during the presentation of each stimulus. The dataset are expected to be clustered into four different classes; prefix, root, suffix and neutral (i.e., not belonging to any class).

## 4.2 Algorithm Settings

Table 1 shows the parameter values used for each algorithm in the experiments. In the ABC algorithm, the number of employed bees and onlooker bees is equal to the number of sample size. We limit the execution to a maximum of 5 runs with no significant changes observed with increased number of runs.

**Table 1.** Parameter values used in clustering algorithms

| Algorithm | Parameter | Value |
|---|---|---|
| *k*-means | Sample size, $S$ | 120 |
| | Number of centroids, $c_j$ | 4 |
| | Number of clusters, $k$ | 4 |
| | Max. number of iterations | 500 |
| ABC | Colony size, $n$ | 120 |
| | Number of food sources, $e$ | 4 |
| | Number employed bees, *nep* | 4 |
| | Number of onlooker bees, *nsp* | 116 |
| | Max. number of iterations, $R$ | 300 |

## 4.3 Experimental Results

We compare the results from the cluster assignments of the ABC algorithm against *k*-means. Table 2 summarizes the results of the two algorithms based on the mean amplitude (Amp) and mean reaction time (RT) for each word stimulus while Figure 5 depicts the cluster patterns. We further compare the fitness of each cluster using the within cluster some of square (SS) values. As can be observed in Figure 5, there exists a fairly distinct partition between the clusters, obviously based on the reaction time (RT) recorded from the stimuli. We further compare the performance of the two algorithms based on the difference in the total sum of squared distances, $E$ (Eq. 2).

**Table 2.** k-means and ABC cluster assignments

| | *k*-means | | | | ABC | | | |
|---|---|---|---|---|---|---|---|---|
| Class | Size | SS | Mean Amp(uV) | Mean RT(ms) | Size | SS | Mean Amp(Uv) | Mean RT(ms) |
| 1 | 43 | 112944.18 | -0.07597 | 1461.205 | 20 | 121368.2 | 0.031451 | 1828.55 |
| 2 | 16 | 116748.14 | -0.0307 | 1278.745 | 13 | 92810.56 | -0.04855 | 1259.45 |
| 3 | 44 | 109722.07 | 0.062156 | 1639.956 | 48 | 107163.4 | 0.031473 | 1619.16 |
| 4 | 17 | 90941.96 | 0.032604 | 1844.932 | 39 | 102379.1 | -0.0507 | 1439.73 |

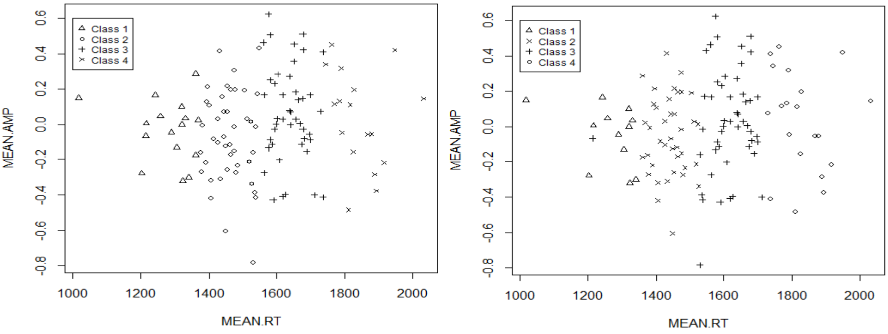


**Fig. 5.** Comparison between *k*-means left) and ABC (right) morpheme clusters

Smaller $E$ value denotes better clusters. Each algorithm was executed 5 times and the mean, max and min values of $E$ were recorded. Table 3 summarizes the performance of each algorithm. The ABC algorithm performs slightly better than the *k*-means algorithm by approximately 2.4%.

**Table 3.** Results for the ABC versus *k*-means clustering algorithms ($E$)

| Algorithm | Mean | Min. | Max. |
|---|---|---|---|
| *k*-means | 107485.6 | 82452.92 | 133838.7 |
| ABC | 102379.1 | 79658.72 | 132702.4 |

## 5 Conclusion and Future Work

In this research, we tested the feasibility of applying the ABC algorithm on EEG dataset to cluster natural language morphemes. Our initial results show that the ABC algorithm performs 2.4% better than the baseline *k*-means. The non-uniform distributions of cluster members denote the non-trivial task of recognizing specific morpheme patterns mostly due to the noise contamination in the raw EEG signals and less discriminant EEG features. This is well expected as pattern recognition tasks involving EEG recordings are prone to the variances across human subjects (heterogeneity) and the non-stationary signals. Using heuristics to determine good seeds for centroid/food source selections may efficiently improve the performance of both algorithms. As part of future work, we plan to establish an EEG/ERP model trained on correct identifications of desired morpheme groups. It is thus crucial to find an efficient way to automatically synchronize the stimulus presentation time to the EEG recordings which were done manually. Real time markings of a stimulus presentation in EEG recordings will produce more accurate EEG/ERP analysis of specific event-related stimulus.

**Acknowledgement.** This research is sponsored by the Ministry of Education (MOE) under the Fundamental Research Grant Scheme (FRGS-13-018-0259).

## References

1. Badecker, W., Caramazza, A.: Morphology and aphasia. In: Spencer, A., Zwicky, A.M. (eds.) Handbook of Morphology, pp. 390–406. Blackwell, Oxford (1998)
2. Barber, H., Domínguez, A., de Vega, M.: Human brain potentials indicate morphological decomposition in visual word recognition. Neuroscience Letters 318(3), 149–152 (2002)
3. Cui, X., Potok, T.E.: Document Clustering Analysis Based on Hybrid PSO+K-means Algorithm. Journal of Computer Sciences (Special Issue), 27–33 (2005)
4. Davis, M.H.: Units of representation in visual word recognition. Proc. of the National Academy of Sciences of the United States of America 101(41), 14687–14688 (2004)
5. Devlin, J.T., Jamison, H.L., Matthews, P.M., Gonnerman, L.M.: Morphology and the internal structure of words. Proc. of the National Academy of Sciences of the United States of America 101(41), 14984–14988 (2004)
6. Diependaele, K., Sandra, D., Grainger, J.: Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. Language and Cognitive Processes 20(1/2), 75–114 (2005)
7. Emotiv Systems. Emotiv EPOC, California, USA, `http://www.emotiv.com/epoc.php` (retrieved September 15, 2014)
8. Forster, K.I., Forster, J.C.: DMDX: A windows display program with millisecond accuracy. Behavior Research Methods Instruments and Computers 35(1), 116–124 (2003)
9. Karaboga, D., Akay, B.: A comparative study of Artificial Bee Colony algorithm. Applied Mathematics and Computation 214, 108–132 (2009)
10. Karaboga, D., Ozturk, C.: A novel clustering approach: Artificial Bee Colony (ABC) algorithm. Applied Soft Computing 11, 652–657 (2011)
11. Kuncheva, L.I., Rodriguez, J.J.: Interval feature extraction for classification of event-related potentials (ERP) in EEG data analysis. Progress in AI 2(1), 65–72 (2013)
12. Morris, J., Stockall, L.: Early, equivalent ERP masked priming effects for regular and irregular morphology. Brain and Language 123(2), 81–93 (2012)
13. Mizooji, K.K., Haghighat, A.T., Forsati, R.: Data clustering using bee colony optimization. In: 7th International Multi-Conference on Computing in the Global IT, pp. 189–194 (2012)
14. Münte, T.F., Say, T., Clahsen, H., Schiltz, K., Kutas, M.: Decomposition of morphologically complex words in English: evidence from event-related brain potentials. Brain Research Cognitive Brain Research 7(3), 241–253 (1999)
15. Passino, K.M., Seeley, T.D., Visscher, P.K.: Swarm Cognition in honey bees. Behavioral Ecology and Sociobiology 62, 401–414 (2008)
16. Pham, D.T., Otri, S., Afify, A., Mahmuddin, M., Al-Jabbouli, H.: Data Clustering using the Bees Algorithm. In: Proc. of the 40th CIRP International Manufacturing Systems Seminar, Liverpool, UK, pp. 1–8 (2007)
17. Rakshit, P., Bhattacharyya, S., Khasnobish, A., Tibarewala, D.N., Janarthanan, R.: Artificial Bee Colony based feature selection for motor imagery EEG data. In: Bansal, J.C., Singh, P., Deep, K., Pant, M., Nagar, A. (eds.) Proc. of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012). AISC, vol. 202, pp. 127–138. Springer, Heidelberg (2013)
18. Rodriguez-Fornells, A., Münte, T.F., Clahsen, H.: Morphological priming in Spanish verb forms: An ERP repetition priming study. Cognitive Neuroscience 14(3), 443–454 (2002)
19. Sanei, S., Chambers, J.A.: EEG signal processing. John Wiley & Sons, West Sussex (2013)
20. Schiezaro, M., Pedrini, H.: Data feature selection based on Artificial Bee Colony Algorithm. EURASIP Journal on Image and Video Processing (2013)
21. Schomer, D.L., da Silva, F.L.: Niedermeyer's electroencephalography: Basic principles, clinical applications and related fields. Lippincot Williams and Wilkins, Philadephia (2010)

# The Initial Investigation of the Design and Energy Sharing Algorithm Using Two-Ways Communication Mechanism for Swarm Robotic Systems

Amelia Ritahani Ismail, Recky Desia, and Muhammad Fuad Riza Zuhri

Department of Computer Science,
Kulliyyah of Information and Communication Technology,
International Islamic University Malaysia,
P.O. Box 10, 50728 Kuala Lumpur, Malaysia
amelia@iium.edu.my, {recky.qie,fuad.zuhri}@gmail.com

**Abstract.** Swarm Robotics (SR) is a new field of study that is mainly concerned with con-trolling and coordinating a multiple small robots. SR has several key characteristics that make it a preferable choice for a variety of tasks. The characteristics include lower cost, easiness to program, scalability of tasks and fault tolerance. The robustness from fault tolerance in SR comes from having a group of small robots working on the same task and thus enabling them to tolerate the loss of a few members of the swarm as the other members can still continue with the mission. However it has shown that continuous failure of members of a swarm such as those due to low energy have a significant impact on the overall performance of the swarm. In addition, the possibility of completion of the task is also dependent on the percentage of the swarm falling out of the group due insufficient energy. Some of the work that has been proposed by the researchers is by adding a charging station or a removable charger. However, these techniques have their own limitations. Therefore a work on having the robot(s) to charger themselves without the help of the charging station or a removable charger is proposed. But the work is only proven successful in simulation without a proper design and testing in a real robots scenario. This paper is therefore will describe our initial investigation on the design and the implementation of energy sharing algorithm using two-ways robotic swarm communication mechanism with NRF2401.

## 1 Introduction

Swarm intelligence (SI) is a field of Artificial Intelligence that is inspired by in-sects, birds and other animals [2, 3, and 5]. SI system consists of a group or a swarm of simple systems or agents interacting and communicating with each other and their environment. The definition of swarm intelligence denoted by [2] is:

> *'The discipline that deals with natural and artificial systems composed of many individuals that coordinate using decentralized control and self-organization. In particular, the discipline focuses on the collective behaviors that result from the local interactions of the individuals with each other and with their environment"*

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_7

Swarm Robotics (SR) is a new field of study [4, 1] that is concerned with controlling and coordinating a multiple small robots and is defined by [5] as:

> *"the study of how large number of relatively simple physically embodied agents can be designed such that a desired collective behavior emerges from the local interactions among agents and between the agents and the environment"*

According to [4, 3, and 5] robotic swarms have several advantages over their more complex individual robot and are the results of using many robots instead of just one. According to them, this is made possible because:

- It is easier to design simple robot units required for a swarm.
- Robot swarms are able to cover more area than an individual robot.
- Swarm robots are fault tolerant.
- Their effectiveness scales well with the number of members.
- The algorithms for swarms scale well and do not depend on the number of robots.
- Some tasks may be too complex for a single robot to perform.
- The communication between the robots is reduced because of the indirect interactions.
- They can accomplish some tasks that would be impossible for a single robot to achieve.

The main objective of this paper is to investigate the design and energy sharing algorithm using two-ways communication mechanism which is applied to swarm robotic systems to maintain the system's robustness with NRF2401. The rest of this paper is organized as follows: Section 2 explains on the problem definition; Section 3 describes the energy-sharing algorithm that is implemented in this paper; Section 4 explains the experimental design and Section 5 discusses the initial result which is obtained during the experiments. This paper ends with Section 6, which concludes this paper and explains our future work.

## 2 Problem Definition

One of the major advantages of swarm robotic systems is their robustness from fault tolerance. This is because SR has a group of small robots working on the same task and thus enabling them to tolerate the loss of a few members of the swarm as the other members can still continue with the mission. However, the work in [8] has shown that continuous failure of members of a swarm such as those due to low energy have a significant impact on the overall performance of the swarm. In this example, it relates to the time needed to complete the task. In addition, the possibility of completion of the task is also dependent on the percentage of the swarm falling out of the group due to the insufficient energy.

Because of the reason explained above, there are few algorithms that are proposed by [8] to allow robots in the swarm to share their energy without human or battery intervention. However, this work is only proven successful in simulation without a proper design and testing in a real robots scenario. This paper is therefore will describe our initial investigation on the design and the implementation of energy sharing algorithm using two-ways robotic swarm communication mechanism.

# 3 Literature Review

## 3.1 Energy Sharing

Commonly, the swarm robots use the battery charger or a charging station for their charging mechanism. [14] proposed a battery-exchanged mechanism to replace robots' with low energy with a fully charged battery. The battery station is provided with numbers of batteries that can be exchanged with any robots that have low in battery level. Therefore, if the robot's battery is about to empty, the robot needs to travel to the charging station and exchange its battery with the fully-loaded battery. As mentioned by [14], it takes around 36 seconds for each process to be completed. Furthermore, the robot also needs to travel to the charging station to exchange its battery leading to the additional travelling time for the process of changing the battery.

There is also another improved technique proposed by [15] for the charger mechanism. To save the time of recharging, a robot will be dedicated as the movable charger robot. It means that there will be one robot that works only for bringing a lot of energy to recharge the other robots when the other robots need more energy. The charger robot will continuously broadcast a message. In the time that a mobile robots battery level is lower than defined threshold level and are unable to move them, a request message will be sent to the charger robot to help them. However as mentioned by [15], there will be a case where the charger robot is out of battery in the middle of the work environment that may disturb the work of the other robots.

Due to the weaknesses described in [14] and [15], there are few proposed techniques by [8], which are inspired by the immune system that allow the robots in the system to share the energy of their battery without the help either the charging station and/or a charging robot(s). They are (1) Single Nearest Charger Algorithm, (2) Shared Nearest Charger Algorithm and (3) Granuloma Formation Algorithm. Basically, the algorithm principle is like this, the energy threshold is set to indicate the minimum energy of one robot before they can donate to another robot. If one robot send message for help so the other robot that has more energy will come and share the energy. But if one robot cannot help another robot so the robot leaves and let the other robot that have more energy to be shared to help. The techniques that are explained in [8] can help the robots in the system to share their energy and help each other as well as they can complete the task together. However, the technique is only proven successful in the simulation.

## 3.2 Communication between Two Robots

The energy sharing is a mechanism where a robot shares their energy to their partner when their partner needs to be charged. It means that there is cooperation between two robots or more. Therefore, if the energy sharing is implemented in a robot work system, there should be any communication between two robots or more to enable them to communicate. So the communication between robots becomes a main system that should be implemented.

Many application for mobile robots need to communicate with each other by using ad-hoc networking. [9] proposed and evaluate two routing protocols tailored for use in ad-hoc networks formed by mobile robot teams. The teams consist of Mobile Robot Distance Vector (MRDV) and Mobile Robot Source Routing (MRSR). Both protocols perform efficient routing. The simulation study shows that MRDV and MRSR incur lower overhead while operating in mobile robot network while compared to old mobile ad hoc network routing protocols.

Moreover, a new IVC communication architecture for platooning systems is proposed in [10]. Some weaknesses on DSRC are identified if used alone in such data demand scenarios. By adding a new communication technology to each vehicle, using IR and transmitting most of the control data through it, IVC constraint can be solved.

Furthermore, [11] proposed a common wireless remote control system for laboratory mobile robots. In this system, a single wireless IEEE 802.11g network is adopted for communication and a robot inside/outside two-level Client/Server architecture is utilized. This kind of system structure is suitable for latter any scale extension. Here different kinds of laboratory mobile robots are all regarded as one kind of special sensors, the system is transparent. To guarantee the IEEE 802.11g network performance in huge data transmission cases, an improved case of installing strong marine bridges for mobile robots is presented.

Another paper, [12], shows the implementation of a wireless system for the control of mobile robots using circuits with Self Timed (ST) Synchronization, implemented in reconfigurable devices FPGAs. The system is composed of a global network of small ST with the blocks of Xbee wireless transmission, of each processing unit, which will develop independent processes communicated by means of Modules of wireless transmission that form the network of activation of peripheral units. Besides that, a paper [13] shows a design of autonomous mobile platform that can be controlled using a GPS and electronic compass. The mobile robot is equipped also with MaxSream XBee Pro 2.4GHz radio modem communication module.

Moreover, for the moveable charger project [15], the robot is equipped with six infrared diodes and phototransistors for communication between the mobile charger and other swarm robots. Infrared and phototransistors of occupied docking station will be disabled during active charging. This behaviour prevent mobile robot from docking to occupied charger station.

Since different mechanisms of communication have been used by other researchers, in this paper another mechanism is proposed which is NRF2401. NRF2401 is a communication module that can be used to transmit and receive data. The NRF2401 offers the following benefits: (1) Highly reduced current consumption, (2) Lower system cost (facilitates use of less expensive micro controller) and (3) Reduced risk of ‘on-air’ collisions due to short transmission time. [16] proposed a design of group robots system based on wireless communication by using NRF2401. The result of the experiment shows that the system can achieve the wireless communication between the Master Robot and the Slave Robot successfully.

# 4 Experimental Setup

## 4.1 Robot Design

The initial experiment is done with two simultaneous robots with the design and components that are shown in Figure 1 and Figure 2. There are two part of robot body which are the upper part and the lower part. In the upper part, we put breadboard, Arduino Mega 2560 board, LCD, 3 LED's as notification, Arduino sensor shield and

NRF2401 as communication with other robots and Arduino Bluetooth module as communication with user as shown in Figure 1. The experiment has been done in 4 x 8 arenas and each cell has a measurement of 30 cm x 30 cm. The length and height of the obstacle are around 30 cm and 15 cm respectively. The horizontal and vertical value is provided to be processed by robot to construct its area of the map by calculate between the two values, start point is also given by user.

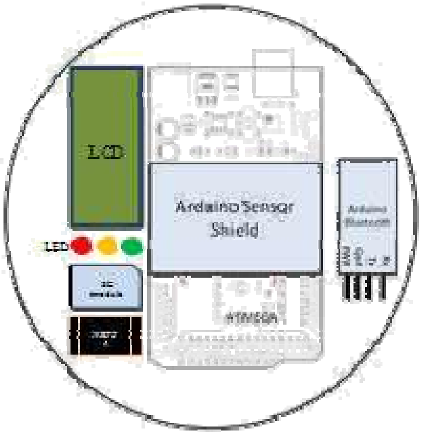


**Fig. 1.** The design of the upper robot body

In the lower part, we put ultrasonic sensor, driver stepper motor, stepper motor, battery box and wheel as shown in Figure 2. There are also few parameters that need to be included during the experiment. The parameters are:

- Start Point
- Destination
- Life Energy
- Energy to be shared

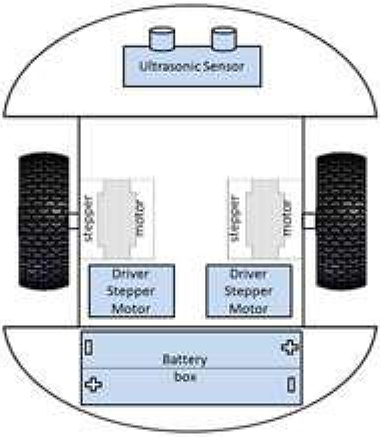


**Fig. 2.** The design of the lower robot body

### 4.2 Communication Algorithm

The algorithm describes in Algorithm 1 and Algorithm 2 below is the 2 ways communication used in- swarm robots. It contains the step by step instructions for trans-mitting and receiving command and also including energy sharing. This is how the scenario is look like. In beginning, each robot will have different energy levels; the energy will show in LED (red, yellow, green). During some activities the energy of the robot will reduce, when energy is empty the robot will stay in position and send help command and also the position, the other robot which is who is in free not execute any command will go to position location of robot that ask for help, it is done by "first replay" method which is when the robot ask for help, some of robot will replay the Massage then the robot will choose the robot that who first replay the massage. After the other robot arrive, the other will give the energy to the robot by sending the random number as the energy to share from the energy but no more than half of the energy and the number of energy of other robot will reduce by number of the energy shared.

```
if energy is empty
  then robot will not
  move
  repeat
    robot stops listening
    robot speak by sends help command and the
  position robot start listening
    wait

  if get notification there other robot is going to
    help break
```

```
    get notification other robot is arrive the robot position
    wait for energy transmission process
    get notification that energy sharing process is done turn on the notification led based on energy
  else
    continue looping
```

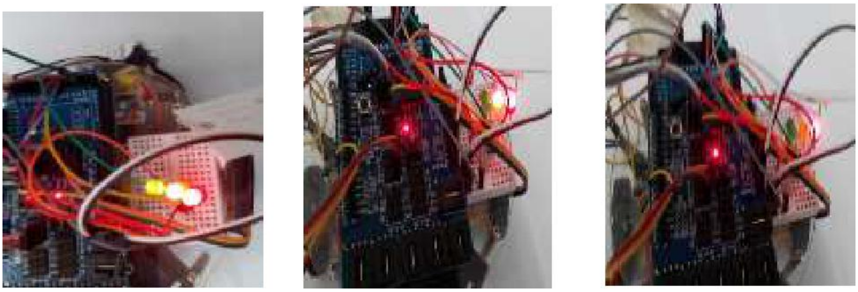
**Algo. 1.** Algorithm for transmitting by NRF2401, in a case of a robot that ask for help

```
if energy is not empty then
    if listen any help command then
          robot get destination robot stops listening
          robot speak by sends notification that the robot going to help
          go to the destination
          robot speak by sends notification that the robot arrive
          robot share the energy but not more than 50% of the energy
          robot speak by sends notification that energy has been shared
    robot start listening
  turn on the notification led based on energy level else
    robot do work
```

**Algo. 2.** Algorithm for receiving information by NRF2401

## 5 Result and Discussion

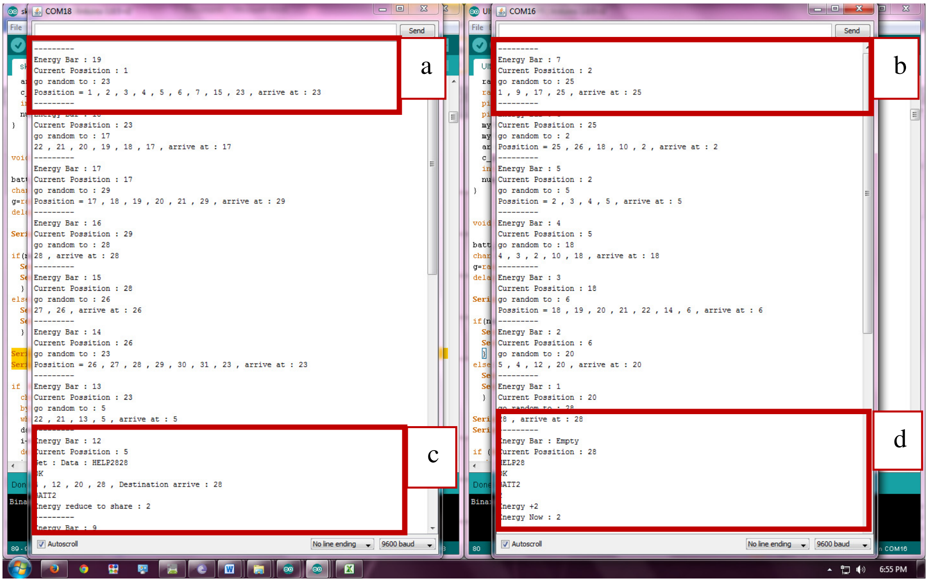
During this initial experiment, we make a comparison between the energy of the robots over time. We assume that during their work, the robots move from the start point to the destination point and throughout the movement, robot's energy is reduced. There are also obstacles in the environment and the robots used the obstacle avoidance algorithm to avoid obstacles in order to arrive to the destination point. The environment in this experiment also can be dynamic due to the robot can detect moving object and will avoid the object such as the robot will avoid other robot that moving inside arena beside avoid the obstacle itself. In depicting the energy level, each robot has three LED lights that show either their energy is: (1) high, (2) medium or (3) low. This is shown in Figure 3 below. In (a) for a high energy level, three (3) LED signals is shown. Meanwhile for medium (b) and low (c), two (2) and one (1) LED signals are shown respectively.



(a) (b) (c)

**Fig. 3.** Energy notification as depicted in: (a) Full robot's energy with three LED is ON. (b) Medium-low robot's energy with two LED is ON. (c)Low robot's energy with only one LED is ON.

Figure 3 shows the example of the LED signals that is shown during the ex-periment. In figure 3(a) and 3(b) robot is randomly assigned high and low energy value. During the experiment, robots are assumed working by moving from start point to the destination point. The path that the robot will follow in arriving to its destina-tion is also provided. Here different colour of LED showing different energy level is shown. For example in 3(a) the robot has high level of energy followed by medium low level of energy in 3(b) with two (2) LED signal and low level of energy signal as depicted in 3(c). This is to provide information on the level of energy during the ex-periment is conducted.



```
---------
Energy Bar : 19
Current Possition : 1
go random to : 23
Possition = 1 , 2 , 3 , 4 , 5 , 6 , 7 , 15 , 23 , arrive at : 23
---------
Current Possition : 23
go random to : 17
22 , 21 , 20 , 19 , 18 , 17 , arrive at : 17
---------
Energy Bar : 17
Current Possition : 17
go random to : 29
Possition = 17 , 18 , 19 , 20 , 21 , 29 , arrive at : 29
---------
Energy Bar : 16
Current Possition : 29
go random to : 28
28 , arrive at : 28
---------
Energy Bar : 15
Current Possition : 28
go random to : 26
27 , 26 , arrive at : 26
---------
Energy Bar : 14
Current Possition : 26
go random to : 23
Possition = 26 , 27 , 28 , 29 , 30 , 31 , 23 , arrive at : 23
---------
Energy Bar : 13
Current Possition : 23
go random to : 5
22 , 21 , 13 , 5 , arrive at : 5
nergy Bar : 12
urrent Possition : 5
et : Data : HELP2828
K
 , 12 , 20 , 28 , Destination arrive : 28
ATT2
nergy reduce to share : 2
---------
```

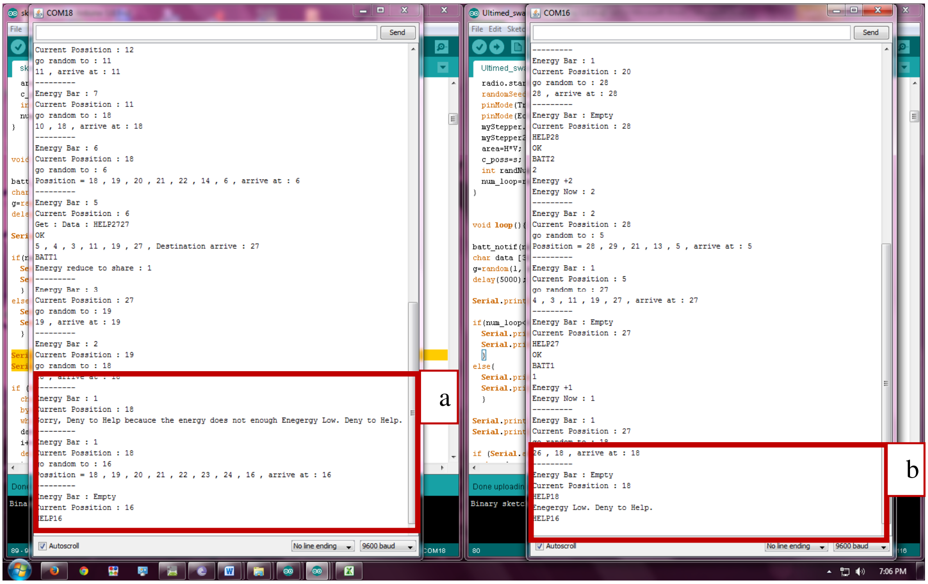
```
---------
Energy Bar : 7
Current Possition : 2
go random to : 25
1 , 9 , 17 , 25 , arrive at : 25
---------
Current Possition : 25
go random to : 2
Possition = 25 , 26 , 18 , 10 , 2 , arrive at : 2
---------
Energy Bar : 5
Current Possition : 2
go random to : 5
Possition = 2 , 3 , 4 , 5 , arrive at : 5
---------
Energy Bar : 4
Current Possition : 5
go random to : 18
4 , 3 , 2 , 10 , 18 , arrive at : 18
---------
Energy Bar : 3
Current Possition : 18
go random to : 6
Possition = 18 , 19 , 20 , 21 , 22 , 14 , 6 , arrive at : 6
---------
Energy Bar : 2
Current Possition : 6
go random to : 20
5 , 4 , 12 , 20 , arrive at : 20
---------
Energy Bar : 1
Current Possition : 20
8 , arrive at : 28
---------
nergy Bar : Empty
urrent Possition : 28
ELP28
K
ATT2

nergy +2
nergy Now : 2
```

**Fig. 4.** The output from the algorithm testing during the experiment: (a) position and energy level of the robots; (b) Robot do its work move from start point to destination point; (c) Energy level is reaching threshold and robot starts to ask for help and receive help processes; and (d) Energy transmitting process.

During the experiment, the output is displayed using Bluetooth communica-tion between the robots and a mobile device to identify the correctness of the algo-rithms that have been developed. This is shown in figure 4(a), (b), (c), (d) and figure 5(a) and (b). Figure 4(a) and (b) show the output from the experiment that displays the position and energy level of the robot and the position of the robot is shown. In 4(c) is an example of message that the robot receives from other robot that ask for help, which is shown in figure 4(d) because of their energy level is zero. In the case of the receiver robot has enough energy, it will move to the robot with zero energy to share its energy. Then, the output shows that one of the robot’s energy is reducing and another is increasing due the course of energy sharing mechanism that is conducted throughout the experiment.

The previous example is the scenario when a robot has enough energy to be shared; however there is also a scenario where the energy cannot be shared dues to the reason that both robots have low energy level. This is shown in figure 4(a) that shows the robot denies helping due to low of energy then the robot will send denied massage to other robot and continue doing its work. Figure 4(b) shows the robot get deny mas-sage by other robot to help and stay in its current position and there is no increased in its own energy.



```
Current Possition : 12
go random to : 11
11 , arrive at : 11
---------
Energy Bar : 7
Current Possition : 11
go random to : 18
10 , 18 , arrive at : 18
---------
Energy Bar : 6
Current Possition : 18
go random to : 6
Possition = 18 , 19 , 20 , 21 , 22 , 14 , 6 , arrive at : 6
---------
Energy Bar : 5
Current Possition : 6
Get : Data : HELP2727
OK
5 , 4 , 3 , 11 , 19 , 27 , Destination arrive : 27
BATT1
Energy reduce to share : 1
---------
Energy Bar : 3
Current Possition : 27
go random to : 19
19 , arrive at : 19
---------
Energy Bar : 2
Current Possition : 19
go random to : 18
8 , arrive at : 18
--------
nergy Bar : 1
urrent Possition : 18
orry, Deny to Help becauce the energy does not enough Enegergy Low. Deny to Help.
--------
nergy Bar : 1
urrent Possition : 18
o random to : 16
ossition = 18 , 19 , 20 , 21 , 22 , 23 , 24 , 16 , arrive at : 16
--------
nergy Bar : Empty
urrent Possition : 16
ELP16
```

```
---------
Energy Bar : 1
Current Possition : 20
go random to : 28
28 , arrive at : 28
---------
Energy Bar : Empty
Current Possition : 28
HELP28
OK
BATT2
2
Energy +2
Energy Now : 2
---------
Energy Bar : 2
Current Possition : 28
go random to : 5
Possition = 28 , 29 , 21 , 13 , 5 , arrive at : 5
---------
Energy Bar : 1
Current Possition : 5
go random to : 27
4 , 3 , 11 , 19 , 27 , arrive at : 27
---------
Energy Bar : Empty
Current Possition : 27
HELP27
OK
BATT1
1
Energy +1
Energy Now : 1
---------
Energy Bar : 1
Current Possition : 27
go random to : 18
26 , 18 , arrive at : 18
--------
Energy Bar : Empty
Current Possition : 18
HELP18
Enegergy Low. Deny to Help.
HELP16
```

**Fig. 5.** The feedback that get by user during testing: (a) The robot denies to help. (b)The robot get a deny massage

In figure 6, the energy levels over time of the two robots that have used are shown. In the beginning of the experiment, one of the robots is given a high level of energy and the second robot is given a low energy. The energy levels are then reduced slowly as both robots do their work. When one of robot’s energy level reach 0 or empty, the robot will stay in the current position and transmit help command and its position to the other robot. The robot will repeat this process until he get notification that his command has been listen by other robot. The other robot that gets the message will then move to the position of the robot and share the energy, energy that will give to the robot must not be more than 50% of the robot energy. However the robot can deny helping the robot with the low energy if its own energy is low.

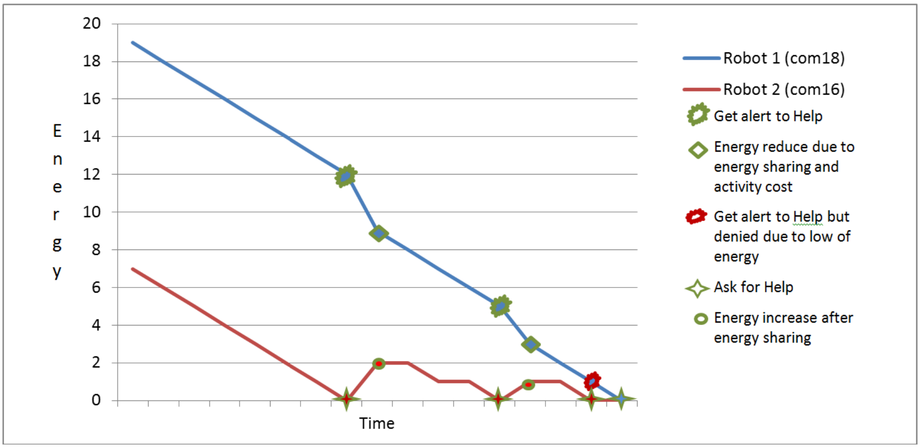


**Fig. 6.** The energy level of the robots during the experiment over time

## 6 Conclusion and Future Work

As a conclusion, the initial result that we obtained based on our design and implementation of the energy sharing algorithm has a potential to be implemented in the larger scale of swarm robotic systems. It is our hoped that the algorithm can be successfully implemented in a large swarm robotic systems to study on the effectiveness and the efficiency of the algorithm as well as the design that we have obtained so far.

**Acknowledgement.** This research is funded by the Ministry of Higher Education under the Exploratory Research Grant Scheme (ERGS): ERGS 13- 017-0050.

## References

1. Bayindir, L., Sahin, E.: A review of studies in swarm robotics. Turkish Journal of Electrical Engineering 15(2), 115–147 (2007)
2. Dorigo, M., Birattari, M., Stutzle, T.: Ant colony optimization. IEEE Computational Intelligence Magazine 1(4), 28–39 (2006)
3. Jevtic, A., Andina, D.: Swarm intelligence and its applications in swarm robot-ics. In: 6th WSEAS Int. Conference on Computational Intelligence, Man-Machine Systems and Cybernetics, pp. 41–46 (2007)
4. Miner, D.: Swarm robotics algorithms: A survey. Technical report, University of Maryland (2007)

5. Şahin, E.: Swarm robotics: From sources of inspiration to domains of application. In: Şahin, E., Spears, W.M. (eds.) Swarm Robotics 2004. LNCS, vol. 3342, pp. 10–20. Springer, Heidelberg (2005)
6. Pineda, L.E., Eiben, A.E., van Steen, M.: Evolving Communication in Robotic Swarms Using On-Line, On-Board, Distributed Evolutionary Algorithms. In: Di Chio, C., et al. (eds.) EvoApplications 2012. LNCS, vol. 7248, pp. 529–538. Springer, Heidelberg (2012)
7. Di Caro, G., Ducatelle, F., Gambardella, L.M.: Wireless communications for dis-tributed navigation in robot swarms (2009)
8. Ismail, A.R.: Immune-Inspired Self-Healing Swarm Robotic Systems. Ph.D. dissertation, University of York, United Kingdom (2011)
9. Das, S.M., et al.: Efficient Unicast Messaging for Mobile Robot. In: International Conference on Robotics and Automation, pp. 93–98 (2005)
10. Fernandes, P., Nunes, U.: Platooning with DSRC-Based IVC-Enabled Autonomous Vehicles: Adding Infrared Communications for IVC Reliability Improvement. In: Intelligent Vehicles Symposium, pp. 517–522 (2012)
11. Liu, H., et al.: A Common Wireless Remote Control System for Mobile Robots in Laboratory (2012)
12. Ortega-Cisneros, S., et al.: Implementation of A Wireless Control System with Self Timed Activation for Mobile Robots, pp. 205–208 (2008)
13. Jaskot, K., Babiarz, A.: Autonomous Mobile Platform II. In: 12th International Carpathian Control Conference, pp. 167–171 (2011)
14. Vaussard, F., et al. (n.d.): Towards Long-Term Collective Experiments
15. Arvin, F., Samsudin, K., Ramli, A.R.: Swarm Robots Long Term Autonomy Using Moveable Charger. In: International Conference on Future Computer and Communication, pp. 127–130 (2009)
16. Cai, G., Harada, Y.: Design of Group Robots System Based on Wireless Communication. In: International Conference of Information Technology, Computer Engineering and Management Sciences, pp. 19–22 (2011)

# Second Order Back Propagation Neural Network (SOBPNN) Algorithm for Medical Data Classification

Nazri Mohd Nawi, Norhamreeza Abdul Hamid, Nursyafika Harsad,
and Azizul Azhar Ramli

Faculty of Computer Science & Information Technology,
University Tun Hussein Onn Malaysia, Johor, Malaysia
`{nazri,azizulr}@uthm.edu.my`,
`{gi090007,ai110052}@siswa.uthm.edu.my`

**Abstract.** Gradient based methods are one of the most widely used error minimization methods used to train back propagation neural networks (BPNN). Some second order learning methods deal with a quadratic approximation of the error function determined from the calculation of the Hessian matrix, and achieves improved convergence rates in many cases. This paper introduces an improved second order back propagation which calculates efficiently the Hessian matrix by adaptively modifying the search direction. This paper suggests a simple modification to the initial search direction, i.e. the gradient of error with respect to weights, can substantially improve the training efficiency. The efficiency of the proposed SOBPNN is verified by means of simulations on five medical data classification. The results show that the SOBPNN significantly improves the learning performance of BPNN.

**Keywords:** Back Propagation, Search direction, Gain variation, Second Order Neural Network, Conjugate Gradient, Quasi-Newton.

## 1 Introduction

Traditional learning algorithms for feed forward neural networks use gradient descent (GD) techniques to find a local minimum of the error function and the corresponding networks' weights. GD is one of the most widely gradient based algorithms used to train back propagation (BP) networks in minimizing error. GD techniques involve backward error correction of network weights. Despite the general success of BP method in the learning process, several major deficiencies are still needed to be solved [1]. The convergence rate of BP is very low and hence it becomes unsuitable for large problems.

Furthermore, the convergence behavior of the BP algorithm depends on the choice of initial values of connection weights and other parameters used in the algorithm such as the learning rate and the momentum term. Thus, BP needs improvement to perform well and overcome those drawbacks.

Several related research in BP in previous years back improving on selection of better activation function [2] and selection of learning rate as well as momentum

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems*,
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_8

[3-5]. Unfortunately, in practice, even with these modifications, BP still has some drawback as long as it uses the first order neural network method, specifically GD method with poor convergence properties [4].

Later, as summarized by Bishop [5] various optimization techniques were suggested for improving the efficiency of error minimization process or in other words the training efficiency. Among these are second order methods such as Conjugate Gradient (CG) and Newton method [7-8].

Recently, the influence of gain was studied by few researchers [9-11]. The gain parameter controls the steepness of the activation function. It has been shown that a larger gain value has an equivalent effect of increasing the learning rate.

This paper suggests that a simple modification to the initial search direction, i.e. the gradient of error with respect to weights, can substantially improve the training efficiency. It was discovered that if the gradient based search direction is locally modified by a gain value used in the activation function of the corresponding node, significant improvements in the convergence rates can be achieved. It has also been shown that the proposed method is robust, easy to compute, and easy to implement into the well known nonlinear conjugate gradient algorithms.

Thus, this research implements and examines the representational capabilities of the proposed SOBPNN which consists of Conjugate Gradient Fletcher-Reeves (CGFR) [12], Conjugate Gradient Polak Ribiere (CGPR) [13] and Davidon-Fletcher-Powell (DFP) [14]. The performance of SOBPNN and BPNN are compared by some simulation experiments on five medical data classification problems. The accuracy, computation time and epochs iteration of each method is summarized and compared one to another.

The continuing section of this paper is organized such as follows: Section 2 illustrates the implementation of the proposed method in the research. Section 3 discusses the results of the simulation testing on the method implemented. The closing section contains concluding remarks and brief discussion for further research.

## 2 The Proposed Second Order Back Propagation (SOBPNN) Algorithm

In this section, a new approach for improving the training efficiency of gradient descent method (back propagation algorithm) is presented. The proposed method (SOBPNN) modifies the gradient based search direction by changing the gain value adaptively for each node.

The learning process in neural networks is one of the optimization processes where the objective of a learning process in neural networks is to find a weight vector, *w* that minimizes the difference between the actual output and the desired output on both training and testing data sets.

$$\min_{w \in \Re^n} E(w) \tag{1}$$

Consider a multilayer feed forward neural network (FNN) [15] with one output layer, one input layer and one or more hidden layers. Each layer has a set of units, nodes, or neurons. It is usually assumed that each layer is fully connected with a previous layer without direct connections between layers which are not consecutive. Each connection has a weight.

For a particular input pattern, an error function is define on that pattern as,

$$E = \frac{1}{2} \sum_k (t_k - o_k^L)^2 \tag{2}$$

Where $o_k^L$ is the activation of the $k^{th}$ node of layer $L$. Let $w_{ij}^L$ be the weight on the connection from the $i^{th}$ node in layer $L-1$ to the $j^{th}$ node in layer $L$. The overall error on the training set is simply the sum, across patterns, of the pattern error $E$.

The net input to the $j^{th}$ node of layer $L$ is defined as $net_j^L = (w_j^L, o^{L-1}) = \sum_k w_{j,k}^L o_k^{L-1}$ The activation of a node $o_j^L$ is given by a function of its net input,

$$f(o_j^L) = \frac{1}{1 + \exp(net_j^L)^{c_j^L}} \tag{3}$$

Where $f$ is any function with bounded derivative, and $c_j^L$ is a real value called the gain of the node. Note that at $c_j^L = 1$, this activation function becomes the usual logistic activation function.

The weight update expression with a non- unit gain value is derived by differentiating the error term as given in Equation (2) with respect to $w_{ij}^L$ as follows:

$$\frac{\delta E}{\delta w_{ij}^L} = \frac{\delta E}{\delta net^{L+1}} . \frac{\delta net^{L+1}}{\delta o_j^L} . \frac{\delta o_j^L}{\delta net_j^L} . \frac{\delta net_j^L}{\delta w_{ij}^L}$$

$$= [-\delta_1^{L+1} \ldots - -\delta_n^{L+1}] . \begin{bmatrix} w_{1j}^{L+1} \\ \ldots \\ w_{nj}^{L+1} \end{bmatrix} . f'(c_j^L net_j^L) . c_j^L . o_j^{L-1} \tag{4}$$

In particular, the first three factors of Equation (4) indicate that

$$\delta_1^L = (\sum_k \delta o_k^{L+1} . w_{k,j}^{L+1}) . f'(c_j^L . net_j^L) \tag{5}$$

As we know that the iterative Equation (5) for $\delta_1^L$ is the same formula as standard back propagation [15] except for the appearance of the value gain.

By integrating Equation (4) and (5) yields the learning rule for weights:

$$\Delta w_{ij}^{L} = \eta \delta_{j}^{L}.c_{j}^{L}.o_{j}^{L-1}$$
$$= \eta \frac{\delta E}{\delta w_{ij}^{L}} \tag{6}$$

where $\eta$ is a 'learning rate' and the search direction or gradient vector at point $w_{ij}^{L}$ is $d = \frac{\delta E}{\delta w_{ij}^{L}} = g$. In the proposed method the calculation of the gradient of error $g^{(n)}$ at step $n$ is a function of gain $c_{j}^{L(n)}$ as follows:

$$d^{(n)} = \frac{\delta E}{\delta w_{ij}^{L(n)}}.(c_{j}^{L(n)}) = g^{(n)}.c_{j}^{L(n)} \tag{7}$$

The gain value at step $n$ is calculated using gradient of error w.r.t. to gain,

$$\frac{\delta E}{\delta c_{j}^{L}} = (\sum_{k} \delta c_{k}^{L+1} w_{kj}^{L+1}) f'(c_{j}^{L}.net_{j}^{L}).net_{j}^{L} \tag{8}$$

Then the gradient descent rule for the gain value is formulated as follows:

$$\Delta c_{j}^{L} = \eta \delta_{j}^{L}.\frac{net_{j}^{L}}{c_{j}^{L}} \tag{9}$$

At the end of each iteration the new gain value is updated using a simple gradient based method as given by the formula,

$$c_{j}^{new} = c_{j}^{old} + \Delta c_{j}^{L} \tag{10}$$

The advantage of the proposed modification is that it is generic and can be implemented easily into second order methods. Therefore, the next section will discuss on the implementation of SOBPNN with other second order methods such as Conjugate Gradient and Quasi-Newton method. In fact, Conjugate Gradient algorithm was proposed by several researches [7] [16] [17] [18].

Suppose in step $n$ in gradient descent algorithm, the current weight vector is $w^{n}$ and particular gradient based search direction is $d^{n}$. The weight vector at step $(n+1)$ is computed by using Equation (11) [18]:

$$w^{(n+1)} = w^{n} + \eta^{n} d^{n} \tag{11}$$

Where $\eta^{n}$ is the learning rate value at step $n$. Then, gradient based search direction is calculated at each step using Equation (12):

$$d^{(n)} = -\frac{\partial E}{\partial W_{ij}^{(n)}} = g^{(n)} \tag{12}$$

The algorithm begins minimization process with an initial estimate $w_0$ and as initial gradient based search direction as (13):

$$d_0 = -\nabla E(w_0) = -g_0 \tag{13}$$

For every epoch mentioned in Equation (12), the search direction at $(n+1)^{th}$ iteration calculated:

$$d_{n+1} = \frac{\partial E}{\partial W_{n+1}}(c_{i,n+1}) + \beta_{n+1} d_n (c_{i,n}) \tag{14}$$

Where the scalar $\beta_n$ to be determined by the requirement that $d_n$ and $d_{n+1}$ must fulfill the conjugacy property [2]. Meanwhile, the scalar $\beta_n$ referred to Polak, Ribiere and Polyak and Fletcher and Reeves method. Hence, represented by the Equation (15) and (16) respectively [18]:

$$\beta_n = \frac{(g_{n+1} - g_n)^T g_{n+1}}{g_n^T g_n} \tag{15}$$

$$\beta_n = \frac{g_{n+1}^T g_{n+1}}{g_n^T g_n} \tag{16}$$

In other hand, Quasi Newton method or Davidon Fletcher Powell method is known by the condition Equation (17).

$$H_{n+1} S_n = y_n \tag{17}$$

Where $y_n$ is the gradient change at iteration $n$ . By using Equation (12), the gradient change is calculated as Equation (18):

$$y_n = g_{n+1} - g_n \tag{18}$$

Meanwhile $S_n$ is the change in position $w$ at iteration $n$ such as Equation (19).

$$S_n = w_{n+1} - w_n \tag{19}$$

Quasi-Newton method are categorized in terms of simple Equation (20) where $\nabla_n$ is a correction. The method used for updating $\nabla_n$ in this research is Davidon-Fletcher-Powell (DFP) as in Equation (21):

$$H_{n+1} = H_n + \nabla_n \tag{20}$$

$$\nabla_n = \left(1 + \frac{S_n^T H_n S_n}{y_n^T S_n}\right) \frac{y_n y_n^T}{y_n^T S_n} - \frac{y_n S_n^T H_n + H_s S_n y_n^T}{y_n^T S_n} \tag{21}$$

The pseudo code for CGFR and CGPR as followed:

```
Start
Step 1: Initialize the weight vectors randomly, the gradient vector
        $g_0$ to zero. Let the first search direction $d_0$ be Set $\beta_0 = 0$ ,
        $epoch = 1$ . Let $Nt$ be the total number of weight values. Select
        a convergence tolerance, $CT$ .
Step 2: At step $n$ , evaluate gradient vector, $g_n$
Step 3: Evaluate $E(w_n)$ . IF $E(w_n) < CT$ then STOP Training ELSE go to
        Step 4.
Step 4: Calculate a new gradient based search direction,
        $d_n = -g_n + \beta_n d_{n-1}$
Step 5: IF $n > 1$ THEN, update using Equation (12) or Equation (13)
        ELSE go to Step 6.
Step 6: IF $[(epoch+1)/Nt] = 0$ THEN 'restart' the gradient vector with
        $d_n = -g_n$ ELSE go to Step 7.
Step 7: Calculate the optimal value for learning rate $\eta_n^*$ by using line
        search technique such as $E(w_n + \eta_n^* d_n) = \underset{\lambda >= 0}{\min} E(w_n + \eta_n d_n)$
Step 8: Update $w_n : w_{n+1} = w_n + \eta_n^* d_n$
Step 9: Set $n = n + 1$ and go to Step 2.
End
```

Meanwhile, pseudo code for DFP as followed:

```
Start
Step 1: Initialize the initial weight vectors $w_0$ to random values and
        undertake a positive definite initialization of the Hessian
        matrix $H(0)$ . Select a convergence tolerance, $CT$ .
Step 2: Compute the gradient based search direction, $d_n$ at step $n$
        $d_n = -H_n g_n$ .
Step 3: Search the optimal value for learning rate $\eta_n$ by using line
        search technique such as $E(w_n + \eta_n^* d_n) = \underset{\lambda >= 0}{\min} E(w_n + \eta_n d_n)$
Step 4: Update $w_n : w_{n+1} = w_n + \eta_n^* d_n$ .
Step 5: Compute $S_n = w_{n+1} - w_n$ , $y_n = g_{n+1} - g_n$ and $\nabla_n = \left(1 + \frac{y_n^T H_n y_n}{S_n^T y_n}\right)\frac{S_n S_n^T}{S_n^T y_n} - \frac{S_n y_n^T H_n}{S_n^T y_n}$
Step 6: Update the Hessian matrix $H_n$ , $H_{n+1} = H_n + \nabla_n$
Step 7: Compute the error function value $E(w_n)$
Step 9: IF $E(w_n) > CT$ go to Step 2, ELSE stop training
End
```

## 3 Results and Discussions

The performance criterion used in this research focuses on the speed of convergence, measured in number of iterations, CPU time and accuracy. Four methods have been utilized in this research which are BPNN, CGPR, CGFR and DFP methods. This research was carried out by testing those methods performance using five medical data classification such as Breast Cancer [19], Diabetes [20], Heart [21], Thyroid [22] and Oral Cancer [23]. The simulation testing were done by using Matlab R2010b software and performed on a CPU of AMD E2 - 1800, with 1.7 GHz processor. Meanwhile, some values are set as shown in Table 1.

**Table 1.** Fixed variables

| **Variables** | Hidden Nodes | Target Error | Maximum Epoch | Momentum | Learning Rate | Trials Total |
|---|---|---|---|---|---|---|
| **Value** | 5 | 0.01 | 5000 | 0.4 | 0.3 | 30 |

The simulation required data such as epoch, CPU time and accuracy. The data is then calculated into average and standard deviation (SD).

### 3.1 Breast Cancer Problem

This dataset was generated from University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg [19]. The input attributes are for instance the clump thickness, the uniformity of cell size, the uniformity of cell shape, the amount of marginal adhesion, the single epithelial cell size, frequency of bare nuclei, bland chromatin, normal nucleoli and mitoses. This problem tries to diagnosis of Wisconsin breast cancer by trying to classify a tumor as either benign or malignant based on cell description gathered by microscopic examination. This data set consists of 9 inputs and 2 outputs. The results were recorded as in Table 2.

**Table 2.** Breast cancer data set result summary

| | **BPNN** | | **CGPR** | | **CGFR** | | **DFP** | |
|---|---|---|---|---|---|---|---|---|
| | Average | SD | Average | SD | Average | SD | Average | SD |
| epoch | 2473 | 1592 | **33** | 3 | 4061 | 1697 | 4518 | 1013 |
| CPU time | 124.84 | 91.504 | **3.065** | 0.686 | 415.720 | 186.076 | 473.359 | 117.767 |
| accuracy | 93.217 | 9.445 | 95.015 | 0.497 | 94.274 | 4.482 | **95.986** | 0.744 |

As summarized in the Table 2, DFP perform much better than other methods with the highest average accuracy by 95.986% to converge. Therefore, DFP was better in average accuracy with the value of 2.769% more than BPNN. In second place, is CGPR with the value 1.798% more than BPNN and CGFR comes in a third place with the average value of 1.057% more than BPNN. Other than that, CGPR performs better in terms of CPU time average where CGPR is 40 times faster than BPNN. But CGFR and DFP failed to converges faster than BPNN where BPNN converges 4 times faster than DFP and CGFR. In terms of iteration, CGPR performs fewer epochs than other methods which is 33 epochs. Ironically, CGFR and DFP performed poorly as both method reached over 4000 epochs. The results clearly show that the SONN (CGPR and DFP) outperformed BPNN method.

### 3.2 Diabetes Problem

Diabetes data set [20] taken from UCI Machine Learning Website, this data set describes that diabetes patient records were obtained from two sources: an automatic electronic recording device and paper records. Therefore, this data set consists of 384 instances. There is 10 attributes where 8 attributes are for input and 2 attributes for output. Table 3 summarizes the simulation testing result on Diabetes data set.

**Table 3.** Diabetes data set result summary

| | **BPNN** | | **CGPR** | | **CGFR** | | **DFP** | |
|---|---|---|---|---|---|---|---|---|
| | Average | SD | Average | SD | Average | SD | Average | SD |
| epoch | 2442 | 1249 | **87** | 5 | 5000 | 0 | 5000 | 0 |
| CPU time | 117.907 | 63.040 | **6.751** | 0.428 | 512.409 | 9.525 | 556.035 | 4.577 |
| accuracy | 68.423 | 5.292 | **68.507** | 0.444 | 61.148 | 7.064 | 70.811 | 0.697 |

In this data set, CGPR performs 17 times faster than BPNN with just an average of 6.7507 seconds to converge with 87, BPNN converge with 2442 epochs, while CGFR and DFP both were 5 times slower than BPNN to converge with 5000 epochs. Whereas CGFR and DFP considered failed to converge since it did not reach target error. Meanwhile, CGPR performs slightly better than other methods with the accuracy of 68.5073% to classify the problem which is 0.084 more accurate than BPNN. The results shows that CGPR defeat other methods in this problem

### 3.3 Heart Problem

The Heart data set [21] consists of 37 attributes, where 35 are the input and 2 are the output. T his data set has 460 instances, taken from UCI Machine Learning Website. The simulation result for this benchmark data set was summarized in Table 4.

**Table 4.** Heart data set result summary

| | **BPNN** | | **CGPR** | | **CGFR** | | **DFP** | |
|---|---|---|---|---|---|---|---|---|
| | Average | SD | Average | SD | Average | SD | Average | SD |
| epoch | 1544 | 412 | **61** | 5 | 5000 | 0 | 5000 | 0 |
| CPU time | 78.855 | 22.472 | **7.245** | 1.716 | 641.382 | 119.344 | 927.492 | 175.789 |
| accuracy | 77.065 | 1.054 | **77.217** | 0.592 | 68.641 | 9.174 | 77.109 | 1.270 |

Table 4 shows that CGPR has highest accuracy with 0.152% higher than BPNN. DFP and CGFR consider failed classifying the problem as it cannot converge to the target error value. In terms of CPU time, CGPR performs 10 times faster than BPNN. This can be proven by the convergence time for CGPR which was 7.245 seconds with 61 epochs and BPNN was 78.855 seconds with 1544 epochs. However CGFR and DFP both took more time to converge within 641.382 seconds and 927.4915 seconds with 5000 epochs respectively. Hence, still CGPR beat other methods.

## 3.4 Thyroid Problem

The dataset has 21 attributes and can be assigned to hyper-, hypo- and normal function of thyroid gland [22] where 21 attributes are the input and 3 are the output. There are 3600 instances. This data set was taken from UCI Machine Learning Website records provided by the Garavan Institute of Sydney, Australia. The simulation result for thyroid data set was summarized into Table 5.

**Table 5.** Thyroid data set result summary

| | **BPNN** | | **CGPR** | | **CGFR** | | **DFP** | |
|---|---|---|---|---|---|---|---|---|
| | Average | SD | Average | SD | Average | SD | Average | SD |
| epoch | 5000 | 0 | **33** | 3 | 5000 | 0 | 5000 | 0 |
| CPU time | 572.263 | 43.992 | **10.141** | 1.105 | 1593.850 | 46.655 | 1807.387 | 26.449 |
| accuracy | 95.148 | $4.34\text{x}10^{-14}$ | **95.242** | 1.326 | 95.910 | 0.633 | 92.825 | 0.706 |

Table 5 shows that CGPR outperformed other methods for all performance metrics. CGPR only required 33 epochs in 10.141 seconds with 95.242% accuracy. While the other methods considered failed since it did not reach the target error.

## 3.5 Oral Cancer Problem

The Oral Cancer data set [23] consists of 110 instances. The attributes of this data set was 7 where there was 6 inputs and 1 output. The data set were taken from the previous thesis of Siti Khadijah Masri entited as “Pengkelasan Simptom-simptom Barah Mulut Menggunakan Pendekatan Rangkaian Neural” [24]. The thesis stated that the data set records were provided by Batu Pahat Dental Clinic. The simulation result for this benchmark problem data set was summarized in Table 6.

**Table 6.** Oral Cancer data set result summary

| | **BPNN** | | **CGPR** | | **CGFR** | | **DFP** | |
|---|---|---|---|---|---|---|---|---|
| | Average | SD | Average | SD | Average | SD | Average | SD |
| epoch | 5000 | 0 | **57** | 14 | 5000 | 0 | 5000 | 0 |
| CPU time | 154.394 | 1.421 | **2.090** | 0.582 | 289.067 | 67.663 | 451.320 | 79.592 |
| accuracy | 78.588 | 0.054 | **78.770** | 0.237 | 78.240 | 0.371 | 78.568 | 0.124 |

Based on the Table 6, BPNN, CGFR and DFP reached the maximum value of 5000 epochs as the average epoch. Thus, it considered failed. Meanwhile, CGPR converges to global minima just within 57 epochs within 2.090 seconds with 0.181% accurate less than BPNN. Overall, still, the CGPR perform slightly better than other methods.

## 4 Conclusion

This research presents the performance of SOBPNN for the medical data sets. The SOBPNN algorithm is standard and has been implemented in all commonly used gradient based optimization processes. An evaluation on the simulation results showed that the SOBPNN algorithm indicates a promising trade-offs achievements in terms of computational time, rate of convergence and accuracy.

**Acknowledgement.** The authors would like to thank Office of Research, Innovation,Commercialization and Consultancy (ORICC), Universiti Tun Hussein Onn Malaysia (UTHM) and Ministry of Higher Education (MOHE) Malaysia for supporting this research under the Fundamental Research Grant Scheme (FRGS) vote no. 1236.

## References

1. Zhang, W.J.: Computational Ecology: Artificial Neural Networks and Their Applications, pp. 41–47. World Scientific Publishing Co. Pte. Ltd., Singapore (2010)
2. Chandra, P., Singh, Y.: An activation function adapting training algorithm for sigmoidal feedforward networks. Neurocomputing 61, 429–437 (2004)
3. Yu, X.H., Chen, G.A., Cheng, S.X.: Acceleration of backpropagation learning using optimized learning rate and momentum. Electronics Letters 29(14), 1288–1289 (1993)
4. Hamid, N.A., Nawi, N.M., Ghazali, R.: The Effect of Adaptive Gain and Adaptive Momentum in Improving Training Time of Gradient Descent Back Propagation Algorithm on Classification Problems. International Journal on Advanced Science, Engineering and Information Technology 1(2), 178–184 (2011)
5. Hamid, N.A., Nawi, N.M., Ghazali, R., Salleh, M.N.M.: Solving Local Minima Problem in Back Propagation Algorithm Using Adaptive Gain, Adaptive Momentum and Adaptive Learning Rate on Classification Problems. International Journal of Modern Physics: Conference Series 9, 448–455 (2012)
6. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press (1995)
7. Nawi, N.M., Ransing, M.R., Ransing, R.S.: An improved Conjugate Gradient based learning algorithm for back propagation neural networks. International Journal of Computational Intelligence 4(1), 46–55 (2007)
8. Wang, Z., Fang, J., Liu, X.: Global stability of stochastic high-order neural networks with discrete and distributed delays. Chaos, Solitons & Fractals 36(2), 388–396 (2008)
9. Thimm, G., Moerland, P., Fiesler, E.: The Interchangeability of Learning Rate and Gain in Backpropagation Neural Networks. Neural Comput. 8, 451–460 (1996)
10. Holger, R.M., Graeme, C.D.: The Effect of Internal Parameters and Geometry on the Performance of Back propagation Neural Networks. Environmental Modeling and Software 13(1), 193–209 (1998)

11. Eom, K., Jung, K., Sirisena, H.: Performance Improvement of Backpropagation Algorithm by Automatic Activation Function Gain Tuning Using Fuzzy Logic. Neurocomputing 50, 439–460 (2003)
12. Fletcher, R., Powell, M.J.D.: A rapidly convergent descent method for minimization. British Computer J., 163–168 (1963)
13. Fletcher, R., Reeves, R.M.: Function minimization by conjugate gradients. Comput. J. 7(2), 149–160 (1964)
14. Huang, H.Y.: A unified approach to quadratically convergent algorithms for function minimization. J. Optim. Theory Appl. 5, 405–423 (1970)
15. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning Internal Representations by Error Propagation. In: Rumelhart, D.E., McClelland, J.L. (eds.) Parallel Distributed Processing, vol. 1, pp. 318–362 (1986)
16. Huzaifa, I., Asivadam, V.S., Saad, N.: Enhanced Conjugate Gradient Methods for Training MLP Networks. In: Proceedings of 2010 IEEE Student Conference on Research and Development (SCOReD 2010), Putrajaya, Malaysia (2010)
17. Hager, W., Zhang, H.: A Survey on Nonlinear Conjugate Gradient Methods. Pacific Journal of Optimization 2, 35–58 (2006)
18. Nawi, N.M., Ghazali, R., Salleh, M.N.M.: An Approach to Improve Back-propagation algorithm by using Adaptive Gain. Biomedical Soft Computing and Human Sciences 16(2), 125–134 (2010)
19. Mangasarian, O.L., Wolberg, W.H.: Cancer diagnosis via linear programming. SIAM News 23, 1–18 (1990)
20. Smith, J.W., Everhart, J.E., Johannes, R.S.: Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In: Proceedings of the Symposium on Computer Applications and Medical Care, pp. 261–265 (1988)
21. Detrano, R., Janosi, A., Steinbrunn, W.: International application of a new probability algorithm for the diagnosis of coronary artery disease. American Journal of Cardiology 64, 304–310 (1989)
22. Coomans, D., Broeckaert, I., Jonckheer, M., Massart, D.L.: Comparison of Multivariate Discrimination Techniques for Clinical Data - Application to The Thyroid Functional State. Methods of Information Medicine 22, 93–101 (1983)
23. Masri, S.K.: Pengelasan Simptom-simptom Barah Mulut Menggunakan Pendekatan Rangkaian Neural. Undergraduate Thesis. Universiti Tun Hussein Onn Malaysia (2012)

# Weighted Linear Fractional Programming for Possibilistic Multi-objective Problem

Nureize Arbaiy

University Tun Hussein Onn Malaysia, Batu Pahat 86400, Malaysia
nureize@uthm.edu.my

**Abstract.** The assessment of the weights of objective function plays an important role in a multi-objective process. This paper discusses a weighting method for linear fractional programming to solve possibilistic programming of the multi-objective decision-making problem. The minimal and maximal values of the objective function are utilized in the determination the weight value. This analysis concludes that it is worthwhile to pursue proposed solution approach to the multi-objective evaluation scheme, which addresses some limitation to determine the weight values.

**Keywords:** Weighting method, Multi-objective problem, Possibilistic Programming, Linear fractional.

## 1 Introduction

The nature of most multi-objective decision analysis includes conflicting issues. It makes the developed mathematical model for the multi-objective problem should capable to conduct compromise analysis. This is to achieve a balance among several conflicting features or objectives in this case. There are numerous weighting methods available in the literature which discusses the technique and its merits e.g. [1,2,3]. The strategy to use the relative weight to the objective functions is to form a singular objective function. The results of this strategy will obtain the best compromise solution if the weight assigned to the objective function is demonstrating the relative preference. Such procedure mimics the nature of human decision making which tend to select the compromising solutions in particular preference structure.

Real world problem and the nature of decision making which depends on the human judgments contribute to various types of uncertainties. Such uncertainty may specifically come from the human preference and expertise, machine faults, varying expert evaluations, and incomplete information [4]. Since formulating mathematical model is necessary to find a solution from real world problem, transforming the inherent uncertainties from nature to the mathematical model is difficult. Modeling real-world practical problem into mathematical model requires numerical values which is usually neglects the uncertainties. However, translating uncertainties or incomplete information to the precise numerical values requires appropriate approach that is difficult and yet challenging [5]. The imprecision in a problem formulation is generally

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_9

due to the inclusion of human judgments and preferences, and the random situation. The formulated problem may results in misleading and improper solution if such parameters are not appropriately determined as crisp values in the mathematical model [6]. Hence, the uncertainties must be handled properly to ensure that the developed mathematical model for the problem account the uncertainties of evaluations into consideration.

It is important to address inherent uncertainty to obtain an optimal solution. Many approaches, such as probability distribution, fuzzy numbers, and different types of thresholds [7], have been used to model uncertainty and imprecision, in the distinct occurrence of the uncertainty. Probability and possibility theories are introduced to treat the random and fuzzy information, respectively. Fuzzy sets [8] also have made significant role and are useful for representing uncertain and imprecise information. Stochastic programming model (i.e., [9,10]) and the fuzzy programming model (i.e., [11,12,13,14]) have been introduced to address various real-world problems. The application of fuzzy set and possibility theories to decision-making allows decisions based on imprecise information. In linear programming, possibility theory can be used to include imprecise parameters in the problem formulation [15]. It is evident that fuzzy theories and other related theory has made successful to work on uncertainty problem in modeling and dealing with real world problem.

In this study, a possibility programming multi-objective model is developed on the basis of modality optimization which produces a fractional linear programming problem. A flexible normalized weighting method among the objective functions is used to address the problem of determining the weight of the objective functions. The normalization scheme uses the differences of optimal function values in the maximal and minimal points, and enables us to obtain the weight for fuzzy objective. Weighting method for the objective function is incorporated which is induced from the individual minimal and maximal solutions. Meanwhile, the use of a necessity measure to the model is useful for encapsulating the decision maker's attitude and preferences under uncertain conditions. The preservation of the uncertainties in the mathematical programming model is important to ensure that the developed model conveys the decision maker's intention appropriately. The rest of the paper is organized as follows. Section 2 introduces a formulation of possibilistic multi-objective problem model. Section 3 describes weighting method to the formulated possibilistic multi-objective problem. Section 4 summarizes the solution procedure and is followed by a numerical experiment example. Section 6 concludes the section.

## 2 Formulating Fuzzy Possibilistic Multi-objective Problem Model

Possibilistic programming describes the type of fuzzy programming produced if the vagueness in the decision maker's aspiration is modeled as an objective function using a fuzzy preference relation and the ambiguities in the coefficients are represented in terms of a possibility distribution [16]. General multiple objective possibilistic problem is then formulated as the following model:

$$\begin{aligned} \max \quad & \sum_{i=1}^{p} h_i^{\eta} \\ \text{subject to}: \quad & \sum_{i=1}^{p} \left( \frac{\sum_{j=1}^{n} x_j \alpha_j}{\sum_{j=1}^{n} |x_j| d_j} \right) \geq h \end{aligned} \tag{1}$$

From (1), the following description describes the objective function and the constraints:

Let $\mathbf{v}^{\eta} \in [0,1]^m$ be a necessity aspiration degree that a decision maker is aspired to achieve certainly. Constraints $\overline{\tilde{\mathbf{A}}}\mathbf{x} \leq \overline{\tilde{\mathbf{b}}}$ is formulated using necessity measurement as follows:

$$Nec\left(\left(\overline{\tilde{\mathbf{A}}}\mathbf{x} \leq \overline{\tilde{\mathbf{b}}}\right) \geq \mathbf{v}^{\eta}\right) \tag{2}$$

The inequalities sign is determined by the decision maker's decision i.e. to minimize or to maximize. The symmetric fuzzy number is written as $A = \left\langle \sum_{j=1}^{n} x_j a_j, \sum_{j=1}^{n} |x_j| d_j \right\rangle$. From (2), let us assume that $s$ is less than $v^{\eta}$, and we obtain,

$$s = \sum_{j=1}^{n} x_j a_j + v^{\eta} \left( \sum_{j=1}^{n} |x_j| d_j \right) \tag{3}$$

Thus, expression (3) is the treated constraint which considers the certainty degree of decision maker's intention to the problem constraint.

In this model, decision maker's target value which contains fuzzy value is included in the objective function and is treated as a constraint by using a modality approach.

Let us consider that the decision maker wants to maximize the certainty degree that the event is larger than $g^{\eta}$, and is modeled as maximizing $Nec\left(\alpha x \geq g^{\eta}\right)$.

Using additional variable $h$, the following model expresses the decision maker's aspiration.

$$\begin{aligned} \max \quad & h \\ s.t \quad & Nec\left(\left(\alpha x \geq g^{\eta}\right) \geq h\right) \end{aligned} \tag{4}$$

A fuzzy-possibilistic multi-objective programming problem (FR-PPP) model is written by using the treated constraints (3) and objectives (4) as follows:

$$\begin{aligned} \text{max} \quad & \sum_{i=1}^{p}\left(\frac{\sum_{j=1}^{n} x_j \alpha_j}{\sum_{j=1}^{n}|x_j| d_j}\right) \\ \text{subject to:} \quad & \sum_{i=1}^{p}\sum_{j=1}^{n} x_j a_j + v^{\eta}\left(\sum_{j=1}^{n}|x_j| d_j\right) \\ & x_i \ge 0 \end{aligned} \tag{5}$$

Problem (5) is a linear fractional programming problem with multiple objectives and can be solved by using linear program. Fractional programming solution is important as various problems consider the optimization of a ratio between physical and/or economic linear functions [20]. In this case, the modality optimization takes the advantages of fractional programming in finding the problem's solution.

## 3 Weighting Method for Fractional Programming Problem

The approach to determine the weight value is to use the maximal value and minimal value of each objective function under consideration. Let us consider relative importance $w_i$ for $\mu_i^{N_i} = C_j^{N_i}$ and $w_i^{'}$ for $\mu_i^{D_i} = C_j^{D_i}$ such that $w_i > 0, w_i^{'} > 0$ and $\sum_{i=1}^{p} w_i + w_i^{'} = 1$. The compatibility of a value of $j$ of $P \le \left(N_i x, N_i^0\right)$ may be given as in [19]. Thus we obtain the simple additive weighting model to solve the multi-objective linear fractional programming problem (5) as follows:

$$\begin{aligned} opt \quad & V(\mu) = \sum_{i=1}^{p}\left(w_i \mu_i^{N_i} + w_i^{'} \mu_i^{N_i}\right), \\ \text{such that :} \quad & \mu_i^{N_i} = C_j^{N_i}, \mu_i^{D} = C_j^{D}, \\ & Ax \le b, \mu_i^{N_i} \le 1, \mu_i^{D_i} \le 1, \\ & \mu_i^{N_i} \ge 0, \mu_i^{D_i} \ge 0, x \ge 0, i = 1, \dots, p, \end{aligned} \tag{6}$$

where $V(\mu)$ is the achievement function.

The best maximal point, $f_i^{+}$, is the point of the objective function space whose coordinates are equal to the maximum that can be achieved by each objective function in the feasible region, $X$ .

The weight $w_i^{*}$ is calculated as follows:

$$w_i^{*} = \frac{1}{f_i^{+} - f_i^{-}} \tag{7}$$

The normalized weights for $w_i^{N_i}$ and $w_i^{D_i}$ are as follows:

$$w_i^{N_i} = \frac{w_i^{N_i^*}}{\sum_{i=1}^{p} w_i^{N_i}} 0.5, \quad w_i^{D_i} = \frac{w_i^{D_i^*}}{\sum_{i=1}^{p} w_i^{D_i}} 0.5 \tag{8}$$

where $w_i^{N_i} + w_i^{D_i} = 1$.

$w_i^*$ represent the weights for the membership functions of the decision vectors. The scheme expressed by Equation (8) is used to obtain the normalized weight for fuzzy objective for multi-objective linear fractional problem model (5). This normalization scheme provides the best normalization results as we normalize the objective functions by the true intervals of their variation over the Pareto optimal set. Problem (8) can be solved using linear fractional programming problem [20]. Problem (8) can be extended to include the possibility as well.

## 4 Solution Procedure

A Multi-objective Fuzzy Possibilistic evaluation scheme is explained as follows:

### 4.1 Problem Description and Modeling

Describe the problem, determines parameter values, objective, constraints and decision maker's target. Readers are referred to the paper [21,22] for additional explanation on building the fuzzy random based model.

### 4.2 Building Multi-objective Fuzzy Possibilistic Model

- Analyze and treat the constraint as Equation (2). Set the degree of certainty $v^{\eta}$ and transform the constraints as Expression (3).
- Analyze and treat the objective function by determining the necessity aspiration level $g^{\eta}$ of the objective function. For the case of decision maker wants to maximize the objective function, transform the objective function to the expression of maximizing $Nec(\alpha x \geq g^{\eta})$.
- Develop a multi-objective possibilistic programming model as Equation (5).

### 4.3 Solving Multi-objective Fuzzy Possibilistic Model

Solve the model (5) using linear fractional approach.

- Calculate minimum and maximum individual solution $N_i^0 = \max_{x \in X} N_i x, \quad D_i^0 = \min_{x \in X} D_i x$
- Choose the threshold $p_i$ and $s_i$ indicating the appropriate closeness to $N_i^0$ and $D_i^0$ for all $i$.
- Elicit the weighs, $w^{N_i}$ and $w^{D_i}$.
- Solve problem model (8) to obtain the solution $\mathbf{x}$.

After proposing the fractional model, the error assessment for the results [23], from this model should be considered because of the estimated value of the coefficient for decision parameter. Hence an error assessment is provided, which is called the weight absolute percentage error for fuzzy decision (WAPE-FD) and give the definition as follows.

**Definition 1.** Weight Absolute Percentage Error for Fuzzy Decision (WAPE-FD)

The weight absolute percentage error for fuzzy decision (WAPE-FD) is a measure of accuracy of a method for constructing fitted values in statistics, specifically in trend estimation. It usually expresses accuracy as a percentage, and is defined by the formula:

$$\mathrm{WAPE} - \mathrm{FD} = 100\% * \left| \frac{\mathrm{WE_k} - \mathrm{EWE_k}}{\mathrm{WE_k}} \right|, \tag{2}$$

where $\mathrm{WE_k}$ is the actual weight expected value and $\mathrm{EWE_k}$ is the estimated weight expected value.

The weight absolute value in this calculation is summed for every fitted or estimated point and multiplying by 100 makes a percentage error.

## 5 Numerical Example

We demonstrate the use of the proposed method on a model of a crops industry. The problem is modeled as multi-objective problem with co-exist uncertainties. The uncertainties are characterized as follow:

*The historical data are used to estimate the model coefficient. The aspiration of the decision maker target(s) contains vagueness and the fuzzy value used in the model coefficient contains ambiguity. That is, the manager is unable to precisely determine the target value. The manager is aspiring to maximize the production volume together with maximizing the profit to some target value. Hence, the problem results in Possibilistic Multi-objective Problem with Fuzzy coefficient.*

The crop production planning problems are investigated with two decision variables and two functional objectives under four system constraints. The problem is then modeled as follows:

$$\begin{array}{l} \text{satisfy :} \\ \left.\begin{array}{l}\text{maximize profit} : Z_1 = \langle 0.860, 0.100\rangle x_1 + \langle 1.100, 0.100\rangle x_2, \\ \text{maximize production} : Z_2 = \langle 1.126, 0.020\rangle x_1 + \langle 0.000, 0.000\rangle x_2\end{array}\right\}(a) \\ \text{subject to :} \\ \left.\begin{array}{l}\text{raw material} : F_1 = \langle 3.75, 0.06\rangle x_1 + \langle 0.91, 0.08\rangle x_2 \le 87.75, \\ \text{labor} : F_2 = \langle 0.65, 0.55\rangle x_1 + \langle 0.90, 0.09\rangle x_2 \le 4.42, \\ \text{mills} : F_3 = \langle 17.35, 0.85\rangle x_1 + \langle 2.16, 0.27\rangle x_2 \le 95.20, \\ \text{capital} : F_4 = \langle 0.87, 0.65\rangle x_1 + \langle 0.98, 0.65\rangle x_2 \le 20.15,\end{array}\right\}(\mathrm{b}) \end{array} \tag{9}$$

The solution procedure mentioned in the Section 4 is used to solve the problem. Let us assume that the decision maker decides that certainty degree is larger $v^{\eta} = 0.7$ for the system constraint (b) in the Problem (9). The constraints under expression (2) and (3) are used to transform the constraints based on decision maker aim as follows:

$$\left.\begin{array}{l} Nec\big(\big(\langle 3.75, 0.06\rangle x_1 + \langle 0.91, 0.08\rangle x_2 \le 87.75\big) \ge v^{\eta}\big), \\ Nec\big(\big(\langle 0.65, 0.55\rangle x_1 + \langle 0.90, 0.09\rangle x_2 \le 4.42\big) \ge v^{\eta}\big), \\ Nec\big(\big(\langle 17.35, 0.85\rangle x_1 + \langle 2.16, 0.27\rangle x_2 \le 95.20\big) \ge v^{\eta}\big), \\ Nec\big(\big(\langle 0.87, 0.65\rangle x_1 + \langle 0.98, 0.65\rangle x_2 \le 20.15\big) \ge v^{\eta}\big), \end{array}\right\}(\mathrm{b}) \tag{10}$$

Say that the decision maker aims to maximize the certainty degree of profit is larger than 5.0 million dollars, and to maximize the certainty degree of production volume is larger than 5.2 million tones. According to (4), the decision maker's aspirations are modeled as follows:

$$\begin{array}{l} Nec(0.86x_1 + 1.10x_2 \ge 5.0) \\ Nec(1.126x_1 + 0.00x_2 \ge 5.2) \end{array} \tag{11}$$

The constraints and the objective functions are now treated using necessity measure and prepared for the possibilistic multi-objective solution. The problem (9) is then rewritten by considering (10) and (11) as follows:

$$\begin{array}{ll} \max & \left\{\dfrac{0.86x_1 + 1.1x_2 - 5.0}{0.1x_1 + 0.1x_2}, \dfrac{1.126x_1 + 0x_2 - 5.2}{0.02x_1 + 0x_2}\right\} \\ \text{subject to:} & 3.79x_1 + 0.95x_2 \le 87.75, \\ & 1.03x_1 + 0.96x_2 \le 4.42, \\ & 17.94x_1 + 2.34x_2 \le 95.20, \\ & 1.32x_1 + 1.43x_2 \le 20.15, \\ & x_i \ge 0. \end{array} \tag{12}$$

Based on the third step listed in the Section 4, parameter values for the first and second objective function are computed as follows: $P^0_{G_1} = 0.3$, $S^0_{G_1} = 2.0$, $N^0_{G_1} = 0.75$ and $D^0_{G_1} = 0.0$, $P^0_{G_2} = 0.6$, $S^0_{G_2} = 3.0$, $N^0_{G_2} = 0.97$ and $D^0_{G_2} = 0.0$.

**Table 1.** Individual Optimal Solutions and Normalized Weight

| Item | $N_{G_1}$ | $N_{G_2}$ | $D_{G_1}$ | $D_{G_2}$ |
|---|---|---|---|---|
| minimal $g_{ij}$ | -5.000 | -5.200 | 0.000 | 0.000 |
| maximal $\mu_{ij}$ | 0.753 | 0.978 | 0.639 | 0.109 |
| weight $w^*$ | 0.173 | 0.161 | 1.564 | 9.174 |
| Normalized weight $w^*$ | 0.258 | 0.241 | 0.072 | 0.427 |

The equivalent linear programming for problem (12) is as follows:

$$
\begin{array}{lll}
\text{max} & : & V(\mu) = 0.258\,\mu_{N_1} + 0.241\,\mu_{N_2} + 0.072\,\mu_{D_1} + 0.127\,\mu_{D_2} \\
\text{subject to} & : & 0.45\,\mu_{N_1} \text{ - } 0.860\,x_1 - 1.100\,x_2 = \text{-}5.3, \\
 & & 0.37\,\mu_{N_{21}} \text{ - } 1.126\,x_1 + 0.000\,x_2 = \text{-}5.8, \\
 & & 2.00\,\mu_{D_1} + 0.100\,x_1 + 0.100\,x_2 = 2, \\
 & & 3.00\,\mu_{D_2} + 0.020\,x_1 + 0.000\,x_2 = 2, \\
 & & \text{system constraint s (b)}, \\
 & & \mu_{N_i} \leq 1, \mu_{D_i} \leq 1, \mu_{N_i} \geq 0, \mu_{D_i} \geq 0, x_i \geq 0.
\end{array}
\tag{13}
$$

The results of the industrial problem case are explained. In this study, fuzzy regression method is used to estimate the coefficient values [21,22] which help development of the initial model. In this case, the previous patterns of outcomes are included in the future prediction or decision.

In the industrial Problem (13), the objective function and the constraints were retreated using necessity measure to illustrate decision maker's intention so as to make the mathematical programming for the respective model is as close as a decision maker's aim. The optimal solution of the problem Model (13) is $(x_1, x_2) \approx (5.35,\ \ 1.03)$ whose objective value is $V(\mu) = 0.87$, with $\mu_{N_1} = 1.00$, $\mu_{N_2} = 0.62$, $\mu_{D_1} = 0.68$ and $\mu_{D_2} = 0.96$. Since the proposed method offers iterative and satisficing solution, decision maker may be able to reconstruct the problem and considers his new aspiration. From that, it is remarkable that the proposed method may be able to produce various solutions depending on the decision maker's aim. The interactive system can be useful to verify the obtained solution with the decision maker's intention. To check the sensitivity of the results, the proposed WAPE_FD can be utilized so as to compare the non fuzzy and fuzzy weight value. The WAPE_FD reminded us how much risk we should consider in this problem. It also could clarify the results between fuzzy decision and non-fuzzy decision that show us how far between both of them. Therefore, the results could be described that the fuzzy results were differentiate from the crisp number's results in this problem.

## 6 Conclusion

Multiple objectives are combined into a single objective to solve a multi-objective problem and yield a best compromise solution. This can be prepared by assigning appropriate weighting which represent the relative importance of the objectives. A normalized weighting scheme is provided in the model explained in this paper that is elicited by using the individual best maximal and minimal solutions for each objective function. In this study also, we emphasize the role of the necessity measure in expressing the decision maker's aims and accurately determining the value of each objective function. We have examined the possibilistic programming problem and multi-objective possibilistic programming problem through a modality perspective, which uses a linear fractional approach to obtain the solution. The proposed method can be repeated iteratively and various solutions can be obtained, as bounded by the aims of the decision maker.

**Acknowledgement.** The author expresses her appreciation to the University Tun Hussein Onn Malaysia (UTHM) and Research Acculturation Grant Scheme (RAGS) Vot R044.

## References

1. Saaty, L.: The analytic hierarchy process: planning, priority setting, resource allocation. McGraw-Hill, New York (1980)
2. Marler, R.T., Arora, J.S.: Weighted sum method for multi-objective optimization: new insights. Structural and Multidisciplinary Optimization 41(6), 853–862 (2010)
3. Jubril, A.M.: A nonlinear weights selection in weighted sum for convex multiobjective optimization. Facta universitatis - series: Mathematics and Informatics 27(3), 357–372 (2012)
4. Moller, B., Graf, W., Beer, M.: Safety assessment of structures in view of fuzzy randomness. Computers and Structures 81(15), 1567–1582 (2003)
5. Zeleny, M.: The pros and cons of goal programming. Computers and Operations Research 8(4), 357–359 (1981)
6. Inuiguchi, M., Sakawa, M.: Possible and necessary efficiency in possibilistic multiobjective linear programming problems and possible efficiency test. Fuzzy Sets and Systems 78(2), 231–241 (1996)

7. Bouyssou, D.: Modeling inaccurate determination, uncertainty, imprecision using multiple criteria. In: Lockett, A.G., Islei, G. (eds.) Improving Decision Making in Organizations. Lecture Notes in Economics and Mathematical Systems, vol. 335, pp. 78–87. Springer (1989)
8. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 338–353 (1965)
9. Dantzig, G.: Linear programming under uncertainty. Management Science 1(3-4), 197–206 (1955)
10. Sengupta, J.K.: Stochastic goal programming with estimated parameters. Journal of Economics 39(3-4), 225–243 (1979)
11. Zimmermann, H.-J.: Description and optimization of fuzzy systems. International Journal General Systems 2(4), 209–215 (1976)
12. Zimmermann, H.-J.: Fuzzy programming and linear programming with several objective functions. Fuzzy Sets System 1(1), 45–55 (1978)
13. Inuiguchi, M., Ichihashi, H., Tanaka, H.: Fuzzy programming: A survey of recent developments. In: Slowinski, R., Teghem, J. (eds.) Stochastic Versus Fuzzy Approaches to Multiobjective Mathematical Programming Under Uncertainty, pp. 45–68. Kluwer Academics, Dordrecht (1990)
14. Sakawa, M.: Fuzzy sets and interactive multi-objective optimization. Applied Information Technology. Plenum Press, New York (1993)
15. Julien, A.: An extension to possibilistic linear programming. Fuzzy Sets and Systems 64(2), 195–206 (1994)
16. Inuiguchi, M., Ramik, J.: Possibilistic linear programming: a brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem. Fuzzy Sets and Systems 111(1), 3–28 (2000)
17. Kornbluth, J., Steuer, R.: Multiple objective linear fractional programming. Management Science 27(9), 1024–1039 (1981)
18. Luhandjula, M.K.: Fuzzy approaches for multiple objective linear fractional optimizations. Fuzzy Sets and Systems 13(1), 11–23 (1984)
19. Zadeh, L.A.: Optimality and non-scalar-valued performance criteria. IEEE Transactions on Automatic Control 8(59), 59–60 (1963)
20. Charnes, A., Cooper, W.: Programming with linear fractional functions. Naval Research Logistics Quarterly 9(3-4), 181–186 (1962)
21. Nureize, A., Watada, J.: Linear Fractional Programming for Fuzzy Random based Possibilistic Programming Problem. International Journal of Simulation Systems, Science & Technology 14(1), 24–30 (2014)
22. Arbaiy, N., Watada, J.: Constructing Fuzzy Random Goal Constraints for Stochastic Fuzzy Goal Programming. In: Huynh, V.-N., Nakamori, Y., Lawry, J., Inuiguchi, M. (eds.) Integrated Uncertainty Management and Applications. AISC, vol. 68, pp. 293–304. Springer, Heidelberg (2010)
23. Lin, P.-C., Wu, B., Watada, J.: Kolmogorov-Smirnov Two Sample Test with Continuous Fuzzy Data. Integrated Uncertainty Management and Applications 68, 175–186 (2010)

# Intelligent Double Treatment Iterative Algorithm for Attribute Reduction Problems

Saif Kifah, Salwani Abdullah, and Yahya Z. Arajy

Data Mining and Optimisation Research Group (DMO)
Centre for Artificial Intelligence Technology
Universiti Kebangsaan Malaysia, 43600 Bangi Selangor, Malaysia
`{Saifkc,yahya.arajy}@gmail.com, salwani@ftsm.ukm.my`

**Abstract.** Attribute reduction is a combinatorial optimization problem in data mining that aims to find minimal reducts from large set of attributes. The problem is exacerbated if the number of instances is large. Therefore, this paper concentrates on a double treatment iterative improvement algorithm with intelligent selection on composite neighbourhood structure to solve the attribute reduction problems and to obtain near optimal reducts. The algorithm works iteratively with only accepting an improved solution. The proposed approach has been tested on a set of 13 benchmark datasets taken from the University of California, Irvine (UCI) machine learning repository in line with the state-of-the-art methods. The 13 datasets have been chosen due to the differences in size and complexity in order to test the stability of the proposed algorithm. The experimental results show that the proposed approach is able to produce competitive results for the tested datasets.

**Keywords:** Attribute Reduction, Composite neighbourhood structure, Iterative improvement algorithm, Rough set theory.

## 1 Introduction

Rough set theory attribute reduction is considered as one of the phenomena in data mining, which is known as NP-hard optimisation problem. The objective is to reduce the amount of the data from a large dataset. Applying attribute reduction process on a large dataset increases the quality of knowledge discovery for certain needs. Moreover, the reduction processes in the datasets relies totally on the correlation between the attributes that support a better representation of the data. Therefore, it avoids the disposal of information loss. Data mining has become an interesting field of study in extracting useful information from a large dataset, and attribute reduction is one of its basic issues.

Data preparation from the reduction procedure supports the mining techniques by reducing the search area during the mining process. This technique reduces the time taken to obtain important data. Recently, the literature on search methodologies for attribute reduction reports the implementation of stochastic algorithmic approaches

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems*,
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_10

such as, simulated annealing [1], ant colony [2,3], genetic algorithm [1], tabu search [4], scatter search [5], great deluge [6,7], and variable neigbourhood search [8].

The objective of the proposed work in this paper is to employ an iterative improvement algorithm embedded with composite neighbourhood structures to increase the data precision. The iterative improvement algorithm is a meta-heuristic approach for solving combinatorial optimization problems. The composite neighbourhood structures are adopted to avoid the ineffectiveness of using only single neighbourhood structure [9]. An intelligent selection mechanism is integrated in the algorithm to enhance the selection that occurs between the neighbourhoods. Additionally a double treatment is applied to accelerate the convergence of the solution by escaping from the ineffectual search space to another. This technique aids the exploration of the search space in order to obtain near optimal results.

The remainder of this paper is organised as follows: Section 2 illustrates the calculation of the dependency degree in rough set theory. Section 3 describes the iterative improvement algorithms with the necessary requirements and highlights the proposed algorithm, i.e. a double treatment iterative improvement algorithm with intelligent selection of composite neighbourhood structures. Section 4 provides the results obtained, followed by statistical analysis, and Section 5 provides the conclusion and recommended future work.

## 2 Rough Set Theory

Rough Set Theory (RST) is an effective mathematical tool to analyse the vague information of an object. Dependency degree rough set theory is adopted as a quality measurement tool to evaluate the importance of the near optimal reducts. RST has become an essential technique to handle all attribute reduction, knowledge discovery, pattern recognition, image processing and data mining problems. The theory was introduced by Pawlak [10,11] to deal with two arbitrary subsets of universe called lower and upper approximations.

**Table 1.** Example of dataset

| U\W | a | b | c | d |
|---|---|---|---|---|
| **u1** | 0 | 1 | 2 | 0 |
| **u2** | 0 | 0 | 0 | 0 |
| **u3** | 1 | 2 | 1 | 2 |
| **u4** | 2 | 0 | 0 | 1 |
| **u5** | 1 | 1 | 2 | 1 |
| **u6** | 0 | 2 | 1 | 2 |

The dataset represented in Table 1 is a two dimensional array, where it consists of three attributes, one decision attribute and six objects. Let U represent the universe and W represents a set of attributes. Next let C represents the set of condition attributes {a, b, c} and D represents the set of decision of attributes (d). Hence, $C \subset W$, $C \cup D=W$, and $C \cap D= \phi$. The entry in column u and row w has the value f(u,w), which will then define an equivalence relation over U. Given x, the universe can be partitioned into a set of disjoint subset:

$$R_q = \{x : x \in U \wedge f(u,w) = f(u_0,w) \forall u_0 \in U\} \tag{1}$$

For example, let us assume U= {u1, u2, u3, u4, u5, u6}, W= {*a, b, c, d*}, C= {*a*, *b*, *c*} and D= {*d*}. The main idea for rough set theory is applying the concept of discernibility. Let I= (U, W) be an information system, where U is a non-empty set of a finite object (the universe) and W is a non-empty finite set of attributes such that *a*: U $\rightarrow \forall e$ for every $e \in$ W. $\forall e$ represents the value of an attribute *a*. For any P $\subset$ W there is a connected equivalence relation IND(P):

$$IND(P) = \{(u,w) \in U^2 \mid \forall e \in P\, a(u) = a(w)\} \tag{2}$$

If $(u,w) \in$ IND(P), then *u* and *w* are indiscernible by the attributes in Table 1. While the equivalence classes of the P indiscernibility relation are denoted as $[u]_p$. For example, let P= {*b*, *c*}, then objects u1 and u5 are indiscernible, objects u2 and u4 are indiscernible and so are u3 and u6. Let Q={d}, then objects u1 and u2 are indiscernible, u3 and u6 are also indiscernible, and so are u4 and u5.

*U=IND*(*P*) ={{u1,u5}, {u2,u4}, {u3,u6}}.
*U=IND*(*Q*) ={{u1,u2}, {u3,u6}, {u4,u5}}.

Letting X U, X can be calculated approximately using only the information contained within P by constructing the P-lower approximations of X:

$$\underline{P}X = \{x \mid [x]P \subseteq X\}| \tag{3}$$

Let P and Q be an equivalence relation over U, then the positive regions can be defined as:

$$POSp\ (Q) = \bigcup_{x \in U\,/\,Q} \underline{P}X \tag{4}$$

The positive region contains all objects in the datasets of U that can be classified as classes of U/Q, and this information will be used in all of the attributes P. For example, let P = (b, c) and Q = (d).

POSP(Q) = {Ø,{u3,u6},Ø}= {u3,u6}.

This exculpate all of objects u3 and u6 that are certainly to be classified as discernible, which belong to a class in attribute d, when considering attributes b and c. Other objects that lack information to make them discernible will not be classified. An important issue in the data analysis is to discover dependencies between attributes. The dependency degree is measured by the following equation:

$$k = \gamma_{\{b,c\}} p(Q) = \frac{|\mathrm{POS}_P(\mathrm{Q})|}{|\mathrm{U}|} \tag{5}$$

If k = 1, Q depends totally on P, if 0<k<1, Q depends k=*y* partially on P, and if k = 0 then Q does not depend on P. In the example above, the dependency degree of decision {d} from the attributes {b, c} is calculated as:

$$k=\gamma_{\{b,c\}}(\{d\})=\frac{|\mathrm{POS}_{\{b,c\}}(\{\mathrm{d}\})|}{|\mathrm{U}|}=\frac{|\{u3,u6\}|}{|\{u1,u2,u3,u4,u5,u6\}|}=\frac{2}{6}$$

A reduct is defined as a subset of minimal Cardinality $R_{\min}$ of the conditional attribute set C such $y^{R(D)}$- $y^{C(D)}$.

$$R=\{X : X \subseteq C, \gamma_{X(D)} = \gamma_{C(D)}\} \tag{6}$$

$$R_{\min}=\{X : X \in R, \forall Y \in R, |X| \leq |Y|\} \tag{7}$$

The intersection of all the sets in $R_{\min}$ called the Core:

$$Core(R) = \bigcap_{X \in R} X \tag{8}$$

The elements of the core are the attributes that cannot be eliminated without introducing more contradiction to the dataset.

Using the dataset in Table 1, the dependency degree D = $\{d\}$ on all possible subsets of C can be calculated as:

$$\gamma\{a\}=\frac{1}{6}\ ,\ \gamma\{b\}=\frac{2}{6}\ ,\ \gamma\{c\}=\frac{2}{6}\ ,\ \gamma\{a,b\}=\frac{6}{6}\ ,\ \gamma\{a,c\}=\frac{6}{6}\ ,\gamma\{b,c\}=\frac{2}{6}\ ,\ \gamma\{a,b,c\}=\frac{6}{6}.$$

The minimal reducts obtained in this example: Rmin= {a,b} or {a,c}.

Finding minimal reducts is classified as a NP-hard problem. The procedure of calculating all the possible reducts (Core(R)) is a time consuming process. Therefore, researchers try to adapt several heuristic algorithms to find approximate solutions for this problem.

## 3 Composite Neighbourhood Structures

A composite neighbourhood structure (CNS) colligates two or more neighbourhood search operations. The advantage of combining several neighbourhood search operations aids to recompense the insufficiency of employing a single type of isolated neighbourhood search operation while exploring a search space [9, 12].

The CNS has been employed within the tabu search algorithm to solve the preventive maintenance scheduling problem [13]. Three neighbourhood structures (add, drop and swap) were conducted in their research. The add move works only on a feasible solution, while the drop move selects a certain task to be removed to maintain the feasibility of the solution. The swap move is employed when two moves (add and drop) are not able to satisfy the hard constraints imposed on this problem. The results obtained show that the CNS effectively improves the algorithm.

CNS was employed in tabu search algorithm to solve the problem of scheduling two-machine pre-emptive open shops in order to minimise total tardiness [12]. Two neighbourhood structures (insertion and swap) are employed in the search operation. The insertion neighbourhood works by moving one job from a certain position and inserting it into another one. The swap structure consists of exchanging the positions of two jobs. This exchange takes place when a specified number of iterations ends without any improvement. The tabu search employs both alternating neighbourhood structures and a dynamic tabu list. Due to the hybridisation of the tabu search with CNS, the algorithm was able to obtain high quality solutions for large sized problems within a fair amount of time.

The hybrid neighbourhood structures with tabu search algorithm for parallel machine total tardiness problems was applied by Bilge et al. [14]. The insert move and pairwise exchanges (swaps) are the two neighbourhood structures used. The insertion move simply locates two particular jobs and places the first job into the position that precedes the second one. Whilst, the swap move place each job in the location previously occupied by other jobs. This procedure can be considered as a combination of two insert moves which is referred as a "hybrid structure". It can be seen from their results that the adopted strategy was very successful in isolating desirable regions of the neighbourhood. This leads to the increase in the speed of the search and improves the quality of the solution.

A randomised iterative improvement algorithm for course timetabling problem was proposed in [15]. The algorithm works by applying a set of neighbourhood structures on the current solution. A better solution is always accepted and the worse solution is accepted with a certain probability which is based on the exponential Monte Carlo algorithm. By employing CNS procedure, the algorithm was able to obtain good results when compared to other available approaches in the literature

## 4 Iterative Improvement Algorithms with Composite Neighbourhood Structures

Three variations of improvement algorithms are carried out in this research. Each algorithm performs a different mechanism to solve the attribute reduction problem. The first algorithm is the Basic Composite Neighbourhood Structure which is coded as Basic-CNS. This algorithm employs two or more neighbourhood structures on the initial solution. The selection of the neighbourhoods is randomly generated. The second algorithm is the Intelligent Selection Composite Neighbourhood Structures coded as IS-CNS. This algorithm employs two lists of neighbourhood structures. The selection of the neighbourhood structures between the two lists is based on certain rules, which are created to improve the selection of the neighbourhood structures to suite the attribute reduction problem based on the current solution in hand. The third algorithm is the Double Treatment Intelligent Selection with Composite Neighbourhood Structures which is coded as IS-DT-CNS. This algorithm uses five lists of neighbourhood structures, where only two lists are applied in the first phase (constructive phase), while in the second phase (improvement phase), all the five lists are employed.

### 4.1 Initial Solution Generation and Representation

The initial solution is generated using a constructive heuristic method that is randomly obtained by distributing ones and zeros into one dimensional array (where the size of the array is equal to the number of attributes). The cell with the value of “1” indicates that the attribute is selected, on the other hand, the cell with the value of “0” means that the attribute is discarded.

### 4.2 Neighbourhood Structures

The neighbourhood structures used in our approach are outlined as follows:

**N1:** Delete one attribute at random - choose one attribute at random to be deleted from the current solution.

**N2:** Intelligently delete one attribute - delete one attribute of a lowest priority according to the priority list from the current solution.

**N3:** Delete two attributes - select two attributes at random to be deleted from the current solution.

**N4:** Delete three attributes - randomly pick three attributes to be removed.

**N5:** Add one attribute - select one attribute randomly in order to be appended to the current solution.

**N6:** Intelligently add one attribute - select one attribute of a highest priority to be inserted to the current solution based on the priority list.

**N7:** Add two attributes - select two attributes at random to be inserted to the current solution.

**N8:** Add three attributes - select three attributes at random to be added to the current solution.

**N9**: Randomly delete 5% of the attributes - calculate this ratio from the current solution and select one attribute at random to be removed from the current solution.

**N10:** Randomly delete 10% of the attributes - the 10% ratio is calculated from the attributes in the current solution in order to be removed.

**N11:** Randomly delete 20% of the attributes - remove 20% of the attributes in the current solution at random.

**N12**: Normal swap - select one attribute at random from the current solution and another attribute from the deleted list and swap between them.

**N13:** Intelligent swap - select one attribute with the lowest priority from the current solution and select one attribute with the highest priority from the deleted list and swap between them.

The intelligent neighbourhood structure adapts itself by working according to the priority list. This list depends on the variation of the objects in each attribute, which is different from one data set to another. The more variation leads to the highest priority, while the less variation leads to the lowest priority. Moreover, if all of the attributes have the same priority then the selection of the neighbourhood will change from intelligent selection (see Section 4.4) to a random selection (that selects a neighbourhood structure at random).

### 4.3 Solution Quality Measurement

The measurement of the solution quality is based on the dependency degree in the rough set theory. The optimal value for the dependency degree is equal to one. Given two solutions i.e. an initial solution Sol and a trial solution Sol*. The trial solution Sol* is accepted if there is an improvement in the dependency degree (i.e. if $\gamma$ (Sol*) > $\gamma$ (Sol)). If the dependency degree for both solutions is the same (i.e. $\gamma$ (Sol*) = $\gamma$ (Sol)) then the solution with the less number of attributes (denoted as #) will be accepted.

### 4.4 Intelligent Selection

The intelligent selection (IS) is a mechanism to select an effective combination of neighbourhood structures to accelerate the convergence of a solution. The intelligent selection is based on a certain number of rules (that are formed after some preliminary experiments where we have found that it is highly probable for certain combination of neighbourhood structures to be less effective in reaching the minimal reducts). The intelligent selection is applied in the second algorithm (IS-CNS) by creating two lists (coded as List-1 and List-2) that contain the same neighbourhood structures as discussed above except for N4 and N8. These two neighbourhoods are being excluded due to the large changes caused on the solution (with respect to the dependency degree) which is not required in the second phase of the algorithm where the intelligent selection takes place. Note that in the second phase, we are interested to reduce the number of attributes while maintaining the dependency degree equal to 1. The intelligent selection has also been adopted in the third algorithm (IS-DT-CNS). In IS-DT-CNS, five lists have been applied (i.e. List-1, List-2, List-3, List-4 and List-5). List-1 and List-2 contain all the neighbourhood structures. While List-3 = {N9, N10}, List-4 = {N1, N5, N12, N13} and List-5 = {N2, N6}. The different lists of neighbourhood are created in order to allow the algorithm to have a better and different way of exploring on the search space. This will avoid the algorithm from getting stuck in local optima. The rules that are employed in selecting the appropriate combination of neighbourhood structures from List-1, List-2, List-3, List-4 and List-5 are listed as follows:

$IS_1$: Rule 1: if (number of attribute of a trial solution < 4 and dependency degree!=1) then List-1($N_8$) → List-2(select any neighbourhood structures at random).
$IS_2$: Rule 2: if List-1($N_1$) → List-2($N_9$).
$IS_3$: Rule 3: if List-1($N_6$) → List-2($N_1$).
$IS_4$: Rule 4: if List-1($N_5$) → List-2($N_2$).
$IS_5$: Rule 5: if List-1($N_7$) → List-2($N_3$).
$IS_6$: Rule 6: if List-1($N_2$) → List-2($N_9$).
$IS_7$: Rule 7: if (List-1($N_{12}$) or List-1($N_{13}$)) → List-2 (randomly select any neighbourhood structure except “add” neighbourhood structures).
$IS_8$: Rule 8: if (List-1($N_9$) or List-1($N_{10}$) or List-1($N_{11}$)) → List-2($N_1$).

$IS_9$: Rule 9: if List-1($N_3$) → List-2($N_5$).
$IS_{10}$: Rule 10: if List-1($N_{12}$) → List-2($N_9$).
$IS_{11}$: Rule 11: if List-4($N_1$) → List-5($N_6$).
$IS_{12}$: Rule 12: if List-4($N_5$) → List-5($N_2$).

Note that, the rules employed here depend on the number of attributes and the quality (i.e. dependency degree) of the current solution. Any changes in terms of the number of attributes and/or the dependency degree cause different rules to be fired.

### 4.5 First Algorithm: Basic Iterative Improvement Algorithm with Composite Neighbourhood Structures (Basic-CNS)

The first algorithm (Basic-CNS) employs thirteen neighbourhood structures with a variant specifications embedded into one list in order to suit to the reduction procedure. The composite procedure requires a combination of two or more neighbourhood structures to be applied in the algorithm. The selection between the neighbourhood structures in Basic-CNS is based on a random mechanism. The quality of the solution is measured based on the explanation in Section 4.3. The improved solution is always being accepted. The search process continues until the termination condition is met. In this work, the termination condition is set to be the maximum number of iterations which is 250 iterations. The overview of Basic-CNS is presented in [16].

### 4.6 Second Algorithm: Intelligent Selection Improvement Algorithm with Composite Neighbourhood Structures (IS-CNS)

In IS-CNS, two neighbourhood lists (List-1 and List-2) are employed based on the intelligent selection mechanism. The algorithm is implemented in order to treat a randomised selection of the Basic-CNS. The IS-CNS is divided into two phases, constructive and improvement. The algorithm is also embedded with thirteen neighbourhood structures. An intelligent selection mechanism is employed to select the effective combination of neighbourhood structures. The selection of the neighbourhood structures from List-1 is randomly performed. While the neighbourhood structures that will be selected from list-2 are based on rules IS1 and IS2.

The second phase (improvement phase) is equipped with List-1 and List-2. The two lists in this phase contain the entire neighbourhood structures except N4 and N8. On the other hand, the best solution (BestSol) obtained in the first phase (constructive phase) is assigned as an initial solution (Sol) in the second phase. The same procedure is implemented in the improvement phase except the needed rules; those rules are IS2 to IS10. The quality measurement of both phases is explained in Section 4.3. The process is repeated until the termination criterion is met. The IS-CNS procedure is presented in [16].

## 4.7 Third Algorithm: Proposed Algorithm (IS-DT-CNS)

In this section, we will describe the idea of implementing the iterative improvement algorithm for attribute reduction problems by reducing the number of cardinality of the datasets with respect to the dependency degree. By understanding the nature of the exacerbated datasets, we believe that the problems can be solved by using a simple yet effective algorithm. In this work, the iterative improvement algorithm which can be classified as a "simple" method is a descent algorithm that always accepts a better solution and less dependent on the parameter setting, no matter how long the search process will take place. This motivates us to investigate more on the nature of the datasets instead of concentrating only in obtaining good solutions. It will help further in identifying the suitable components to be added to the algorithm in handling the problems in hand, and at the same time achieving better solutions.

Since the IS-DT-CNS scheme requires a list of neighbourhood structures, hence we are focusing on employing five lists of neighbourhood structures (i.e. List-1, List-2, List3, List-4 and List-5). The algorithm is divided into two phases i.e. constructive and improvement. In the first phase the same techniques of IS-CNS is adopted. Two lists (i.e. List-1 and List-2) are equipped in the constructive phase. The selection of the neighbourhood structure is based on IS1 and IS2 rules in order to generate a trial solution. The algorithm only accepts a better solution. This procedure is repeated until the dependency degree of the solutions reach the optimal value (i.e. in the constructive phase) which is equal to one.

In the second phase (i.e. an improvement phase), two procedures are employed called Treatment-1 and Treatment-2. Two lists (i.e. List-1 and List-2) are employed in Treatment-1, whilst three lists of neighbourhood structures (i.e. List-3, List-4 and List-5) are equipped in Treatment-2. Both of the treatments have different objectives to be achieved. The first treatment is to reduce the amount of the cardinality, while the second treatment is trying to escape from local optima with the dependency degree fully maintained. In attaining both objectives, we found out that the combination of the neighbourhoods in the first two lists (List-1 and List-2) helps to reduce the cardinality, while the other three lists (List-3, List-4, List-5) assist in escaping from the local optima. The content of each list is mentioned in Section 4.4.

**The first treatment (Treatment-1)**

The first treatment of the IS-DT-CNS employs two lists of neighbourhood structures (List-1 and List-2). These lists contain the entire neighbourhood structures except N4 and N8. The best solution acquired in the constructive phase is assigned as an initial solution (Sol) in Treatment-1. The neighbourhood structure is selected at random from List-1, whilst the neighbourhood structure from List-2 is rigorously selected based on rules IS2 to IS10. The quality measurement of Treatment-1 is the number of cardinality with dependency degree equals to one.

**The second treatment (Treatment-2)**

The second treatment of IS-DT-CNS entails applying the other three lists (i.e. List-3, List-4 and List-5) with a specific neighbourhood structures. List-3 contains N9 and N10. List-4 comprises N1, N5, N12 and N13. List-5 contains N2 and N6. The best solution obtained from Treatment-1 is assigned as an initial solution (Sol) in Treatment-2. Note that the lists in the second treatment are employed in sequence. The selection of neighbourhood structures is randomly selected from List-3 and List-4. Meanwhile the neighbourhood structures in List-5 are strictly selected based on rules (IS11 and IS12). The quality measurement of Treatment-2 is the number of cardinality, where the solution with less number of cardinality and the dependency degree is equal to one is considered as the best solution. The process is repeated until the termination condition is met. In this algorithm, the termination condition is set as a number of iterations. Fig. 1 shows the pseudo code of the IS-DT-CNS.

```
IS-DT-CNS algorithm pseudo code
Phase 1: Constructive:
        Generate an initial solution Sol;
        Set the intelligent selection list as $IS_k$, $K$= 1,..,$k$, where
        $k$=12;
        Set the quality measurement function  $f$(x); where x is
        the obtained solution;
        Set  maximum number of iteration, Max_iter← 250;
        Iter_cons ← 1;
        Do while (γ !=1 | | Iter_cons != Max_iter)   // first do-
        while loop
            Define a neighbourhood of Sol by randomly select a neighbourhood structure using $IS_1$
            and $IS_2$ to generate a trial solution called Sol*;
            If $f$(Sol*) is better than $f$(Sol);
                Sol ← Sol*;
            Increase Iter_cons by 1;
        End do while
Phase 2: Improvement:
          Set best solution from Phase 1 as an initial solution Sol;
        Iter_impro ← 1;
        Do while (Iter_impro!= (Max_iter – Iter_cons))    //
        second do-while loop
        Treatment-1:
            Define neighbourhood from(List-1 and List-2) of Sol by randomly select neighbour-
            hood structures using $IS_2$ to $IS_{10}$ to generate a trial solution called Sol*;
            If $f$(Sol*) is better than $f$(Sol);
                Sol ← Sol*;
        Treatment-2:
            Define neighbourhood from (List-3,List-4 and
            List-5) of Sol by randomly select
            neighbourhood structures using $IS_{11}$ and $IS_{12}$ to generate a new solution called Sol*;
            If $f$(Sol*) is better than $f$(Sol);
                Sol ← Sol*;
          Increase Iter_impro by 1;
   End do while
```

**Fig. 1.** The pseudo code for the IS-DT-CNS

The objective of the constructive phase (Phase 1) is to increase the dependency degree until it reaches one. When the dependency degree is equal one the algorithm immediately terminates the constructive phase and continues with the improvement phase. In the improvement phase (Phase 2), the number of iteration is equal to

(250 – Iter_cons), where Iter_cons is the number of iterations that is needed to obtain the dependency degree equal to 1. Note that, the total numbers of iterations is 250 (as been applied in other available approaches in the literature).

## 5 Results and Discussion

### 5.1 Results on Dimensionality Reduction

The algorithms are tested on 13 standard benchmark datasets taken from University of California Irvine (UCI). The datasets can be downloaded from http://www.ics.uci.edu/~mlearn. There are variant differences in the complexity for each dataset. More information about the complexity of the dataset can be found in [17, 18].

### 5.2 Comparison with the Literature

The performance of IS-DT-CNS is compared to other available algorithms in the literature as shown in Table 2. In ant colony optimisation (AntRSAR) applied by Jensen and Shen [2], the algorithm utilised the basic features of any colony algorithm. In addition, they took the expense of the time taken to discover the reducts into consideration. While, in [3] Ke et al did a better use of the ant colony by producing an efficient ant colony (ACOAR) of higher quality results with less expense of time. Both algorithms i.e. tabu search approach (TSAR) by Hedar et al [4] and scatter search (SSAR) by Jue et al. [5] are faster than other algorithms except ACOAR. Genetic algorithm (GenRSAR) and simulated annealing (SimRSAR) by Jensen and Shen [1] consume the highest expense in runtime. A Modified Great Deluge algorithm (MGDAR) is proposed by Mafarja and Abdullah [6]. Non-linear great deluge (NLGD-RSAR) proposed by Jaddi and Abdullah [7].

**Table 2.** Comparisons with state-of-the-art

| Datasets | IS-DT-CNS | AntRSAR | GenRSAR | ACOAR | TSAR | Sim-RSAR | SSAR | MGDAR | NLGD-RSAR |
|---|---|---|---|---|---|---|---|---|---|
| M-of-N | 6 | 6 | $6^{(6)} 7^{(12)}$ | 6 | 6 | 6 | 6 | 6 | 6 |
| Exactly | 6 | 6 | $6^{(10)} 7^{(10)}$ | 6 | 6 | 6 | 6 | 6 | 6 |
| Exactly2 | 10 | 10 | $10^{(9)}11^{(11)}$ | 10 | 10 | 10 | 10 | 10 | 10 |
| Heart | 6 | $6^{(18)} 7^{(2)}$ | $6^{(18)}7^{(2)}$ | 6 | 6 | $6^{(29)}7^{(1)}$ | 6 | $6^{(14)}7^{(6)}$ | 9 |
| Vote | 8 | 8 | $8^{(2)} 9^{(18)}$ | 8 | 8 | $8^{(15)}9^{(15)}$ | 8 | 8 | $10^{(14)}11^{(6)}$ |
| Credit | $8^{(13)}9^{(4)}$ $10^{(3)}$ | $8^{(12)}9^{(4)}$ $10^{(4)}$ | $10^{(6)}11^{(14)}$ | $8^{(16)} 9^{(4)}$ | $8^{(13)}9^{(5)}10^{(2)}$ | $8^{(18)}9^{(1)}$ $11^{(1)}$ | $8^{(9)}9^{(8)}$ $10^{(3)}$ | $8^{(13)}9^{(3)}10^{(4)}$ | 11 |
| Mush-room | 4 | 4 | $5^{(1)} 6^{(5)}7^{(14)}$ | 4 | $4^{(17)} 5^{(3)}$ | 4 | $4^{(12)}5^{(8)}$ | $4^{(7)}5^{(13)}$ | 4 |
| LED | 5 | $5^{(12)}6^{(4)}7^{(3)}$ | $6^{(1)}7^{(3)} 8^{(16)}$ | 5 | 5 | 5 | 5 | 5 | $7^{(15)}8^{(5)}$ |
| Letters | 8 | 8 | $8^{(8)} 9^{(12)}$ | 8 | $8^{(17)} 9^{(3)}$ | 8 | $8^{(5)}9^{(15)}$ | $8^{(18)}9^{(2)}$ | 9 |
| Derm | 6 | $6^{(17)} 7^{(3)}$ | $10^{(6)}11^{(14)}$ | 6 | $6^{(14)} 7^{(6)}$ | $6^{(12)}7^{(8)}$ | 6 | $6^{(11)}7^{(9)}$ | $11^{(17)}12^{(3)}$ |
| Derm2 | $8^{(8)} 9^{(12)}$ | $8^{(3)} 9^{(17)}$ | $10^{(4)}11^{(16)}$ | $8^{(4)}9^{(16)}$ | $8^{(2)}9^{(14)}10^{(4)}$ | $8^{(3)}9^{(7)}$ | $8^{(2)}9^{(18)}$ | $8^{(4)}9^{(12)}10^{(4)}$ | $11^{(15)}12^{(5)}$ |
| WQ | $12^{(2)}13^{(14)}$ $14^{(4)}$ | $12^{(2)}13^{(7)}$ $14^{(11)}$ | 16 | $12^{(4)}13^{(12)}$ $14^{(4)}$ | $12^{(1)}13^{(13)}$ $14^{(6)}$ | $13^{(16)}14^{(4)}$ | $13^{(4)}14^{(16)}$ | $12^{(4)}13^{(12)}$ $14^{(4)}$ | $15^{(11)}16^{(9)}$ |
| Lung | 4 | 4 | $6^{(8)}7^{(12)}$ | 4 | $4^{(6)} 5^{(13)} 6^{(1)}$ | $4^{(7)}5^{(12)}6^{(1)}$ | 4 | $4^{(6)}5^{(11)}6^{(3)}$ | 4 |

The results that are summarised in Table 2 show nine approaches that have been tested on the 13 datasets. On the whole, it appears that IS-DT-CNS is capable to outperform other algorithms in most of the tested datasets. The algorithm is able to produce a new lower bound on Derm2. This is due to the intelligent selection in the algorithm enhanced with the double treatments. Both mechanisms are able to improve the performance of only the modest technique (iterative improvement algorithm, coded as Basic-CNS) that has the merit of being a competitive algorithm.

We further investigate the performance of our approach by carrying out Friedman's statistical test. Friedman's test is used to test the differences between groups when the dependent variable being measure is ordinal, thus it rank the algorithms for each dataset separately and set the best performing algorithm as a control algorithm to compare against it. If significant differences are detected (based on Friedman's test), Holm's and Hochberg's tests are conducted as post hoc methods to obtain the adjusted *p*-values.

The average ranking of IS-DT-CNS, AntRSAR, GenRSAR, ACOAR, TSAR, SimRSAR, SSAR, MGDAR and NLGD-RSAR produced by Friedman's test are summarised in Table 3.

**Table 3.** Average ranking of the algorithms (based on Friedman's test)

| # | Algorithm | Ranking |
|---|---|---|
| 1 | ACOAR | 2.61 |
| 2 | IS-DT-CNS | 2.69 |
| 3 | SimRSAR | 3.61 |
| 4 | AntRSAR | 3.96 |
| 5 | SSAR | 4.11 |
| 5 | TSAR | 4.11 |
| 7 | MGDAR | 5.30 |
| 8 | NLGD-RSAR | 6.88 |
| 8 | GenRSAR | 6.88 |

From Table 3, it can be seen that IS-DT-CNS is ranked 2.69. The P-value of Friedman's test based on the best-performing algorithm is below the significance interval of 95% (critical level = 0.05), which is statistically significant different. Therefore, we perform a post hoc method (Holm's and Hochberg's test) to obtain the adjusted *p*-values for each comparison between ACOAR (as the controlling method) with the other approaches.

**Table 4.** Adjusted *p*-values of the methods in comparison

| # | Algorithm | Unadjusted P | P Holm | P Hochberg |
|---|---|---|---|---|
| 1 | GenRSAR | 7.76E-07 | 6.21E-06 | 6.21E-06 |
| 2 | NLGD-RSAR | 3.43E-04 | 0.002 | 0.002 |
| 3 | GenRSAR | 0.034 | 0.207 | 0.207 |
| 4 | TSAR | 0.099 | 0.497 | 0.428 |
| 5 | SSAR | 0.107 | 0.497 | 0.428 |
| 6 | AntRSAR | 0.162 | 0.497 | 0.487 |
| 7 | SimRSAR | 0.299 | 0.598 | 0.598 |
| 8 | IS-DT-CNS | 0.914 | 0.914 | 0.914 |

Table 4 shows the adjusted *p*-values of Holm's and Hockberg's statistical tests for each comparison when using the ACOAR as a control algorithm. It can be seen that

ACOAR is statistically better than 3 approaches i.e., GenRSAR, NLGD-RSAR and GenRSAR with adjusted *p*-value <0.05. Nevertheless, there is no significant difference between ACOAR and IS-DT-CNS (adjusted *p*-value is higher than 0.05).

## 6 Conclusion and Future Work

The double treatment iterative improvement algorithm with intelligent selection composite neighbourhood structure (IS-DT-CNS) has been applied to handle the attribute reduction problem. Due to the inability of the iterative improvement algorithm to find regions of the search space with better solution, therefore, implementation of the intelligent selection provides the algorithm with an efficient procedure that has been able to disassemble the problem. The double treatment mechanism helped the algorithm to better explore the search space in order to avoid being stuck into a local optima by finding other variation of the attributes with the same quality (with respect to the objective function). The algorithm is able to obtain higher quality results when compared with Basic-CNS and IS-CNS approaches. Further comparison with the available approaches in the literature showed that the algorithm is able to obtain competitive results. In future work, we planned to apply different measurement tools (other than the dependency degree rough set) such as entropy and a fuzzy rough set, in order to solve attribute reduction problems in terms of finding minimal reducts in lesser computational time, and with higher classification accuracy.

## References

1. Jensen, R., Shen, Q.: Semantics-preserving dimensionality reduction: rough and fuzzy-rough-based approaches. IEEE Transactions on Knowledge and Data Engineering 16, 1457–1471 (2004)
2. Jensen, R., Shen, Q.: Finding rough set reducts with ant colony optimization. In: Proc. UK Workshop on Computational Intelligence, pp. 15–22 (2003)
3. Ke, L., Feng, Z., Ren, Z.: An efficient ant colony optimization approach to attribute reduction in rough set theory. Pattern Recognition Letters 29, 1351–1357 (2008)
4. Hedar, A.R., Wang, J., Fukushima, M.: Tabu search for attribute reduction in rough set theory, Soft Computing-A Fusion of Foundations. Methodologies and Applications 12, 909–918 (2008)
5. Jue, W., Hedar, A., Shouyang, W.: Scatter Search for Rough Set Attribute Reduction. In: Second International Conference on Bio-Inspired Computing: Theories and Applications, BIC-TA 2007, pp. 236–240 (2007)
6. Mafarja, M., Abdullah, S.: Modified great deluge for attribute reduction in roughset theory. In: Proceedings of International Conference on Fuzzy Systems and Knowledge Discovery, pp. 1464–1469 (2011)
7. Jaddi, N.S., Abdullah, S.: Nonlinear Great Deluge Algorithm for Rough Set Attribute Reduction. Journal of Information Science and Engineering 29, 49–62 (2013)
8. Arajy, Y.Z., Abdullah Hybrid, S.: variable neighbourhood search algorithm for attribute reduction in Rough Set Theory. In: 2010 10th International Conference on Intelligent Systems Design and Applications (ISDA), pp. 1015–1020. IEEE (2010)

9. Grabowski, J., Pempera, J.: Sequencing of jobs in some production system. European Journal of Operational Research 125, 535–550 (2000)
10. Pawlak, Z.: Rough sets. Internat. J. Comput. Inform. Sci. 11, 341–356 (1982)
11. Pawlak, Z.: Rough sets: Theoretical aspects of reasoning about data. Springer (1991)
12. Liaw, C.-F.: An efficient tabu search approach for the two-machine preemptive open shop scheduling problem. Computers & Operations Research 30, 2081–2095 (2003)
13. Gopalakrishnan, M., Mohan, S., He, Z.: A tabu search heuristic for preventive maintenance scheduling. Computers & Industrial Engineering 40, 149–160 (2001)
14. Bilge, Ü., Kiraç, F., Kurtulan, M., Pekgün, P.: A tabu search algorithm for parallel machine total tardiness problem. Computers & Operations Research 31, 397–414 (2004)
15. Abdullah, S., Burke, E.K., McCollum, B.: Using a Randomised Iterative Improvement Algorithm with Composite Neighbourhood Structures for the University Course Timetabling Problem. In: Metaheuristics, pp. 153–169 (2007)
16. Jihad, S., Abdullah, S.: Investigating composite neighbourhood structure for attribute reduction in rough set theory. In: 2010 10th International Conference on Intelligent Systems Design and Applications (ISDA), pp. 1183–1188 (2010)
17. Oh, I.S., Lee, J.S., Moon, B.R.: Hybrid genetic algorithms for feature selection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1424-1437 (2004)
18. Bandyopadhyay, S., Maulik, U., Wang, J.T.L.: Analysis of biological data: a soft computing approach. World Scientific Pub. Co. Inc. (2007)

# Modeling Associative Memories Using Formal Concept Analysis

Cherukuri Aswani Kumar[1,*], M.S. Ishwarya[1], and Chu Kiong Loo[2]

[1] School of Information Technology & Engineering, VIT University, Vellore, India
[2] Dept. Of Artificial Intelligence, Faculty of CS and IT, University of Malaya
cherukuri@acm.org

**Abstract.** Associative memory is one of the primary functions of the human brain. In the literature, there are several neural networks based models that represent associative memory with the help of pattern associations. In this paper, we model the associative memory activity using Formal Concept Analysis (FCA), which is a standard technique for data and knowledge processing. In our proposal, patterns are associated with the help of object-attribute relations and the memory is represented using the formal concepts generated using FCA. We show that the extent and intent relations in the concepts help us to recall the patterns bi-directionally. Further, we model the pattern recall process for the given input even when the exact match is not found in the memory, using the concept hierarchies in the concept lattice.

**Keywords:** Associative Memories, Concept Hierarchy, Concept Lattice, Formal Concept Analysis, Pattern Association.

## 1 Introduction

One way of human brain to learn and memorize the new concepts is through association. Artificial Neural Networks (ANN) that are aimed at modeling the human brain behavior implements this association through either feed-forward or recurrent networks. These networks store the set of patterns as memories [1]. The main goal of these networks is to identify the corresponding associated pattern in the memory for the given input pattern. Memories that model association phenomenon are of two types: Auto-associative and hetero-associative memories. In auto-associative memory, the input pattern and the corresponding output pattern are same. However, in hetero-association, input and the corresponding output patterns differ [1]. The most popular and widely used associative memory models are Hopfield networks, Bidirectional Associative Memory (BAM) etc. Among these models, our particular interest in this research is modeling the BAM phenomenon.

BAM is a recurrent network that is capable of modeling the hetero-association. BAM are memories that can be mapped in both the directions (forward and backward) between same or different pattern types. BAM consists of two layers of neurons (say

---

[*] Corresponding author.

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_11

X and Y). All the neurons in layer X is connected to every other neuron in layer Y. The states of neurons in the layer *X* and *Y* are $x^i$ and $y^j$ can take 1,-1 in case of bipolar encoding and 0, 1 in case of binary encoding. After the establishment of weights, when an input pattern is presented at layer X corresponding associated pattern is recalled at layer Y and vice versa. The pattern recall depends on the initial pattern pairs and their connection in weight matrix. The calculated weight matrix is the memory. Whenever a pattern is presented to BAM, it recalls the associated pattern with the help of weight matrix by multiplying the input vectors to the weight matrix. Later, this intermediate result is given to an activation function. If the field in the obtained vector satisfies the threshold of the applied activation function, it is taken as 1 and the corresponding neural unit is fired else it is 0 and no changes takes place in corresponding neural unit. The respective output from the field yields a pattern, which is the associated pattern for the given input.

Formal Concept Analysis (FCA), proposed by Wille, is a mathematical framework for conceptual data analysis based on the lattice theory. It is mainly used as an analytical tool for knowledge representation and discovery [2]. FCA achieves this objective by investigating the relationship between the real world objects and their attributes. One of the outputs from FCA is the set of formal concepts while the concepts are regarded as the fundamental units of human thinking. In the literature, few studies are reported that connect the human cognitive process and FCA as discussed in Section 3. Further, FCA is used to optimize the ANN learning process [3,4,5]. FCA is also used for information retrieval applications using concept hierarchy that responds for partial input queries [6].

Extending upon these observations, in this paper, we attempt to model the hetero-associative memory process using FCA. Rest of the paper is organized as follows. Section 2 provides a brief background on FCA. Section 3 reviews the related works. In section 4, we propose our model for representing associative memory using FCA. Section 5 provides the experimental analysis of the proposed model.

## 2 Background

FCA is the principle way of deriving a concept hierarchy from the collection of objects and attributes. The object-attribute relation forms the formal context. FCA works on these formal contexts to generate the list of formal concepts. Every context is a triple (G, M, I) where G is the set of objects or extent, M is the set of attributes or intent and I is the binary relation between G and M (I is also known as incidence relation). A formal concept is given by (A, B) where $A \subseteq G$ and $B \subseteq M$ such that $B^{\downarrow}= A$ ($A \subseteq G$ is the set of all objects that have all the attributes in $B \subseteq M$), $A^{\uparrow}= B$ ($B \subseteq M$ is the set of all attributes present in all the objects in $A \subseteq G$). List of concepts and corresponding lattice has to be generated from the given context. Concept lattices show the Galois connection among the formal concepts where each node in the lattice represents a formal concept. A lattice is a partial ordered set in which every pair of elements has a Lowest Upper Bound (LUB) and Greatest Lower Bound (GLB). Also, an inheritance hierarchy exists in these connections which is a directed rooted graph where each node inherits from the all the other nodes above it. The crucial relationship between the concepts is seen when they are ordered hierarchically [7]. Let B (G, M, I) denote the set of concepts in the context (G, M, I). If $(A_1, B_1)$ and $(A_2, B_2)$ are concepts of the context B (G, M, I) then $(A_1, B_1)$ is called a subset $(A_2, B_2)$ if $A_1 \subseteq A_2$ (or equivalently $B_1 \supseteq B_2$). This sub-super concept relations is denoted by $(A_1, B_1) \leq (A_2, B_2)$. A subset will have fewer objects and greater attributes than any of its supersets. From an Object-Oriented_view, the extent of the concept specializes and the intent of the concept generalizes when the lattice is traversed from top to bottom (The extent of the concept generalizes and the intent of the concept specializes when the lattice is traversed from bottom to top). Galois connection between two partially ordered sets which are duals of each other is called Galois duality. A concept (A, B) where A is the extent and B is the intent, Galois duality states that by minimizing A, B can be maximized and vice versa.

Several applications of FCA are reported in the literature. Interested readers can refer following references for details [8, 9, 10, 11].

## 3 Related Works

Concepts are the base of human thinking and FCA. Perceptions and cognitions are one among the natural ability of human brain which can be modeled to an extent with neural networks. Gärdenfor's work on conceptual space suggests that the above mentioned ability that are modeled using non-symbolic representation (specifically neural network) can be efficiently substituted by symbolic representations FCA [12]. Another cognitive application of FCA serves for the enhancing cognition in web [13]. People's reasoning and cognitive capability can be modeled with FCA which results in Conceptual Space Markup language (CSML). FCA along with the inspiration from human memory approaches the physical memory bounds [14].This approach proposes a model that would exceptionally provide high density of memory without traditionally required increase in the physical storage capacity. Similar to relating dynamics of FCA, Patel et al proposes a model for relating subconscious and conscious cognitive processes by designing a reference model for brain [15].

FCA is taken to next level by realizing its psychological ramification of concepts [16]. Mathematical models utilize the conceptual spaces to substitute the time that are required to reach conclusions similar to human conclusions [17]. Zarate et al. [3] modeled an approach for extracting and representing knowledge in Artificial Neural Network (ANN) [3]. It is also used to extract knowledge from a specifically trained ANN like for cold rolling process [4]. An interesting investigation by Dyce et al. improves the cognitive structure of FCA and is being used for discovering ways for better operation of semantic memory of Artificial Intelligence [18]. SOPHIANN is a computational tool modeled for extracting the knowledge from ANN through FCA [19]. Novel cognitive system was proposed that explains how FCA is used exactly to describe human cognitive process [20]. Inspired by the basic idea of FCA, two ways of directionally encoding closure operators on a finite set in a neural network is modeled by Rudolph [21]. Computational equivalent of human mind is attempted through modeling Biologically Inspired Cognitive Architectures (BICA) by extending cognitive architectures [22].

Similarities between BAM and FCA were inferred in the way in which they handle data. Belohlavek [23] has introduced fuzzy based bi-directional associative memory (FLBAM) which is stable and showed that all these stable points form a complete lattice. In another work, Belohlavek [24] has used BAM for representing concept lattice hierarchical structures by interpreting the BAM patterns to represent the concepts. In an interesting work, Acevedo et al. [25] have used Alpha-Beta associative models to store and recall a concept lattice.

Belohlavek [24] proposed a model for representing FCA using BAM emphasizing the fact that FCA relates every datum to other with a relation similar to brain. Further he proved that for each concept lattice there is BAM and all the concepts in the lattice represent the set of all stable points of BAM. In the current work, we propose to model the neural associative memories using FCA while the above mentioned works [24, 25] uses associative memories for modeling FCA principles.

Neural decoding is one among the primary applications of FCA (extracting of information that are encoded and stored in brain) [5] and the effects of sparse neural code can be modeled using concept lattice. Aravind Kumar has proposed a novel way of information retrieval using concept lattices [6]. Motivated from this analysis, we have used the conceptual hierarchy to handle the partial matches in the memory to the given input patterns.

Based on the link between human cognitive process and the way FCA deal with relation between real world objects and attributes, we aim at modeling the associative memories using FCA. Intents and extent in FCA are associated with each other through a relation while, each neural unit in associative memory is associated with each other through weights. The object-attribute relationship in FCA and weights in associative memories are regarded as their corresponding memories while each can be transformed through some function to the other. This theory forms a solid testament for implementing the functionality of associative memories with FCA.

## 4 Proposed Work

To model the associative memory through formal concepts, we regard the set of objects or the extent as one layer of neurons (where each object is considered to be a neuron) and the set of attributes or intent as the other layer of neurons (where each attribute is considered to be a neuron). Since there exists a relation between every object and attribute pair, every neuron in one layer is connected to every neuron in other layer.

According to the aforementioned theory, the memory stores the relationships of the object and attributes. Using concept generation algorithm, we derive the list of concepts for the given intent and extent relationships. In BAM, the patterns are recalled based on the memory (Weight matrix). The generated list of concepts which manifest the link between intent and extent will correspond to the memory. Similar to BAM, this model can perform forward as well as backward recall. Also, the model responds for partial pattern and recalls their associated pattern.

In the following, we illustrate the functionality of the proposed model shown in Fig 1. The given data is preprocessed and formal context is created describing the data as object, attribute and binary relations between them. From the context, concepts are generated using concept generation algorithm [26]. In the cases where in there is a

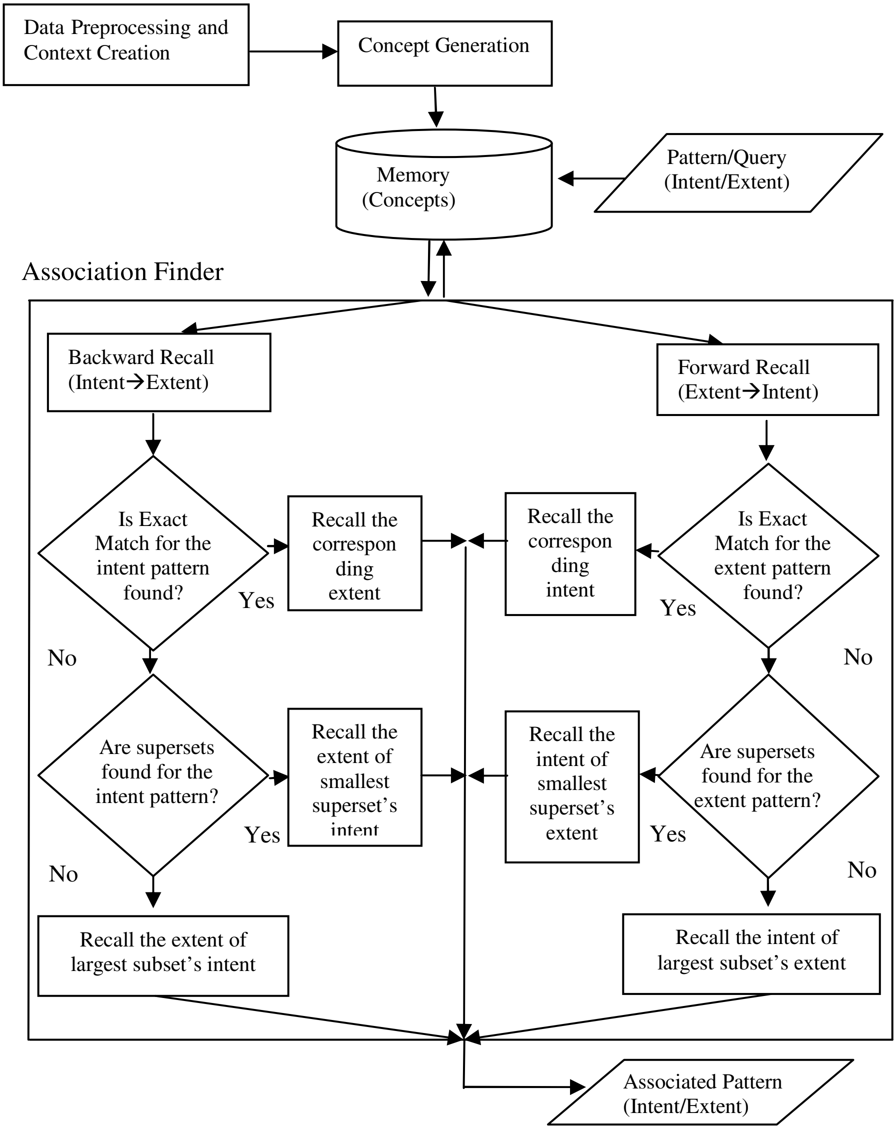


**Fig. 1.** Modeling Associative Memories with FCA

unit concept with null intent and a zero concept with null extent are not considered by the association finder for recalling the associated patterns. After all the concepts were discovered, the memory is queried with a pattern to recall its associated pattern. Recall can be either forward or backward. In a forward recall, the association finder can take three different ways to retrieve the associated pattern. They are: 1) If an exact match is found for the given input (extent) pattern, the corresponding associated pattern (intent) is recalled

2) When there is no exact match in the memory, the association finder checks whether any supersets exist for the given input pattern. If any such supersets are available, then the association finder would select the smallest superset among them and recall the corresponding intent as the associated pattern. 3) When there is no exact match and no supersets are available, the association finder would select the largest subset for the given pattern and recall the corresponding intent as associated pattern.

Similarly, in backward recall, the association finder takes three different ways to retrieve the associated pattern. They are 1) If an exact match is found for the given input (intent) pattern, the corresponding associated (extent) pattern is recalled 2) When there is no exact match found in memory, the association finder checks whether supersets exists for the given input pattern. If any supersets are available, then the association finder would select the smallest superset among them and recall the corresponding extent as associated pattern. 3) When there is no exact match and no supersets are available, then the association finder would select the largest subset among them and recall the corresponding extent as associated pattern. In the following section we present the experiment analysis of the proposed model.

## 5 Experimental Analysis

We conducted experiments on the proposed model with a data set available at the Department of biostatistics web page, Vanderbilt University [27]. The dataset contains the recorded diabetic test results of patients from two different locations. This data set is normalized with the help of variety of sources [28]. The normalized binary dataset has 1 for positive symptoms (Exceeding the normal conditions) and 0 for negative symptoms (Normal conditions). This normalized data set has 404 rows and 11 columns (describes about 404 patients with respect to 11 symptoms of type II diabetes). This normalized dataset is further processed as formal context.

Next neighbor algorithm is used to generate the list of concepts from the context [28]. The algorithm has generated 194 formal concepts from the given context. The model considers this set of formal concepts as memory as illustrated in the Section 4. Subsequently experiments are conducted to check the recall ability of the model in both directions. A total of 6 different cases: 1. Exact forward recall 2. Exact backward recall 3. Smallest superset forward recall 4. Largest subset forward recall 5. Smallest superset backward recall 6. Largest subset backward recall, are considered for experiments.

Case 1: The patient's numbers 4790, 40762, 41029 and 41506 were given as input pattern (extent). Exact pattern is found in the memory and corresponding intent is recalled. The recalled associated pattern is Cholesterol, Glucose, HDL, Ratio, Glyhb, Location, Age, Gender, BMI, Frame.

Case 2: The symptoms Cholesterol, Glucose, HDL, Ratio, Glyhb, Location, Age, Gender, BMI and Frame were given as input pattern (intent). Exact match for the given input pattern is found in the memory and the corresponding extent is recalled. The recalled associated pattern is 4790, 40762, 41029, 41506.

Case 3: The patient's numbers 4790, 40762, 41029 were given as input pattern (extent). No exact pattern is found in the memory and the association finder checks for supersets for the given extent. Supersets were found for the given extent and the intent of the smallest superset is recalled. The recalled associated pattern is Cholesterol, Glucose, HDL, Ratio, Glyhb, Location, Age, Gender, BMI, Frame.

Case 4: A new patient's number "500" that is not present in the memory along with patient's numbers 4790,40762,41029,41506 is given as input pattern (extent). Since no exact match is found and no supersets are available, the system recalls the largest subset's intent as the recalled pattern. The recalled associated pattern is Cholesterol, Glucose, HDL, Ratio, Glyhb, Location, Age, Gender, BMI, Frame.

Case 5: HDL is given as input pattern (intent) to the memory. Since, no exact match is found, the association finder checks for the availability of supersets. From the available supersets, the association finder selects the smallest superset and recalls its extent. The recalled associated pattern is the extent of Glucose and HDL.

Case 6: A new symptom "BP" is added to the pattern Cholesterol, Glucose, HDL, Ratio, Glyhb, Location, Age, Gender, BMI, Frame and given as input pattern (intent) to the memory. Since no exact match and no supersets are found in the memory, the association finder chooses the largest subset's extent as the recalled pattern. The recalled associated pattern is 4790, 40762, 41029, 41506.

**Table 1.** Results of Experiments conducted on the modeled memory

| Dataset# | Object count | Attribute count | Formal concept count | Number of tests conducted | Recall Accuracy (%) |
|---|---|---|---|---|---|
| Dataset-A | 404 | 11 | 194 | 53 | 98.07 |
| Dataset-B | 101 | 13 | 169 | 72 | 97.1 |
| Dataset-C | 121 | 6 | 18 | 31 | 100 |

*- The actual result in some cases seemed to be more generalized than the expected result (i.e., it includes some attributes/objects in addition to expected attributes/objects).

#- Dataset-A [27], Dataset-B [29], Dataset-C [20]

We have tested the model for different possible recalls and a relevant response for the same is observed. Table 1 shows the recorded results of the experiments that have been conducted over the proposed model with the datasets A, B and C. In table 1, the field ‘object count’ and ‘attribute count’ represents the number of objects and attributes present in the pre-processed dataset. The field ‘formal concept count’ indicates the number of formal concepts generated after the application of concept generation algorithm on the respective pre-processed formal contexts and ‘memory recall accuracy’ shows the average recall accuracy of the tests conducted on the corresponding dataset.

The merits of the proposed method can be identified by comparing with the traditional neural networks based associative memory. During the training process in a recurrent neural network, the network computes the weights of all the non-input neural units based on state variable and the previous input. This training process is continued by adjusting the weights till the network produces the desired output for the given input [30]. The proposed method performs the training by means of concept generation. The worst case time complexity involved in training is $O\ (|C||G||M|^2)$ where G is the set of objects, M is the set of attributes and C is the set of formal concepts. The black-box nature of the neural network hides the knowledge and internal working of the network [31]. The knowledge of the proposed method is represented in terms of formal concept which holds the association between the patterns. During a pattern recall in traditional neural network, if an unknown pattern is presented to the network, activation vectors may converge and result to pattern that is not stored in the network (spurious stable state). The proposed method will not generate any spurious stable state and handles the unknown pattern with the approach followed in [6].

The next challenge that we aim to address is the incremental memory update, such as [32], given any update or modification upon learning to any of the object or attribute patterns.

## 6 Conclusions

This paper aimed at modeling associative memory using FCA. For this purpose, we have regarded formal concepts as the memory blocks. Exploiting the object-attribute relations, patterns are associated and recalled in both directions. Also the proposed model is able to recall the associated patterns even for the inputs with additive and subtractive noises with the help of concept hierarchies. The future work concentrates on handling the memory update in an efficient manner by manipulating only those parts of memory that are affected by the update action, so that entire memory need not be built again.

## References

1. Sivanandam, S.N., Deepa, S.N.: Principles of Soft Computing, 2nd edn. Wiley India (2011)
2. Ganter, B., Wille, R., Franzke, C.: Formal Concept Analysis: Mathematical Foundation. Springer (1999)
3. Zarate, L.E., Mariano Dias, S., Junho Song, M.A.: FCANN: A new approach for extracting and representing knowledge from ANN trained via Formal Concept Analysis. Neurocomputing 71, 2670–2684 (2008)
4. Zarate, L.E., Mariano Dias, S.: Qualitative behavior rules for the cold rolling process extracted from trained ANN via the FCANN method. Engineering Applications of Artificial Intelligence 22, 718–731 (2009)
5. Endres, D., Foldiak, P., Priss, U.: An application of Formal Concept Analysis to Semantic Neural Decoding. Annals of Mathematics and Artificial Intelligence 57, 233–248 (2009)
6. Aravind Kumar, M.: Information Retrieval using Concept Lattices. Technical Report, University of Cincinnati (2006)
7. Carpineto, C., Romano, G.: Conceptual Data Analysis. John Wiley and Sons Ltd. (2004)
8. Aswani Kumar, Ch., Annapurna, J.: Exploring Attributes with Domain Knowledge in Formal Concept Analysis. Journal of Computing and Information Technology 21, 109–123 (2013)
9. Aswani Kumar, Ch., Srinivas, S.: Concept Lattice using Fuzzy K-means Clustering. Expert Systems with Applications 37(3), 2696–2704 (2010)
10. Aswani Kumar, Ch.: Mining association rules using non-negative matrix factorization and formal concept analysis. In: Venugopal, K.R., Patnaik, L.M. (eds.) ICIP 2011. CCIS, vol. 157, pp. 31–39. Springer, Heidelberg (2011)
11. Poelmans, J., Ignatov, D.I., Kuzentsov, S.O., Dedene, G.: Formal concept analysis in knowledge processing: A survey on applications. Expert System with Applications 40, 6538–6560 (2013)
12. Gärdenfors, P.: Conceptual Space as a framework of Knowledge Representation. Mind and Matter 2(2), 9–27 (2004)
13. Raubal, M.: Benjamin Adams, The semantic Web needs More Cognition. Semantic Web 1, 69–74 (2010)
14. Kitto, K., Bruza, P., Gabora, L.: A Quantum Information Retrieval Approach to Memory. In: Proceedings of International Joint Conference on Neural Networks (2012)
15. Wang, Y., Wang, Y., Patel, S., Patel, D.: A Layered Reference Model of the Brain. IEEE Transaction on System, Man and Cybernetics 36(2) (2006)
16. Belohlavek, R., Trnecka, M.: Basic Level in Formal Concept Analysis: Interesting Concepts and Psychological Ramifications. In: Proceedings of Twenty-Third International Joint Conference on Artificial Intelligence, pp. 1233–1239 (2013)
17. Goguen, J.: Mathematical models of cognitive space and time. In: Andler, D., Ogawa, Y., Okada, M., Watanabe, S. (eds.) Reasoning and Cognition: Proc. of the Interdisciplinary Conference on Reasoning and Cognition, pp. 125–128. Keio University Press (2006)
18. Dyce, W., Marmin, T., Patel, N., Sipieter, C., Tisserant, G., Prince, V.: Let the System Learn a Game: How FCA can Optimize a Cognitive Memory Structure. In: 20th European Conference on Artificial Intelligence (2012)
19. Zarate, L.E., Song, M., Alvarez, A., Soares, B., Nogueira, B., Vimieiro, R., Dias, S., Santos, T., Vieria, N.: An approach to knowledge extraction from ANN through Formal Concept Analysis-Computational Tool proposal: SOPHIANN, vol. 1, pp. 43–48, 9–13 (2006)

20. Xu, W., Pang, J., Luo, S.: A novel cognitive system model and approach to transformation of information granules. International Journal of Approximate Reasoning 55, 853–866 (2014)
21. Rudolph, S.: Using FCA for Encoding Closure Operators into Neural Networks. In: Priss, U., Polovina, S., Hill, R. (eds.) ICCS 2007. LNCS (LNAI), vol. 4604, pp. 321–332. Springer, Heidelberg (2007)
22. Chella, A., Pirrone, R., Sorbello, R., Jóhannsdóttir, K.R. (eds.): Biologically Inspired Cognitive Architectures 2012. AISC, vol. 196, pp. 33–34. Springer, Heidelberg (2013)
23. Belohlavek, R.: Fuzzy Logical Bidirectional Associative Memory. Information Science 128, 91–103 (2000)
24. Belohlavek, R.: Representation of Concept Lattice by Bidirectional Associative Memories. Neural Computation 12, 2279–2290 (2000)
25. Acevedo, M.E., Yanez-Marquez, C., Acevedo, M.A.: Associative Models for Storing and Retrieving Concept Lattice. Mathematical Problems in Engineering, Article ID 356029 (2010)
26. Aswani Kumar, Ch., Singh, P.K.: Knowledge Representation Using Formal Concept Analysis: A Study on Concept Generation. In: Global Trends in Knowledge Representation and Computational Intelligence, pp. 306–336. IGI Global Publishers (2014)
27. Dataset from Department of Biostatistics, Vanderbilt University, Nashville, Tennessee, `http://biostat.mc.vanderbilt.edu/wiki/pub/Main/DataSets`
28. Data set Normalization: 1) Cholesterol: `http://www.heart.org`; `http://www.vaughns-1-pagers.com` 2) Glucose: `http://en.wikipedia.org/wiki/Blood_sugar`; `http://abcnews.go.com/health` 3) HDL: `http://en.wikipedia.org/wiki/High-density_lipoprotein`; `http://www.heart.org` 4) Ratio: `http://www.mayoclinic.org`; `http://www.exrx.net` 5) Glycolated Hemoglobin: `http://www.webmd.com`; `http://en.wikipedia.org/wiki/Glycated_hemoglobin` 6) BMI: `http://www.en.wikipedia.org/wiki/Body_Mass_Index` 7) Hip/Waist Ratio: `http://www.virginactive.co.uk`
29. `http://www.liacs.nl/~edegraaf/assign1.html`
30. Basheer, I.A., Hajmeer, M.: Artificial Neural Networks: fundamentals, computing, design and application. Journal of Microbiological Methods 43, 3–31 (2000)
31. Benitez, J.M., Castro, J.L., Requena, I.: Are Atrificial Neural Network Black Boxes? IEEE Transaction on Neural Netwrok 8, 1156–1164 (1997)
32. Kriegel, F.: Incremental computation of concept lattice. Studia-Univ. Babes, Bolyai, Informatica, vol. LIX(1) (2014)

# A New State Reduction Approach for Fuzzy Cognitive Map with Case Studies for Waste Management Systems

Miklós Ferenc Hatwágner[1], Adrienn Buruzs[2],
Péter Földesi[3], and László Tamás Kóczy[4]

[1] Department of Information Technology, Széchenyi István University, Győr, Hungary
miklos.hatwagner@sze.hu
[2] Department of Environmental Engineering, Széchenyi István University, Győr, Hungary
buruzs@sze.hu
[3] Department of Logistics, Széchenyi István University, Győr, Hungary
foldesi@sze.hu
[4] Department of Automation, Széchenyi István University, Győr,
Hungary and Department of Telecommunications and Media Informatics,
Budapest University of Technology and Economics, Hungary
koczy@sze.hu, koczy@tmit.bme.hu

**Abstract.** The authors have investigated the sustainability of Integrated Waste Management Systems (IWMS). These systems were modeled by Fuzzy Cognitive Maps (FCM), which are known as adequate fuzzy-neural network type models for multi-component systems with a stable state. The FCM model was designed of thirty-three factors to describe the real world processes of IWMS in as much detailed and as much accurately as possible. Although, this detailed model meets the requirements of accuracy, the presentation and explanation of such a complex model is difficult due to its size.

While there is a general consensus in the literature about a very much simplified model of IWMSs, detailed investigation lead to the assumption that a much more complex model with considerably more factors (components) would more adequately simulate the rather complex real life behavior of the IWMS.

As the starting point we used the thirty-three component model based on the consensus of a workshop of experts coming from all areas of the IWMS (operation, regulation, management, etc.) and the set goal was to find the most accurate real model that could be obtained by analyzing and properly reducing this – very likely too much detailed, or atomized – model.

In this paper, a new state reduction approach with three different metrics is presented. The practical aspects of the results gained by these methods are evaluated.

**Keywords:** fuzzy cognitive maps, integrated waste management system, state reduction methods.

## 1 Introduction

In the course of the previous investigations [1] FCM was used to simulate waste management systems determined by six factors. As a validation of the simulation results

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_12

[2] data were collected based on the relevant literature to set up a time series. This time series served as an input to the Bacterial Evolutionary Algorithm (BEA) which generated a (quasi-)optimal connection matrix producing the possibly most similar time series to the original one obtained from the literature. Despite the expectations, the six-factor FCM model produced rather inaccurate results in practice [3], and this is why a much more detailed model, containing thirty-three factors was created [5] with the help of a group of experts.

After the thorough examination of both the basic and the detailed (33 component) models it became apparent that the two models were very different, in their respective complexities and concepts. For this reason, we assumed that an intermediate model containing less than thirty-three but more than six factors would be presumably able to describe the mechanism and action of a real IWMS with sufficient accuracy.

Table 1 introduces the main factors of the basic model and the thirty-three sub-factors of the detailed model. Table 2 describes the detailed connection matrix.

**Table 1.** The identified sub-factors of the main factors and the concept IDs (CID) of them

| Main factor | Sub-factor | CID | Main factor | Sub-factor | CID |
|---|---|---|---|---|---|
| Technology (C1) | Engineering knowledge | C1.1 | Society (C4) | Public opinion | C4.1 |
| | Technological system and its coherence | C1.2 | | Public health | C4.2 |
| | Local geographical and infrastructural conditions | C1.3 | | Political and power factors | C4.3 |
| | Technical requirements in the EU and national policy | C1.4 | | Education | C4.4 |
| | Technical level of equipment | C1.5 | | Culture | C4.5 |
| Environment (C2) | Impact on environmental elements | C2.1 | | Social environment | C4.6 |
| | Waste recovery | C2.2 | | Employment | C4.7 |
| | Geographical factor | C2.3 | Law (C5) | Monitoring and sanctioning | C5.1 |
| | Resource use | C2.4 | | Internal and external legal coherence (domestic law) | C5.2 |
| | Wildlife (social acceptance) | C2.5 | | General waste management regulation in the EU | C5.3 |
| | Environmental feedback | C2.6 | | Policy strategy and method of implementation | C5.4 |
| Economy (C3) | Composition and income level of the population | C3.1 | Institution (C6) | Publicity, transparency (data management) | C6.1 |
| | Changes in public service fees | C3.2 | | Elimination of duplicate authority | C6.2 |
| | Depreciation and resource development | C3.3 | | Fast and flexible administration | C6.3 |
| | Economic interest of operators | C3.4 | | Cooperation among institutions | C6.4 |
| | Financing | C3.5 | | Improvement of professional standards | C6.5 |
| | Structure of industry | C3.6 | | | |

On the basis of this method, we might be able to support the strategic decision making process of the stakeholder in order to ensure the long-term sustainability of IWMS.

**Table 2.** Detailed connection matrix created by experts normalized for [-10, +10]

| CID | C1.1 | C1.2 | C1.3 | C1.4 | C1.5 | C2.1 | C2.2 | C2.3 | C2.4 | C2.5 | C2.6 | C3.1 | C3.2 | C3.3 | C3.4 | C3.5 | C3.6 | C4.1 | C4.2 | C4.3 | C4.4 | C4.5 | C4.6 | C4.7 | C5.1 | C5.2 | C5.3 | C5.4 | C6.1 | C6.2 | C6.3 | C6.4 | C6.5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C1.1 | 0 | 2 | 0 | 6 | 4 | 6 | 2 | 0 | 8 | 2 | 6 | 4 | 8 | 4 | 8 | 4 | 4 | 0 | 0 | 0 | 4 | 0 | 0 | -6 | 0 | 4 | 4 | 4 | 0 | 0 | 0 | 0 | 2 |
| C1.2 | 4 | 0 | 4 | 4 | 6 | 2 | 2 | 0 | 4 | 2 | 4 | 6 | 8 | 6 | 6 | 6 | 6 | 0 | 2 | 2 | 0 | 0 | 2 | -2 | -6 | 6 | 2 | 6 | 8 | 4 | 0 | 6 | 4 |
| C1.3 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 | 4 | 6 | 6 | 6 | 6 | 4 | 4 | 0 | 2 | 0 | 0 | 4 | 2 | 0 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 2 |
| C1.4 | 2 | 0 | 0 | 0 | 0 | 6 | 2 | 0 | 6 | 6 | 8 | 8 | 8 | 4 | 8 | 8 | 8 | 2 | 6 | 4 | 2 | 2 | 0 | 0 | 2 | 2 | 6 | 0 | 4 | 0 | 2 | 0 | 6 |
| C1.5 | 8 | 2 | 0 | 8 | 0 | 4 | 2 | 0 | 4 | 4 | 6 | 6 | 8 | 6 | 6 | 6 | 6 | 0 | 2 | 0 | 0 | 4 | 2 | -2 | 6 | 0 | 2 | 4 | 0 | 0 | 0 | 2 | 8 |
| C2.1 | 0 | 0 | 6 | 2 | 0 | 0 | 0 | 0 | 2 | 4 | -6 | 0 | 2 | 0 | 0 | 0 | 0 | 4 | 8 | 0 | 2 | 0 | 0 | 0 | 6 | 4 | 4 | 2 | 0 | 0 | 0 | 2 | 0 |
| C2.2 | 0 | 2 | 0 | 0 | 2 | 4 | 0 | 6 | -10 | 0 | -6 | 0 | -4 | 4 | 8 | 6 | 10 | -6 | 4 | 0 | 2 | 0 | -2 | 4 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 4 | 4 |
| C2.3 | 0 | 0 | 6 | 0 | 0 | 4 | 4 | 0 | 4 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 6 | -6 | -6 | 0 | 0 | 0 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C2.4 | 0 | 2 | 4 | 0 | 6 | -6 | -8 | -6 | 0 | -4 | -6 | 0 | -2 | 0 | 0 | -2 | 2 | 2 | -6 | 0 | 2 | 2 | 4 | 4 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 2 |
| C2.5 | 0 | 0 | 0 | 6 | 0 | 4 | 0 | 4 | 0 | 0 | 4 | 0 | 0 | 0 | 2 | 0 | 0 | 6 | 4 | 0 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C2.6 | 0 | 6 | -8 | 6 | 6 | -8 | 6 | 0 | 6 | -8 | 0 | -6 | 2 | 0 | 0 | 0 | 2 | 8 | 4 | 0 | 2 | 0 | 4 | 0 | 6 | 2 | 8 | 8 | 0 | 0 | 0 | 2 | 0 |
| C3.1 | 0 | 2 | 0 | 0 | 2 | -8 | 4 | 0 | 0 | 2 | 2 | 0 | 8 | 8 | 6 | 6 | 0 | 10 | 4 | 0 | 6 | 6 | 8 | 0 | 0 | 0 | 2 | 0 | 2 | 0 | 6 | 0 | 0 |
| C3.2 | 0 | 6 | 0 | 0 | 6 | -6 | 4 | 0 | 6 | 0 | 4 | 0 | 0 | 6 | 8 | 8 | 10 | 10 | 0 | 0 | 2 | 0 | -4 | -4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C3.3 | 0 | 6 | 0 | 2 | 4 | 4 | 4 | 0 | 2 | 2 | 2 | 0 | 6 | 0 | 4 | 8 | 8 | 0 | 0 | 0 | 6 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| C3.4 | 8 | 8 | 0 | 2 | 8 | -6 | 8 | 0 | -2 | 2 | 2 | 0 | 10 | 6 | 0 | 6 | 4 | 4 | 0 | 0 | 2 | 0 | 0 | -4 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 6 | 0 |
| C3.5 | 0 | 4 | 0 | 0 | 6 | 4 | 8 | 0 | 6 | 0 | 0 | 0 | 6 | 6 | 6 | 0 | 8 | 0 | 0 | 0 | 2 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 |
| C3.6 | 0 | 6 | 0 | 0 | 8 | 6 | 10 | 8 | -8 | 4 | 4 | 0 | 4 | 2 | 6 | 4 | 0 | -4 | 0 | 0 | 4 | 0 | 0 | -6 | 0 | 6 | 6 | 6 | 0 | 0 | 0 | 0 | 8 |
| C4.1 | 2 | 2 | 0 | 6 | 6 | 8 | 6 | 4 | 8 | 10 | 6 | 2 | 6 | 4 | 6 | 4 | 4 | 0 | 8 | 4 | 6 | 8 | 8 | 0 | 0 | 8 | 4 | 4 | 6 | 0 | 2 | 4 | 2 |
| C4.2 | 4 | 2 | 2 | 6 | 6 | 6 | -2 | 2 | 8 | 8 | 10 | 6 | 4 | 4 | 4 | 4 | 4 | 8 | 0 | 2 | 4 | 8 | 8 | 0 | 6 | 8 | 6 | 4 | 4 | 0 | 0 | 0 | 0 |
| C4.3 | 0 | 8 | 0 | 4 | 0 | 0 | 0 | 0 | -2 | 4 | -2 | 6 | 10 | 8 | 6 | 8 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 6 | 6 | 8 | 2 | 8 | 0 |
| C4.4 | 2 | 0 | 0 | 2 | 2 | 4 | 2 | 0 | 6 | 6 | 6 | 8 | 2 | 2 | 2 | 2 | 2 | 4 | 10 | 2 | 0 | 6 | 6 | 0 | 2 | 4 | 4 | 4 | 0 | 0 | 0 | 0 | 0 |
| C4.5 | 2 | 0 | 4 | 6 | 8 | -2 | 6 | 2 | 4 | 8 | 6 | 2 | 2 | 2 | 2 | 2 | 2 | 6 | 4 | 2 | 8 | 0 | 6 | 0 | 4 | 4 | 4 | 4 | 2 | 0 | 0 | 0 | 0 |
| C4.6 | 0 | 0 | 4 | 6 | 4 | 2 | 6 | 2 | 4 | 6 | 4 | 2 | 2 | 2 | 2 | 2 | 2 | 8 | 4 | 2 | 8 | 10 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 |
| C4.7 | 0 | 0 | 0 | 2 | 0 | 0 | 4 | 0 | 6 | 4 | 4 | 6 | 2 | 6 | 4 | 2 | 4 | 0 | 0 | 4 | 0 | 2 | 0 | 0 | 2 | 0 | 2 | 2 | 4 | 4 | 6 | 4 | 0 |
| C5.1 | 0 | 4 | 0 | 0 | 4 | 2 | 2 | 0 | 2 | 2 | 2 | 0 | 0 | 6 | 2 | 0 | -4 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 |
| C5.2 | 4 | 6 | 0 | 0 | 4 | 8 | 8 | 6 | 6 | 6 | 8 | 0 | 10 | 6 | 6 | 10 | 6 | 6 | 10 | 0 | 6 | 4 | 0 | 2 | 8 | 0 | 0 | 6 | 8 | 0 | 0 | 0 | 0 |
| C5.3 | 2 | 4 | 0 | 4 | 4 | 8 | 8 | 6 | 8 | 6 | 8 | 0 | 4 | 0 | 2 | 8 | 6 | 4 | 0 | 0 | 4 | 2 | 0 | 0 | 4 | 10 | 0 | 4 | 2 | 0 | 0 | 0 | 0 |
| C5.4 | 2 | 6 | 0 | 0 | 8 | 8 | 6 | 0 | 6 | 6 | 6 | 0 | 8 | 2 | 2 | 2 | 4 | 0 | 4 | 0 | 4 | 0 | 0 | 2 | 8 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 4 |
| C6.1 | 0 | 6 | 0 | 4 | 0 | 2 | 0 | 0 | 4 | 2 | 4 | 4 | 6 | 6 | 6 | 8 | 2 | 2 | 8 | 4 | 2 | 0 | 6 | 0 | -4 | 4 | 0 | 8 | 0 | 8 | 4 | 6 | 0 |
| C6.2 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | -4 | 0 | -2 | 4 | 6 | 8 | 8 | 6 | 4 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | -2 | 4 | 0 | 8 | 10 | 0 | 0 | 4 | 0 |
| C6.3 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 8 | 8 | 6 | 8 | 6 | 6 | 0 | 0 | 2 | 0 | 0 | 4 | 0 | -6 | 4 | 0 | 8 | 8 | 10 | 0 | 4 | 0 |
| C6.4 | 0 | 4 | 0 | 4 | 0 | 2 | 0 | 2 | 2 | 0 | 6 | 6 | 8 | 4 | 4 | 4 | 8 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | -2 | -2 | 0 | 6 | 10 | 6 | 2 | 0 | 0 |
| C6.5 | 4 | 2 | 0 | 6 | 2 | 2 | -2 | 0 | 6 | 4 | 8 | 6 | 6 | 8 | 10 | 10 | 10 | 0 | 0 | 2 | 4 | 0 | 4 | 0 | 4 | 6 | 6 | 10 | 4 | 6 | 2 | 2 | 0 |

## 2 The Investigated State Reduction Methods

In the following a novel state reduction approach with three different distance definitions are proposed and investigated. The three methods are different only in the metrics representing the similarity between the factors. This approach may be considered as a strong generalization of the state reduction procedure of sequential circuits or finite state machines, cf. [7]. The essence of the methods is to create clusters from the factors and apply these clusters as the new factors in the new, simplified model. The base for binding these clusters is the use of a tolerance relation (a reflexive and symmetric, but non-transitive relation) [6], cf. [4]. The reduced model is easier to interpret and to use.

Initially, the clusters are disjoint sets of factors, each of them containing one factor only. $K_i = \{C_i\}$ for every $i = 1 \dots n$ where $K_i$ is the $i$th cluster, $C_i$ is the $i$th factor (called 'concept' in the FCM theory) and $n$ is the number of factors in the model (thirty-three in our case). Next, an agglomerating strategy is applied for all of the clusters. For those factors not included in the current cluster, the 'distance' $(d)$ between the factor and the cluster (that is, all members of the cluster) is measured with the chosen metric.

The first method uses the absolute difference between the connections starting from two factors ($C_i$ and $C_j$) to the third ($C_k$), where $i \neq j \neq k$, $i = 1 \dots n, j = 1 \dots n, k = 1 \dots n$. If this difference is less than a specified ε value ($|w_{ik} - w_{jk}|/2 < \varepsilon$ and $|w_{ki} - w_{kj}|/2 < \varepsilon$, where $w$ specifies the sign and magnitude of connections between factors), the current factor is added to the cluster (metric 'A'). The calculation of metric 'A' and the whole process is described more precisely by the following C-style pseudo-code (see Fig. 1 and Fig. 2).

```
function isNearA(i, j, eps)  // i, j = factor indexes
  near = true;               // eps = ε
  for(k=0; k<n and near==true; k++) // n = no. of factors
    if(k!=i and k!=j)
      if((abs(w(i, k) - w(j, k))/2.) >= eps or
      (abs(w(k, i) - w(k, j))/2.) >= eps)
        near = false // w(i, k) = $w_{ik}$
  return near
```

**Fig. 1.** Calculation of metric 'A'

```
function buildCluster(initialFactor, eps)
  c = {initialFactor}
  for(i=0; i<n; i++)
    if(i != initialFactor)
      member = true
      while(member and hasNextElement(c))
        j = nextElement(c)
        member = isNear(j, i, eps)
        if(member)
          c = c + {i}
  return c

function buildAllClusters(eps)
  clusters = {}
  for(i=0; i<n; i++)
    k = buildCluster(i, eps)
    if(!isElementOf(k, clusters))
      clusters = clusters + {k}
  return clusters
```

**Fig. 2.** Pseudo-code of the state reduction algorithm, Part 1

The `buildAllClusters` function (see Fig. 2) initiates the state reduction process, requests the building of clusters and provides the uniqueness of them. Function `buildCluster` creates a new cluster using the specified initial factor. It investigates the other factors using the `IsNearA` function and adds the current factor to the cluster depending on the return value of `isNearA`.

Having identified the factors of the reduced model, the elements of the connection matrix have to be calculated. Function `getWeight` describes this calculation. The function returns the weight of the connection between clusters $a$ and $b$. The weight is calculated as the average weight of the connections between the factors of cluster $a$ to cluster $b$. According to the original definition of FCM self-loops with zero weight connections are ignored (see Fig. 3).

```
function getWeight(a, b)
  count = 0
  sum = 0
  while(hasNextElement(a))
    i = nextElement(a)
    while(hasNextElement(b))
      j = nextElement(b)
      if(i != j)
        count = count + 1
        sum = sum + w(i, j)
  if(count == 0)
    return 0
  else
    return sum/count
```

**Fig. 3.** Pseudo-code of the state reduction algorithm, Part 2

Obviously, ε is the critical parameter of the state reduction. If ε is too small, no or only a few factor-merges can be made, and then no real model simplification will be done. On the other hand, if the value of ε is too high, e.g. 1, the whole sensitivity matrix ‘collapses’ into a 1 by 1 matrix and the clusters of factors become meaningless. The useful values of ε are somewhere in the [0; 1] interval (according to the possible range of $w$), but the best suitable values are different in every practical case.

Unfortunately, here quite high values of ε were required in practice (see Table 3) in order to reduce the number of factors efficiently, because a small proportion of the interactions between the factors are often ‘too different’ from the investigated member of the cluster.

In order to obtain a more efficient reduction a second metric was applied (metric ‘B’, see `isNearB` in Fig. 4). In this case the current factor is added to the cluster if the difference between factors is less than a specified ε value except for a predefined (low) proportion ($p$) of the cases. The other parts of the algorithm remain the same.

**Table 3.** The number of factors in the reduced connection matrix, using different metrics

| Metric 'A' | | Metric 'B' | | | Metric 'C' | |
|---|---|---|---|---|---|---|
| $\varepsilon$ | No. of factors | $\varepsilon$ | $p$ | No. of factors | $\varepsilon$ | No. of factors |
| 0.3 | 28 | 0.1 | 0.2 | 30 | 0.011 | 30 |
| 0.4 | 25 | 0.2 | 0.05 | 30 | 0.016 | 28 |
| 0.5 | 18 | 0.2 | 0.1 | 26 | 0.022 | 24 |
| 0.6 | 15 | 0.2 | 0.2 | 23 | 0.027 | 22 |
| 0.7 | 12 | 0.3 | 0.05 | 23 | 0.04 | 20 |
| 0.8 | 4 | 0.3 | 0.1 | 21 | 0.048 | 18 |
| | | 0.3 | 0.2 | 15 | 0.054 | 15 |
| | | 0.4 | 0.05 | 19 | 0.06 | 12 |
| | | 0.4 | 0.1 | 10 | | |

```
function isNearB(i, j, eps, p) // i, j = factor indexes
  near = 0                     // eps = ε, p = $p$
  far = 0
  for(k=0; k<n; k++)           // n = number of factors
    if(k!=i and k!=j)
      if((abs(w(i, k) - w(j, k))/2.) < eps)
        near = near + 1// w(i, k) = $w_{ik}$
      else
        far = far + 1
      if((abs(w(k, i) - w(k, j))/2.) < eps)
        near = near + 1
      else
        far = far + 1
  if(near==0 or far/near >= p)
    return false
  else
    return true
```

**Fig. 4.** Calculation of metric 'B'

The values of $p$ are chosen from the interval (0; 1). Larger values of $p$ can help extend the clusters even if in some number of cases the 'distance' measured by the applied metric is greater than the allowed ε value. Too high values of $p$ would enable again to merge almost every factor into the same cluster, regardless of the value of ε, thus it would lead again to a meaningless model. The state reduction described above is similar to classification, where the number of classes is estimated in an appropriate way. The results given by metric 'B' are presented also in Table 3.

Metric 'B' made it possible to decrease the value of ε while the number of factors obtained after reduction was comparably low. The third method uses normalized, squared Euclidean distance (metric 'C'), and is detailed in Fig. 5. Table 3 shows the results of the third method, as well.

```
function isNearC(i, j, eps)
  sum = 0;  // i, j = factor indexes, eps = ε
  for(k=0; k<n; k++) // n = number of factors
    if(k!=i and k!=j)
      dout = w(i, k)-w(j, k) // w(i, k) = $w_{ik}$
      sum = sum + dout * dout
      din = w(k, i)-w(k, j)
      sum = sum + din * din
  if(sum / ((n-2)*8) < eps)
    return true
  else
    return false
```

**Fig. 5.** Calculation of metric ‘C’

## 3 Results

In the next, the authors give an overview about and shortly analyze the driving forces and impact of IWMS upon the results of the state reduction method (Table 4, Table 5). Integrated modeling requires not only the consideration of the technical and economic system elements, but also social, environmental, legal and institutional factors, furthermore their sub-factors. In these cases, to deal with the situations where data at hand are often insufficient for an entire quantitative analysis and the uncertainty is high, a series of non-quantifiable elements become important.

**Table 4.** An example of clusters as a result of state reduction (metric ‘C’, ε = 0.05)

| Cluster ID | Reduced concepts |
|---|---|
| Q1 | C3.1 + C3.2 + C3.3 + C3.4 + C3.5 + C1.1 + C6.1 |
| Q2 | C3.3 + C3.4 + C3.5 + C3.6 + C5.4 + C1.2 |
| Q3 | C2.1 + C2.3 + C2.5 + C5.1 + C1.3 + C4.4 + C4.5 |
| Q4 | C3.6 + C2.2 |
| Q5 | C3.3 + C3.5 + C2.3 + C2.5 + C5.3 + C1.1 + C1.4 + C6.4 + C4.3 + C4.4 |
| Q6 | C2.4 + C2.6 |
| Q7 | C3.1 + C3.3 + C3.5 + C2.5 + C1.1 + C1.4 + C6.1 + C6.4 + C4.3 + C4.4 |
| Q8 | C3.1 + C2.5 + C2.6 + C4.5 |
| Q9 | C3.3 + C3.5 + C2.5 + C5.2 + C5.3 + C1.1 + C1.4 + C1.5 + C6.4 + C4.4 |
| Q10 | C3.3 + C3.5 + C3.6 + C5.3 + C5.4 + C1.2 |
| Q11 | C3.2 + C3.3 + C3.4 + C3.5 + C5.4 + C1.1 + C1.2 + C6.1 |
| Q12 | C3.1 + C3.3 + C2.5 + C1.1 + C1.3 + C1.4 + C6.1 + C6.4 + C4.3 + C4.4 + C4.5 + C4.6 |
| Q13 | C3.3 + C3.4 + C3.5 + C5.4 + C1.1 + C1.2 + C1.5 + C6.4 + C4.3 |
| Q14 | C3.1 + C3.3 + C3.4 + C3.5 + C1.1 + C6.1 + C6.2 + C6.4 + C4.3 |
| Q15 | C3.1 + C3.3 + C1.1 + C1.3 + C1.4 + C6.1 + C6.2 + C6.3 + C6.4 + C4.3 |
| Q16 | C3.3 + C3.5 + C2.3 + C5.3 + C1.1 + C1.2 + C1.4 + C6.4 + C6.5 + C4.3 + C4.4 |
| Q17 | C5.2 + C5.3 + C1.1 + C1.2 + C1.4 + C1.5 + C4.1 + C4.4 + C4.5 |
| Q18 | C2.5 + C5.2 + C5.3 + C1.1 + C1.4 + C1.5 + C4.2 + C4.4 + C4.5 + C4.6 |
| Q19 | C2.3 + C5.1 + C5.3 + C1.1 + C1.2 + C1.3 + C1.4 + C4.3 + C4.4 + C4.5 + C4.7 |

IWMSs are organized along spatial and temporal scales. While modeling the system, it leads to the appearance of the connections and interaction between its factors and sub-factors. The factors of these systems are connected via material, energy, money and information flows and form a complex phenomenon through legal regulation.

During studying the results of the state reduction methods, it could be recognize that one of the constitutive factors in the new model (Table 4) is factor 3 (economy) and its sub-factors. This factor represents the major function among the factors in the clusters. Economy plays a significant role in financing the systems, influences the possibilities of depreciation and re-source development and also the structure of the waste processing industry. Its combination with engineering knowledge reflects the different forms of engagement in engineering education and development of technologies as well and can provide different perspectives on the same complex issue. This combination of factors tells about the relationship of the sub-factors as parts of the systems and highlights the question ‘What is important in this system?’.

**Table 5.** Connection matrix of clusters as a result of state reduction (see Table 4)

| Cluster ID | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 | Q13 | Q14 | Q15 | Q16 | Q17 | Q18 | Q19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Q1 | 0 | 0.50 | 0.07 | 0.47 | 0.23 | 0.31 | 0.25 | 0.16 | 0.28 | 0.41 | 0.44 | 0.17 | 0.36 | 0.28 | 0.16 | 0.25 | 0.25 | 0.19 | 0.11 |
| Q2 | 0.37 | 0 | 0.16 | 0.62 | 0.23 | 0.22 | 0.23 | 0.17 | 0.32 | 0.41 | 0.42 | 0.16 | 0.37 | 0.24 | 0.14 | 0.27 | 0.26 | 0.22 | 0.14 |
| Q3 | 0.16 | 0.16 | 0 | 0.17 | 0.16 | 0.26 | 0.16 | 0.25 | 0.19 | 0.16 | 0.14 | 0.18 | 0.13 | 0.11 | 0.12 | 0.13 | 0.18 | 0.23 | 0.16 |
| Q4 | 0.21 | 0.49 | 0.24 | 0 | 0.30 | -0.50 | 0.15 | 0.03 | 0.31 | 0.51 | 0.28 | 0.08 | 0.32 | 0.19 | 0.05 | 0.35 | 0.20 | 0.22 | 0.19 |
| Q5 | 0.34 | 0.38 | 0.19 | 0.42 | 0 | 0.39 | 0.25 | 0.28 | 0.23 | 0.33 | 0.32 | 0.19 | 0.25 | 0.24 | 0.17 | 0.21 | 0.19 | 0.18 | 0.15 |
| Q6 | -0.06 | 0.18 | -0.17 | 0.05 | 0.00 | 0 | -0.04 | -0.31 | 0.10 | 0.25 | 0.11 | 0.00 | 0.18 | -0.03 | -0.01 | 0.10 | 0.29 | 0.14 | 0.13 |
| Q7 | 0.41 | 0.42 | 0.14 | 0.33 | 0.24 | 0.34 | 0 | 0.29 | 0.25 | 0.36 | 0.38 | 0.23 | 0.30 | 0.31 | 0.22 | 0.24 | 0.20 | 0.21 | 0.14 |
| Q8 | 0.16 | 0.22 | 0.11 | 0.25 | 0.22 | 0.31 | 0.17 | 0 | 0.25 | 0.23 | 0.21 | 0.17 | 0.19 | 0.10 | 0.09 | 0.20 | 0.35 | 0.32 | 0.20 |
| Q9 | 0.40 | 0.42 | 0.25 | 0.46 | 0.26 | 0.51 | 0.28 | 0.35 | 0 | 0.37 | 0.38 | 0.21 | 0.28 | 0.26 | 0.17 | 0.25 | 0.25 | 0.25 | 0.18 |
| Q10 | 0.33 | 0.46 | 0.23 | 0.64 | 0.23 | 0.35 | 0.22 | 0.22 | 0.32 | 0 | 0.38 | 0.15 | 0.33 | 0.21 | 0.12 | 0.25 | 0.27 | 0.24 | 0.16 |
| Q11 | 0.42 | 0.51 | 0.10 | 0.50 | 0.23 | 0.39 | 0.26 | 0.18 | 0.31 | 0.42 | 0 | 0.18 | 0.38 | 0.29 | 0.18 | 0.25 | 0.25 | 0.20 | 0.12 |
| Q12 | 0.38 | 0.37 | 0.19 | 0.29 | 0.26 | 0.36 | 0.30 | 0.36 | 0.27 | 0.33 | 0.35 | 0 | 0.28 | 0.29 | 0.22 | 0.24 | 0.24 | 0.26 | 0.17 |
| Q13 | 0.47 | 0.50 | 0.13 | 0.47 | 0.26 | 0.32 | 0.31 | 0.23 | 0.31 | 0.42 | 0.48 | 0.22 | 0 | 0.34 | 0.24 | 0.28 | 0.23 | 0.19 | 0.14 |
| Q14 | 0.49 | 0.52 | 0.06 | 0.38 | 0.25 | 0.18 | 0.31 | 0.18 | 0.27 | 0.43 | 0.50 | 0.22 | 0.39 | 0 | 0.25 | 0.26 | 0.20 | 0.16 | 0.12 |
| Q15 | 0.52 | 0.52 | 0.08 | 0.30 | 0.25 | 0.25 | 0.33 | 0.28 | 0.25 | 0.43 | 0.49 | 0.25 | 0.36 | 0.41 | 0 | 0.25 | 0.17 | 0.16 | 0.11 |
| Q16 | 0.43 | 0.47 | 0.17 | 0.45 | 0.26 | 0.44 | 0.30 | 0.30 | 0.29 | 0.42 | 0.41 | 0.23 | 0.32 | 0.33 | 0.23 | 0 | 0.21 | 0.21 | 0.15 |
| Q17 | 0.45 | 0.43 | 0.32 | 0.46 | 0.34 | 0.62 | 0.37 | 0.49 | 0.40 | 0.40 | 0.42 | 0.31 | 0.34 | 0.32 | 0.23 | 0.30 | 0 | 0.40 | 0.23 |
| Q18 | 0.36 | 0.34 | 0.36 | 0.37 | 0.32 | 0.60 | 0.32 | 0.51 | 0.36 | 0.33 | 0.33 | 0.29 | 0.27 | 0.24 | 0.19 | 0.26 | 0.36 | 0 | 0.25 |
| Q19 | 0.36 | 0.35 | 0.19 | 0.34 | 0.24 | 0.43 | 0.28 | 0.34 | 0.28 | 0.33 | 0.32 | 0.24 | 0.27 | 0.28 | 0.22 | 0.22 | 0.21 | 0.23 | 0 |

## 4 Conclusions

A new state reduction approach was introduced to make the otherwise too complex connection matrix of IWMS model easier to handle and understand. Three different metrics were investigated and one of the simplified connection matrices resulted by metric ‘C’ was presented and evaluated.

The state reduction approach proved to be good to combine different type of factors and create clusters. It thereby provides a comprehensive and more thoroughly understanding of an IWMS as a technical-economic-social-environmental system.

The conclusions based on the results of state reduction should be viewed together with existing scientific knowledge. In the next period, it is the authors' intention to study further the assumptions, but also be open to insights gained from a systemic approach to deliver a method for decision making on sustainable regional waste management.

## 5 Future Research

The authors' purpose is to apply a modified version of the FCM. As usually it is accepted that causality is not self reflexive, i.e., a concept cannot cause itself, which means that the weight matrix always has '0' in its diagonal [8]. However, the real life systems query the justification of this theory therefore the authors' intention is to overview the operation of self loops which would fit better to reality.

In the present paper three metrics were applied which allow quite different degrees of reduction. The authors' goal is to develop further metrics in the future and to apply a more general concept reduction algorithm.

**Acknowledgement.** The authors would like to thank to TÁMOP-4.2.2.A-11/1/KONV-2012-0012, TÁMOP-4.1.1.C-12/1/KONV-2012-0017, to the Hungarian Scientific Research Fund (OTKA) K105529 and K108405 for the support of the research.

## References

1. Buruzs, A., Pozna, R.C., Kóczy, L.T.: Developing Fuzzy Cognitive Maps for Modeling Regional Waste Management Systems. In: Tsompanakis, Y. (ed.) Proceedings of the Third Int. Conference on Soft Computing Technology in Civil, Structural and Environmental Engineering, Paper 19. Civil-Comp Press, Stirlingshire (2013)
2. Buruzs, A., Hatwágner, M.F., Pozna, R.C., Kóczy, L.T.: Advanced Learning of Fuzzy Cognitive Maps of Waste Management by Bacterial Algorithm. In: Proceedings of IFSA World Congress and NAFIPS Annual Meeting, pp. 890–895. IEEE (2013)
3. Buruzs, A., Hatwágner, M.F., Kóczy, L.T.: Modeling integrated sustainable waste management systems by fuzzy cognitive maps and the system of systems concept. Technical Transactions series Automatic Control (accepted for publication, 2014)
4. Das, M., Chakraborty, M.K., Ghoshal, T.K.: Fuzzy tolerance relation, fuzzy tolerance space and basis. Fuzzy Sets and Systems 97(3), 361–369 (1998)
5. Hatwágner, M.F., Buruzs, A., Földesi, P., Kóczy, L.T.: Strategic decision support in waste management systems by state reduction in FCM models. In: Siah, Y.K. (ed.) ICONIP 2014, Part III. LNCS, vol. 8836, pp. 447–457. Springer, Heidelberg (2014)
6. Klir, G.J., Folger, T.A.: Fuzzy Sets, Uncertainty and Information. Prentice Hall (1987)
7. Paull, M.C.: Minimizing the Number of States in Incompletely Specified Sequential Switching Functions. IRE Transactions on Electronic Computers EC-8(3), 356–367 (1959)
8. Carvalho, J.P.: On the Semantics and the Use of Fuzzy Cognitive Maps in Social Sciences. In: IEEE World Congress on Computational Intelligence, WCCI 2010, July 18-23, pp. 18–23. CCIB, Barcelona (2010)

# Building Renovation Cost Optimization with the Support of Fuzzy Signature State Machines

Gergely I. Molnárka and László Tamás Kóczy

[1] Department of Building Constructions and Architecture, Széchenyi István University, Egyetem tér 1, Győr, Hungary
[2] Department of Automation, Széchenyi István University, Egyetem tér 1, Győr, Hungary

**Abstract.** The renovation of a significant segment of housing sector in Budapest, Hungary is overdue. Besides other causes, the absence of any effective tool that may support the decisions of the ownership communities in determining the technically verified steps and solutions in repair processes hinders the improvement of the physical condition of old residential houses.

In this paper we propose a new formal method and approach for generating such tool that considers the costs and feasibilities of alternative renovation processes with professional data obtained from building diagnostics surveys, technical reports and contractors' billing database. As a case study, a comprehensive renovation chain of the roof structure of a pre-war residential house will be evaluated.

**Keywords:** urban-type residential house, building renovation chain, fuzzy signature, fuzzy state machine.

## 1 Introduction

Among others, the composition of the residential building stock demonstrates well the life quality of the given location. In this respect, the distribution of almost one million apartments in Budapest, Hungary presents a fairly heterogeneous picture: the apartments in urban-type residential buildings (built before the WWII) take 27%; flats in industrialized blocks (built between 1950-1990) add up to 33%; the ratio of dwelling houses is 29%; else 11% [1].

The pre-war urban-type residential houses represent a peculiar set of the apartment building stock. The main characteristics of this building type are the decorative street façades, the existence of courtyard and air-shafts; the presence of apartments differentiated by size and orientation in the same building and the traditional masonry and timber load bearing structures.

With a mathematical approach a house can be considered as an ordered composition of specific building components (made of various materials and technology). With the support of the building construction manuals [2], [3], etc. and thesauri [4], [5] their role, the position and the interrelations among them can be easily conceived; their functional lattice may be mapped. As a result, a hierarchical structure of building components gives a comprehensive picture of the

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_13

overall house. Following this structure, sets of components exist that can be decomposed into subsets (e.g. the tiling and the tile battens form a subset of roof coverings that is a member of the set of roof structure).

Due to several causes the average physical condition of these old residential buildings is below standard. The symptoms of overall physical obsolescence and deterioration are clearly observable (even in representative urban areas):

- Missing wall and plinth damp proof courses;
- Uncertain load-bearing performance of the slab systems and side corridors;
- Out-dated mechanical and electrical systems;
- Broken and crumbling building envelope, floor covering and supplementaries;
- Cracked and perforated roof tiling and flashings;
- Low energetic performance of air-shaft partitioned walls and openings;
- Detached and missing plaster and finishing of the façade surfaces.

In practice, the physical condition of the residential houses are deemed to be constant, however, it is far from the truth: that is why the buildings require continuous surveillance. Its long-lasting lag commits waste, or, in certain cases, serious deteriorations that may launch irreversible decay process.The Fig.1 represents how the renovating interventions can partially restore the performance of a building during the total service life.

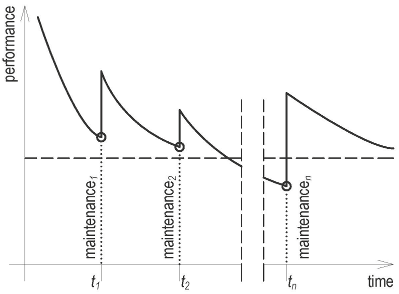


**Fig. 1.** Relation of the maintenance and the performance of a building, based on [6]

If the failure of a building component is discovered, many alternative solutions can be found for its repair or replacement. It is also noticeable that if a complete renovation process is examined that consists of numerous repair tasks, the necessary steps may follow each other in various ways, constituting high number of renovation chains. Following the internal logics of the building renovation directives that are determined in commonly known technical reports, such as [7], [8], etc., some other attributes of the repair steps and renovation chains are observable. Without any details, the most important features are: the dependency, the interchangeability and the cost (and as a consequence, the quality) of the result. These facts make the renovation more complex, therefore it is hard to define the most effective renovation chain. It also has to be noticed that any improvement in physical condition of a building component or a set of components does not influence the state of other components or sets directly (e.g. the stairway renovation does not improve the performance of air-shaft partition walls, logically). It means that the state of the building components determine the states of the given building together, while the state of components may have individual transitions.

Considering the facts that a house is an ordered set of (partially) related building components and several repairs may be realized individually, however, these renovation courses are (partially) interdependent. In addition to this, the working stages in renovation are characteristically overlapping, consecutive, or independent from each other. Nevertheless, these stages may also be divided into smaller units (with harmonization to the owners' budget).

For designing a tool that may support decision-making in the whole renovation process, first the possible steps have to be examined, taking their respective importance, interrelations and some non-constructional (e.g. economic, aesthetic, etc.) aspects into consideration. In our approach, the main goal of the tool is to examine each maintenance step that is essential for renovating the given building from its present (deteriorated) state to the acceptable (renovated) state. There are several factors influencing the schedule of the rehabilitation process. Without any details, the most important such factors are: the *grade of danger* caused by the observed failure, the *interrelations* among deteriorated building components and decays, and the presence of a protocol for the repair, the *financial schedulability*, some complex *logistics aspects*, etc.

After revealing the attributes of the buildings (as sets of numerous building components), their deterioration process and the complexity of building renovation, in the followings the proposed model will be introduced.

## 2 Proposal: A Model for Renovation Cost Optimization

In our previous work [9] the theoretical background of the model and attempt for handling the compound technical problem has introduced. The proposed tool that we call *Fuzzy Signature State Machines (FSSM)* is based on the principles of fuzzy signatures with the combination of fuzzy state machines.

Fuzzy signatures are multicomponent fuzzy descriptors, where any component might be a further nested vector. The next is a very simple example:
$\mu_{A_{fs}} = [\mu_1, \mu_2, [\mu_{3_1}, \mu_{3_2}, [\mu_{3_{3_1}}, \mu_{3_{3_2}}, \mu_{3_{3_3}}]], \mu_4, [\mu_{5_1}, \mu_{5_2}], \mu_6]^T$

The first advantage of using fuzzy signatures is that here any closer grouping and sub-grouping of fuzzy features may be given. Fuzzy signatures are associated with an aggregation system. Each sub-component set may be aggregated by its respective aggregation operation, thus reducing the sub-component to one level higher. The above example has the following associated aggregation structure: $\{a_0\{a_3\{a_{3_3}\}\}\{a_5\}\}$, where each $a_\circ$ denotes an aggregation, particularly the one associated with the child node $x_\circ$ associated with $\mu_\circ$, thus the following example signature might be reduced upwards as follows:

$$\mu_{A_{fs}} \Rightarrow [\mu_1, \mu_2, [\mu_{3_1}, \mu_{3_2}, \mu_{3_3} = a_{3_3}(\mu_{3_{3_1}}, \mu_{3_{3_2}}, \mu_{3_{3_3}}), \mu_4, \mu_5 = a_5(\mu_{5_1}, \mu_{5_2}), \mu_6]^T$$
$$\Rightarrow [\mu_1, \mu_2, \mu_3 = a_3(\mu_{3_1}, \mu_{3_2}, \mu_{3_3}), \mu_4, \mu_5, \mu_6]^T \Rightarrow \mu_0 = a_0(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6)$$

This kind of membership degree reductions is necessary when the data are partially of different structure, e.g. some of the sub-components are missing. Then operations among fuzzy signatures with partially different structure may be carried out, by finding the *largest common sub-structure* and reducing all signatures up to that substructure.

The most important reason for applying fuzzy signatures here as the starting point is the fact that the structure of building condition evaluations follows the architectural and civil engineering common sense, where the sub-structures and components of each building are arranged in hierarchical tree-like structures, where the whole building might be presented by the root of the tree and each mayor sub-component is a first level branch, with further sub-branches describing sub-sub-components.

The fuzzy signature structures have a characteristic feature that plays an important role in handling uneven building diagnostics surveys in depth and accuracy of the evaluation: the membership degree reduction may help in operations among fuzzy signatures with partially different structure by finding the largest common sub-structure and reducing all signatures to that level. As an example, maybe in "survey $A$" the roof structure is considered as a single component of the house and is evaluated by a single linguistic quality label, while in "survey $B$" this is done in detail and tiles, load-bearing structure, tinsmith work, chimney shafts, etc. are evaluated individually.

In our earlier work [10] we presented an approach where the fuzzy signatures could be deployed for describing existing residential houses in order to support decisions of Local Authorities concerning when and how these buildings should be renovated involving non-measurable (and subjective) factors. In that research a series of theoretically arrangeable features were taken into consideration and eventually a single aggregated fuzzy membership value could be calculated on the basis of available detailed expert evaluation sheets and questionnaires. In these models, however, the available information does not support any decision strategy concerning actual sequence of the measures leading to complete renovation; and they are also insufficient to optimize the sequence from the aspect of local or global cost efficiency.

In the followings the mathematical model of the proposed renovation chain will be introduced.

Finite State Machines are determined by the sets of input states $X$, internal states $Q$, and the transition function $f$. The latter determines the transition that will occur when a certain input state change triggers state transition. There are several alternative (but mathematically equivalent) models known from the literature. For simplicity the following is assumed as the starting point of our new model:

$$A = \langle X, Q, f \rangle \tag{1}$$

$$f : X \times Q \to Q, \text{ where } X = \{x_i\} \text{ and } Q = \{q_i\} \tag{2}$$

Thus, a new internal state is determined by the transition function as follows:

$$q_{i+1} = f(x_i, q_i) \tag{3}$$

The transition function maybe interpreted with help of a relation $R$ on $X \times Q \times Q$,where

$$\begin{aligned} R(x_i, q_j, q_k) = 1, \text{ if } f(x_i, q_j) = q_k \text{ and} \\ R(x_i, q_j, q_k) = 0, \text{ if } f(x_i, q_j) \neq q_k \end{aligned} \tag{4}$$

The states of the finite state machine are elements of $Q$. In the present application an extension to fuzzy states is considered in the following sense. Every aspect of the phenomenon to model is represented by a state universe of sub-states $Q_i$. The states themselves are (fuzzy) subsets of the universe of discourse state sets, so that within $Q_i$ a frame of cognition is determined (its fineness depending on the application context and on the requirements toward the optimisation algorithm), so that typical states like "Totally intact", "Slightly damaged", "Medium condition", etc., up to "Dangerous for life" are considered. Any transition from one state to the other (improvement of the condition, refurbishment or renovation) involves a certain cost $c$. In the case of a transition from $q_i$ to $q_j$ it is expressed by a membership value $\mu_{i_j} = c(q_i, q_j)$. In our model the added cost $\Sigma\mu_{i_j}$ along a path $q_{i_1} \to q_{i_2} \to \cdots \to q_{i_n}$ is not usually equivalent with the cost of the transition $\mu_{i_n}$ along the edge $q_{i_l} \to q_{i_n}$. This is in accordance with the non-additivity property of the fuzzy (possibility) measure and is very convenient in our application, as it is also not additive in the case of serial renovations. The Fig. 2 illustrates the initial (deteriorated), internal and the accept (renovated) state of a possible renovation process; the bottom diagram represents an internal ($\mu_{12}$)transition between $q_1$ and $q_2$ states.

In the case of fuzzy signature machines each of the leaves contains a sub-machine with the above property. The parent leave of a certain sub-graph is constructed from the child leaves, so that the sub-machine

$A^i = A^{i_1} \times A^{i_2} \times \cdots \times A^{i_m}$, and thus the states of $A^i$ are $Q^i = Q^{i_1} \times Q^{i_2} \times \cdots \times Q^{i_n}$, so that the transition $Q^{j_1} \to Q^{j_2}$ in this case means the parallel (or subsequent) transitions $q_{j_{1_1}} \to q_{j_{1_2}} \times q_{j_{2_1}} \to q_{j_{2_2}} \times \cdots \times q_{j_{n_1}} \to q_{j_{n_2}}$. A special aggregation is associated with each leaf; similarly as it is in the fuzzy signatures, however, in this case the aggregation calculates the resulting cost $\mu_{j_{1_2}}$ of the transition $q_{j_1} \to q_{j_2}$, so that

$$\mu_{j_{1_2}} = c(q_{j_1}, q_{j_2}) = a_j(c(q_{j_{1_1}}, q_{j_{1_2}}), c(q_{j_{2_1}}, q_{j_{2_2}}), \cdots, c(q_{j_{n_1}}, q_{j_{n_2}})) \tag{5}$$

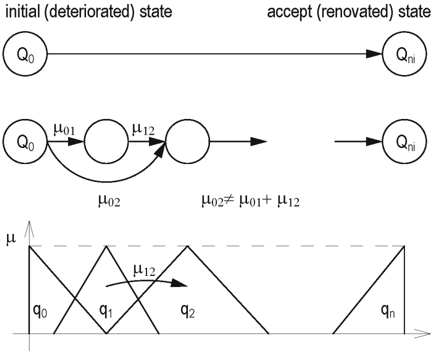


**Fig. 2.** Initial, internal and accept states of a renovation process

where $a$ stands for the respective aggregation. (Note that these aggregations sometimes do not satisfy the symmetricity property of the general axiom structure of aggregations, thus it may be referred as a "non-symmetric aggregation".)

It should be mentioned that the selection of aggregation operator is a key issue that may determine the final result of the model; however, the signature structure makes the application of different aggregation methods possible for each node.

In the followings, the proposed model will be illustrated with a case study based on a typical roof structure.

## 3 Comprehensive Roofing Renovation Chain Evaluation

With the support of common technical literatures mentioned previously the renovation of the urban type pre-war residential houses can be decomposed into eight (more or less) distinct sequences corresponding to the statements in *Section 1.* With further analyses these sequences may be decomposed into sub-sequences; at the leaves of this tree-like structure are represented by the states of the constructions and building elements. With the application of the existing building diagnostics surveys the initial state of examined building components or groups of components can be determined (note that the fuzzy signature structure obtains solution for handling missing or undiscovered data in the child node level).

In the followings, the operation of the proposed model is illustrated with the FSSM of the roofing system of an existing residential house (built in 1912) located in the historic district of Budapest. This roof gives a representative example of the roofs of pre-war urban-type houses with typical components and failures. This example is suitable for illustrating the entire FSSM of the residential house due to the followings: *a)* the complexity of the roofing system reflects to the complexity of the house as a compound of building constructions; *b)* the profoundness of available building diagnostics assessment of the overall roofing system is uneven: certain components are evaluated in details (e.g. the roof framework), some components are described with comprehensive data only (e.g. auxiliary elements).

The Roofing System ($A^1$) state machine has seven sub-state machines ($A^{1_1}$ to $A^{1_7}$ in Table 1) that correspond to the groups of components; the initial states of these groups determine jointly the physical condition of the entire roof structure.

**Table 1.** Observations of initial states of the child nodes of $A^1$ roofing system based on available building diagnostics surveys

| Node | Name | Observation |
|---|---|---|
| $A^{1_1}$ | Roof Framework | Sporadic fungus attack; warping; cracking |
| $A^{1_2}$ | Roof Covering | Discontinuity; cracking; leaking |
| $A^{1_3}$ | Tinsmith Work | leaking; missing elements; detachment |
| $A^{1_4}$ | Chimney Shaft | Low stability; freezing decay |
| $A^{1_5}$ | Auxiliaries | disused elements |
| $A^{1_6}$ | Supplementaries | Dangerous parts; detachment; corrosion |
| $A^{1_7}$ | Firewalls | Detachment; missing parts |

In the starting point of the model applied the state machines on the leaves of the roofing system are

$$\begin{aligned} A^{1_1} &= \langle X^{1_1}, Q^{1_1}, f^{1_1} \rangle \\ A^{1_2} &= \langle X^{1_2}, Q^{1_2}, f^{1_2} \rangle \\ &\vdots \\ A^{1_7} &= \langle X^{1_7}, Q^{1_7}, f^{1_7} \rangle \end{aligned}$$

As a result, the parent node of these leaves is represented with

$$\begin{aligned} A^1 &= \langle X^1, Q^1, f^1 \rangle \text{ state machine, where} \\ X^1 &= X^{1_1} \times X^{1_2} \times \cdots \times X^{1_7} \text{ and} \\ Q^1 &= Q^{1_1} \times Q^{1_2} \times \cdots \times Q^{1_7} \text{ and} \\ f^1 &= f^{1_1} \times f^{1_2} \times \cdots \times f^{1_7} \end{aligned}$$

Several internal states ($q^i$) of each group of building components can be determined during the renovation process. The alternative renovation chains that are constituted by these states characterize different refurbishment behaviour. As an example, the total demolition and rebuilding may result in a failure-free but expensive renovation (in the model, the internal states and transitions indexed with $-1$ represent the demolished states ($q^i_{-1}$) and transition ($\mu^i_{-1s}$; $\mu^i_{t-1}$)). Another typical solution for the refurbishment is the "step-by-step" renovation: some groups of building components or the entire building are renovated in more phases. In this case the recurrence of some cost factors (e.g. auxiliary manpower, scaffolding, transportation, etc.) may result in unwanted rise of the total cost.

The state transitions ($\mu^i_{st}$) that define the $c$ costs among states are determined by the internal logics of building renovation protocols. Therefore some transitions have to be excluded (as an example none of the internal states may follow the "demolished" state). In the model these transitions are denoted as "restrained" transition.

After a detailed assessment of renovation processes the existence of $5 \times 7 \times 8 \times 6 \times 2 \times 4 \times 5 \times = 67,200$ state transitions may be determined. Due to this high number of transitions their illustration meets with difficulties.

Another look of internal states and transitions can be obtained if two sub-state machines are connected, disregarding the rest sub-machines. The Fig. 3 illustrates the combination of $A^{1_4}$ and $A^{1_5}$ state machines.

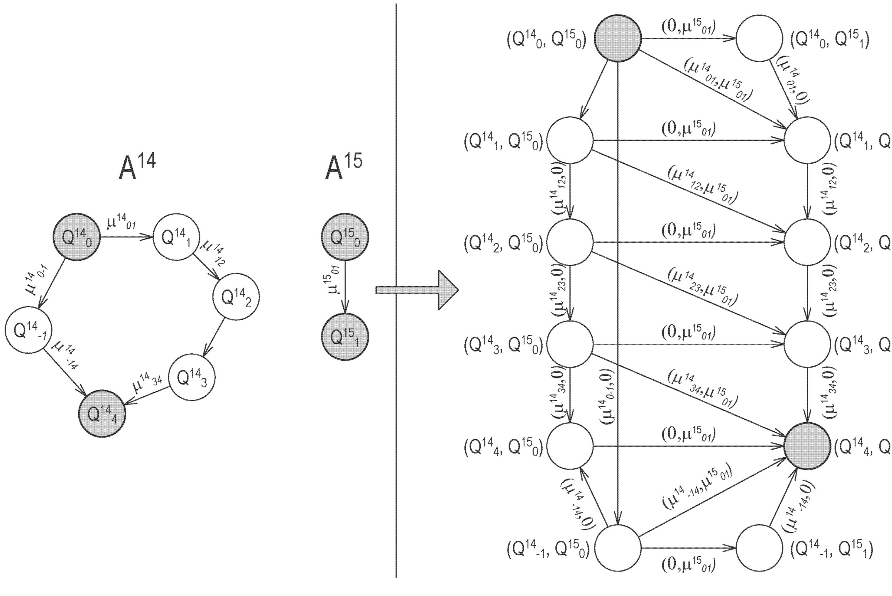


**Fig. 3.** Internal states (selection). Combination of $A^{1_4}$ and $A^{1_5}$ state machines

As the first step in the evaluation, the initial states of the child nodes has to be ascertained based on the concerning building diagnostics surveys (for simplifying the model, the $A^{1_1}$ roof framework node was evaluated separately and the aggregation result is applied on its parent node only). In the evaluation the commonly used linguistic categories were applied (the universe is normalized to [0;1], where 1 represents the intact state in each node); their partitions are represented by Ruspini partition-type [11] triangular and trapezoidal membership functions.

With the support of available contractors' billing database the specific costs were estimated: in the calculations the costs were normalized to [0;1], where 1 represents the total cost of demolition and transportation of existing structure and its reproduction on an acceptable quality that corresponds to the current regulations (the total floor area of the examined roofing system is 506.0 m$^2$; the total reproduction cost takes 84,000 USD).

The attributes of the building components and the simplicity of the method verify alike the application of arithmetic average type of Ordered Weighted Averaging Aggregation (OWA) operator as it was presented by Yager in [12]. In this case the $w$ weighting factors represent the importance of the renovation steps.

The Table 2 summarizes the initial and the accept values of the leaves of $A^1$ fuzzy signature state machines; the cost ratio and the weighting factors in aggregation are also indicated. It is clearly visible that with some simple and well-organized interventions the aggregated value of the roof structure may be improved effectively.

**Table 2.** The initial and accept values of the sub-state machines of roofing system

| **Node** | **Initial state** | **Costs** | **Accept state** | **Weight** |
|---|---|---|---|---|
| $A^{1_1}$ | $x_0^{1_1} = 0.65$ | 36.15% | $x_4^{1_1} = 0.85$ | $w^{1_1} = 0.85$ |
| $A^{1_2}$ | $x_0^{1_2} = 0.35$ | 30.79% | $x_6^{1_2} = 0.70$ | $w^{1_2} = 0.50$ |
| $A^{1_3}$ | $x_0^{1_3} = 0.55$ | 13.65% | $x_7^{1_3} = 0.75$ | $w^{1_3} = 0.45$ |
| $A^{1_4}$ | $x_0^{1_4} = 0.75$ | 8.87% | $x_4^{1_4} = 0.85$ | $w^{1_4} = 0.30$ |
| $A^{1_5}$ | $x_0^{1_5} = 0.20$ | 1.16% | $x_1^{1_5} = 0.65$ | $w^{1_5} = 0.80$ |
| $A^{1_6}$ | $x_0^{1_6} = 0.25$ | 4.57% | $x_4^{1_6} = 0.80$ | $w^{1_6} = 0.80$ |
| $A^{1_7}$ | $x_0^{1_7} = 0.60$ | 4.82% | $x_4^{1_7} = 0.85$ | $w^{1_7} = 0.25$ |
| $A^1$ | $x_0^1 = 0.42185$ | **100.0%** | $x_1^1 = 0.8320$ | |

The results admit of the conclusion that an effective state change may be obtained even with some simple and inexpensive interventions (such as the total dismounting of disused auxuliaries). Moreover, in case of applying sectioned renovation chain (with more internal states) it is advisable to assign an accept state with higher value, since it may improve the value of the entire structure significantly.

## 4 Conclusions and Future Work

During the calculations some important simplifications has to be done for constructing the overall model of the roofing system FSSM. Among others, the statement that in case of sectioned renovation processes the cost factors are divided proportionally among the species of work does not agree with the facts. Nevertheless, the proposed model may bring us up to the complex cost optimization of the residential buildings.

In practice, the optimization of the renovation chain of any sort of residential buildings always has a limited number of sequences; however this number might be rather high. This study revealed that the renovation process of the roofing system may have more than 67 thousand sequence. Because of the computational complexity of the optimization problem, a meta-heuristics with reasonably low complexity but good convergence expectations is proposed, such as the Bacterial Evolutionary/ Bacterial Memetic Algorithms, the Particle Swarm, or the Imperialist Competitive Algorithms.

**Acknowledgments.** The research was supported by TÁMOP-4.2.2.A-11/1/ KONV-2012-0012 and Hungarian Scientific Research Fund (OTKA) K105529, K108405.

## References

1. Csizmady, A., Hegedüs, J., Kravalik, Z., Teller, N.: Long-term housing conception and mid-term housing program of Budapest. Studio Metropolitana (2005) (in Hungarian)
2. Ching, F.: Building construction illustrated, 3rd edn. Wiley, New York (2001)
3. Pfeifer, G.: Masonry construction manual. Birkhäuser, Basel (2001)
4. Kunszt, G.: Thesaurus of the thematic information system of hungarian institute of building science. Technical report, TI, Budapest (1974) (in Hungarian)
5. Koppány, A.: Morphologic system of building constructiona. Technical report, Széchenyi István Technical University, Györ (1997) (in Hungarian)
6. Gibson, E.: Working with the performance approach in building (CIB w060). Technical Report 64, CIB W060, Rotterdam (1982)
7. Tóth, E.: Building Refurbishment Handbook. Verlag Dashöfer, Budapest (2005) (in Hungarian)
8. Giebeler, G. (ed.): Refurbishment manual: maintenance, conversions, extensions. Birkhäuser, Basel (2009)
9. Molnárka, G.I., Kóczy, L.T.: Optimizing complex building renovation process with fuzzy signature state machines. In: Siah, Y.K. (ed.) ICONIP 2014, Part II. LNCS, vol. 8835, pp. 573–580. Springer, Heidelberg (2014)
10. Molnárka, G.I., Kóczy, L.T.: Decision support system for evaluating existing apartment buildings based on fuzzy signatures. International Journal of Computers Communications & Control VI(3), 442–457 (2011)
11. Ruspini, E.H.: A new approach to clustering. Information and Control 15(1), 22–32 (1969)
12. Yager, R.: On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Transactions on Systems, Man and Cybernetics 18(1), 183–190 (1988)

# Cross-Platform Pathway Activity Transformation and Classification of Microarray Data

Worrawat Engchuan[1], Asawin Meechai[2], Sissades Tongsima[3],
and Jonathan H. Chan[1,*]

[1] Data and Knowledge Engineering laboratory,
School of Information Technology, Thailand
King Mongkut's University of Technology Thonburi
`worrawat.eng@st.sit.kmutt.ac.th`,
`jonathan@sit.kmutt.ac.th`
[2] Department of Chemical Engineering, Faculty of Engineering
King Mongkut's University of Technology Thonburi, Thailand
`asawin.mee@kmutt.ac.th`
[3] BioStatistics and Informatics laboratory,
Genome Institute,
National Center for Genetic Engineering and Biotechnology, Thailand
`sissades@biotec.or.th`

**Abstract.** One of the most challenging problems in microarray study is to analyze microarray data from different platforms. This will improve the reliability of the study, as number of samples is larger and it can be applied for rare disease study, for which only a few microarray data have been published. As different microarray platforms cover different number of genes, so the integrative study of two different platforms needs to be able to deal with the missing value issue. Many works have been done for cross-platform microarray data utilization but none of them have focused on gene-set based microarray data classification. In this study, we applied the Bayesian-based method to reconstruct the expression level of the missing genes before transforming it to the gene-set activity. Two gene-set activity transformation methods; Negatively Correlated Feature Set (NCFS-i) and Analysis-of-Variance Feature Set (AFS), were used to evaluate the performance of this method using actual microarray datasets. The results show that the imputation of missing data can improve the classification performance of the cross-platform study.

**Keywords:** microarray data, cross-platform analysis, gene-set activity transformation, negative correlation, Bayesian computation.

## 1 Introduction

Microarray technology is one of the most popular molecular biological techniques, which can be used to examine the expression levels of thousands of genes simulta-

---

[*] Corresponding author.

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems*,
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_14

neously. This technology has been also widely used in medical science in order to study the genetic disease and genetic disorder, resulting in the dramatic increase in publicly available microarray data [1-2]. The availability of tons of microarray datasets allows researchers to utilize them for testing new algorithms or to re-analysis of the data by recent advance techniques. In the study of complex diseases like cancers, the single analysis of only one microarray data is inadequate to unveil the complex mechanisms of the disease. So, the integrative study of multi-layer of biological data is desired for the study of the complex disease [3]. Pathway-based analysis, which transforms the gene expression levels into another form of data using pathway information so-called pathway activity, is one of the widespread techniques for integrative study. The classification using pathway activities has been shown to be more powerful than the use of gene expression levels [4-6]. Unfortunately, most of microarray analyses seem to be not significant for population study because the number of samples is low. In order to improve the quality of microarray data analysis, a technique for integrating multiple microarray datasets needs to be developed [7-8]. Many works have been done for correcting the confounding “batch effect”, which causes the variation of expression levels between datasets. So, adjusting this batch effect, the integration of microarray datasets will be more effective [9-10]. Thus far, there are many platforms of microarray chip being used such as the bead array and the solid-phase array, each of which may cover different numbers of genes [11-12]. Thus, the use of pathway-based gene expression analysis is limited for cross-platform study because of a missing data issue. In machine learning, the missing data problem has been well studied. There are three main approaches to deal with the missing data problem, which are deletion, imputation and using as it is [13]. The missing data generally can be categorized into several types includes missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). The missing genes in cross-platform pathway-based microarray study would be categorized as MAR because the “missingness” is conditionally dependent on some variables [14]. The best way, to treat MAR data is to have some models describing the “missingness” and then impute those missing data [13]. K-Nearest Neighbor (KNN) imputation is one of the simplest methods, which will estimate the missing data by inferring from the k-nearest samples, which contain the value of those missing features [15]. However, KNN imputation was found to induce noise or “dirty” data to the dataset. In 2003, Oba et al. proposed another Bayesian-based imputation called Bayesian Principal component analysis (BPCA) [16]. By using a probabilistic model with Bayes inference framework, this imputation method is found to be more effective and accurate when compared to other existing methods [17-19]. This study applies the BPCA missing data imputation method with two pathway activity transformation methods, which are NCFS-i and AFS for cross-platform microarray data classification. The performance of the method is evaluated by using four actual cancer microarray datasets.

## 2 Methodology

### 2.1 Microarray Datasets

The microarray datasets used in this study are obtained from Gene Expression Omnibus (GEO) Datasets [20]. The general information of four microarray datasets from two different cancers; Lung cancer and Breast cancer is showed in Table 1. The classification scenario of Breast datasets is to distinguish breast cancer patients caused by different genetic factors; Sporadic, BRCA1 mutation and BRCA2 mutation, while Lung datasets are to classify healthy people and patients with lung cancer. In the analysis, two microarray datasets of the same cancers type are analyzed cross-platform.

**Table 1.** Details of microarray datasets used in this study

| Name | Accession | Platform | # of Genes | Class distribution (samples) |
|---|---|---|---|---|
| Breast1 | GSE40115 | GPL15931 | 22,170 | Sporadic: 104<br>BRCA1: 23<br>BRCA2: 20 |
| Breast2 | GSE40115 | GPL15932 | 19,190 | Sporadic: 56<br>BRCA1: 20<br>BRCA2: 16 |
| Lung1 | GSE4115 | GPL96 | 12,753 | Normal: 90<br>Tumor: 97* |
| Lung2 | GSE18842 | GPL570 | 20,318 | Normal: 45<br>Tumor: 46 |

* Lung1 dataset, 5 samples diagnosed as suspect lung cancer, are removed due to the quality control.

Analyzing two microarray datasets from different platforms will have some missing genes. The level of missing data in this cross-platform study varied from ~4% to ~40% as shown in Fig. 1. In this multivariate study, the standardization of all features needs to be done before analyzing the data, in order to keep variables at a similar measurement scale. So, all gene expression levels will be standardized cross samples using the following formula.

$$z_{ij} = \frac{x_{ij}-\mu_j}{\sigma_j} \qquad (1)$$

where, $Z_{ij}$ is z-transformed expression level of gene $j$, sample $i$,

$X_{ij}$ is expression level of gene $j$, sample $i$, $\mu_j$ is average expression level of gene $j$ and $\sigma_j$ is standard deviation of expression level of gene $j$.

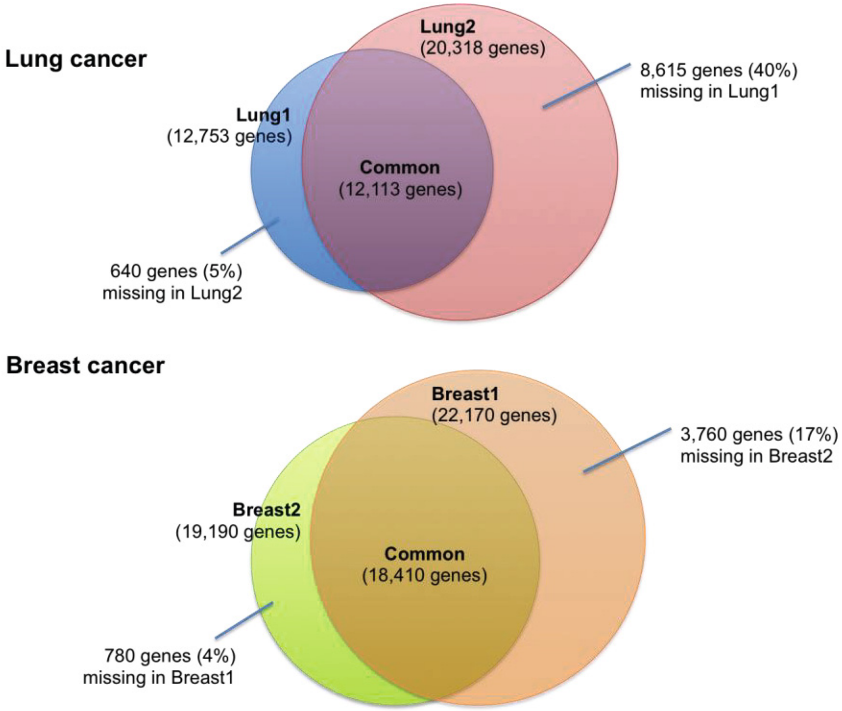


**Fig. 1.** Venn diagram display number of genes missing when doing cross-platform analysis

### 2.2 Pathway Data

The pathway data used is obtained from a public source called Molecular Signatures Database (MSigDB) [21]. This study uses the curated canonical pathways (c2.cp.v4.0), which is the curated pathways set from online databases includes KEGG, Biocarta, Pathway interaction database, Reactome, SigmaAldrich, Signaling Gateway, Signal Transduction KE and SuperArray. This curated canonical pathways set contains 1320 pathways.

### 2.3 Bayesian Principal Component Analysis (BPCA)

In classification, the model is trained by one dataset called training set and validated using another independent dataset of the same cancer called test set in order to evaluate the performance and robustness of the model. However, the missing data issue would occur when doing cross-platform validation, as each platform covers different number of genes. This study applies BPCA approach to reconstruct the missing data in test set using information from the training set. BPCA is a probabilistic missing data estimation approach, which composed of three main components, Principal Component regression, Bayesian estimation and expectation-maximization (EM)-like repetitive algorithm [16]. In this study, BPCA was implemented using pcaMethods library in R package [22].

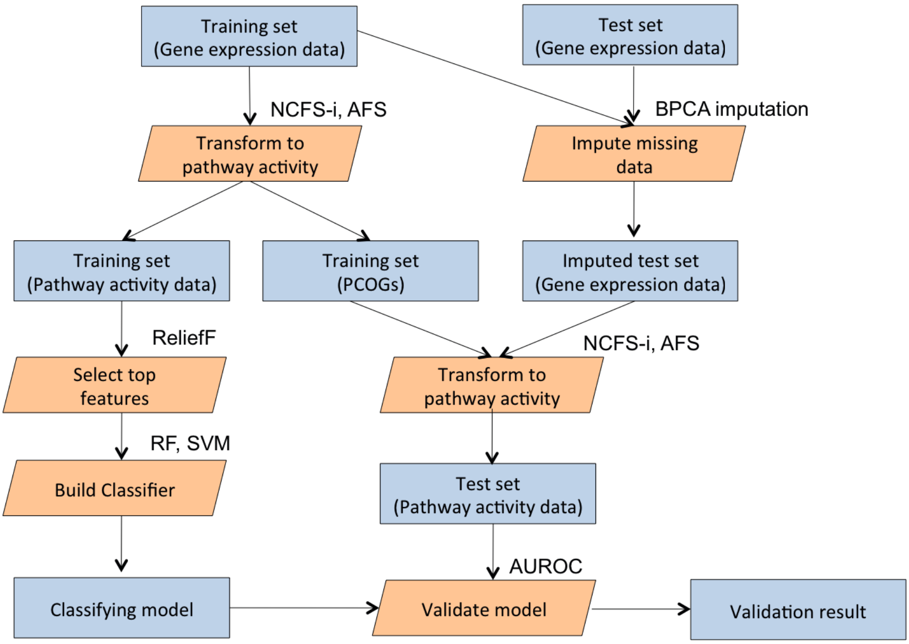


**Fig. 2.** Overall processes of cross platform classification

### 2.4 Pathway Activity Transformation

The integrative analysis of biological data is a powerful approach to study complex diseases. Transforming gene expression level into pathway activity using pathway information is a popular technique to integrate more useful information with microarray data. This work applies two pathway activity transformation methods, which are NCFS-i and AFS. NCFS-i was applied for lung cancer datasets as it was developed to working with binary label problem only [9], while AFS was applied for breast cancer datasets, which is multi-label problem. The main concept of NCFS-i and AFS is to identify the Phenotype Correlated Genes (PCOGs) of each pathway and transform gene expression levels of PCOGs into pathway activity.

### 2.5 Feature selection and Classification

In machine learning, feature selection has been found to be one of the most important steps in order to have higher performance classifying model. Including, the reduction of feature space would provide the faster and more cost-effective classifier [23]. This study uses Relief-F feature filtering and ranker selection approach, which is widely used in many fields, includes bioinformatics [24-26]. In this study, the feature selection will be performed on training data to select different number of top features, which are varied from top 1 to 50 features. Then, those top feature sets are used to build the classifying models. For the classifiers, two well-known classifiers used in this work are Random Forest (RF) and Support Vector Machine (SVM). The classifying models are then validated by another independent test set. The performance of the classification is assessed by calculating Area under ROC curve (AUROC). The feature selection and classification were implemented using java library of WEKA [27]. Overall processes of cross-platform classification are shown in Fig. 2.

# 3 Results

This work applied the Bayesian-based missing data imputation (BPCA) for pathway-based cross-platform microarray data classification. Here we compared the cross-platform classification performance between with and without missing data imputation using public microarray datasets of two cancers, which are Lung cancer and Breast cancer.

## 3.1 Imputation Results

Before classification step, the BPCA missing data estimation of missing genes in test set needs to be done by modeling the expression level of those missing genes in training set. In this study, the BPCA-imputed test set and non-imputed test set are compared in term of discriminative power of its own data. For breast cancer datasets, which are multi-label problem, the discriminative score of each pathway is calculated as F-value derived from ANOVA. The discriminative score for Lung cancer datasets, which are binary-label problems, is calculated as absolute t-score derived from Student's *t*-test. Table 2 shows the comparison of discriminative powers of BPCA-imputed and non-imputed test sets. The discriminative powers of the BPCA-imputed test sets of the Lung2, Breast1 and Breast2 datasets are significantly higher than the non-imputed test set except for the Lung1 dataset, in which the level of missing data is much larger than the others (~40% missing data).

**Table 2.** Discriminative scores comparing BPCA-imputed and non-imputed test sets

| | Average ± Standard deviation | | *p*-value of One way Paired *t*-test |
|---|---|---|---|
| | BPCA | Non-Imputed | |
| Breast1 | ***9.35±7.37** | 9.22±7.23 | 1.0E-06 |
| Breast2 | ***13.8±11.7** | 13.6±11.5 | 0.02 |
| Lung1 | 1.24±0.95 | ***1.28±0.98** | 2.0E-04 |
| Lung2 | ***4.21±3.05** | 4.11±3.02 | 5.0E-05 |

* Highest average discriminative score of each dataset

## 3.2 Cross-Platform Classification Results

After the BPCA imputation of missing data is done, training data is then transformed to pathway activity using either NCFS-i or AFS, for Lung and Breast cancer datasets, respectively. Then the PCOGs information identified in training data is used to transform independent test set of the same cancer to pathway activity. Relief-F feature selection are then performed in training data to select top 1%, 5%, 10%, 20%, and 50% top pathways for further classifier training. Here, two popular classification algorithms; RF and SVM were applied. The trained models were validated using cross-platform independent BPCA-imputed test set and non-imputed test set in order to evaluate their performance and robustness. For example, Breast1-Breast2 represents the cross-platform validation that use Breast1 dataset to train the model and validate the model by using Breast2 dataset. The cross-platform validation result comparing the classification performance on BPCA-imputed and Non-imputed test sets is shown in Fig. 3. The difference of the classification performances was statistically assessed by using one-way paired *t*-test. The results show that there is no difference observed in the Breast2-Breast1 validation because the missing data of this test set is very low (4%) and the PCOGs of those top selected features do not have these missing genes as their members. In the Breast1-Breast2 and the Lung1-Lung2 validations, it was found that using BPCA-imputed test set yields significantly higher AUCs than the non-imputed test set on both RF (*p*-value = 1.9E-05 and 1.5E-13) and SVM (*p*-value = 8.6E-12 and 2.1E-06). However, for the test set with high volume of missing data (Lung2-Lung1 validation), the classification on non-imputed test set shows significantly better performance in the RF (*p*-value = 4.9E-19). Nonetheless, the overall classification performance is rather poor (average AUC = ~0.53) due to the high number of missing data.

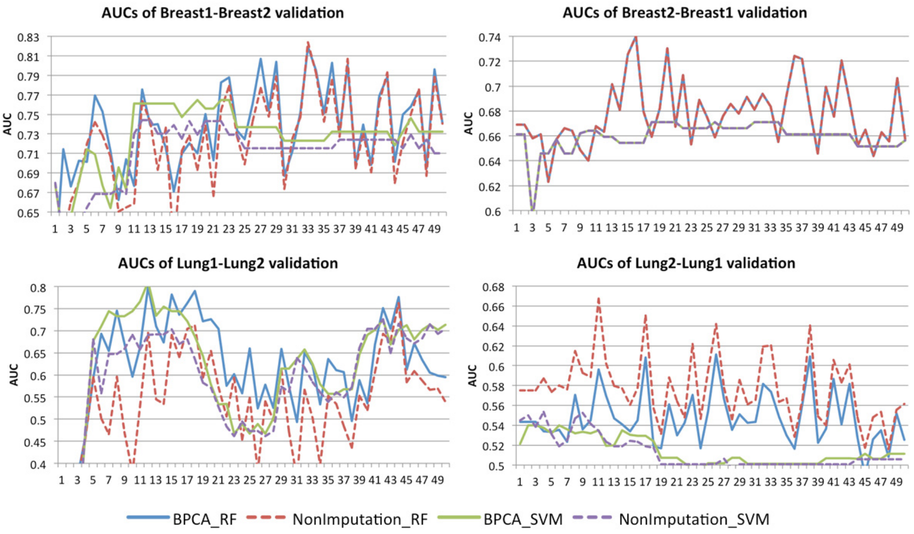


**Fig. 3.** Comparison of cross-platform validation on BPCA-imputed and non-imputed test sets using SVM and RF as classifiers. Dashed lines represent the classification performance on non-imputed test set, while solid lines represent the BPCA-imputed test set validation results.

## 4 Discussions

The pathway activity transformation methods (NCFS-i and AFS) have been claimed to be able to produce a robust classifier for complex disease classification [9-10]. The cross-platform validation results show that these methods are also robust to the missing data issue as well, which can be used for classification purpose without missing data imputation with acceptable accuracy. However, by reconstructing the missing data by the BPCA method, it produces a slightly better discriminative power test set and results in the improvement of cross-platform validation performance in most cases. In this study, we have not considered any batch effect correction in order to clearly illustrate the effect of missing data when doing cross-platform analysis only. So, in the practical study of cross-platform analysis, batch effect correction should be done before subsequent analysis to remove the confounding factors to improve the quality of cross-platform study. For the test set with high missing data level (Lung1), even the classification on the non-imputed test set is better than the BPCA-imputed test set, the overall classification performance is unacceptable. So, the integrative study of these two datasets (Lung1 and Lung2) is not recommended. Considering the missing data of the Breast1-Breast2 and the Lung1-Lung2 validations, EXOSC8 and METTL2B are found missing in PCOGs of the top 10 selected pathways of Breast2 and Lung2, respectively. The function of EXOSC8 is tentatively reported to be related to the sporadic-breast cancer [28] while there is no report on the gene-disease relationship of METTL2B.

## 5 Conclusions

This study applies the Bayesian-based missing data imputation method (BPCA) in pathway-based cross-platform classification of microarray data. Two pathway activity transformation methods (NCFS-i and AFS) were used in order to extract more information in microarray data by transforming gene expression level to pathway activity using pathway information from MSigDB. The cross-platform validations of the test sets with and without BPCA imputation were compared using Relief-F as the feature selection method and RF and SVM as the classifiers. The results show that there is an improvement when using the BPCA method to estimate missing data. So to improve the pathway-based cross-platform microarray data analysis, the BPCA missing data imputation should be applied.

**Acknowledgments.** Worrawat Engchuan would like to thank National Center for Genetic Engineering and Biotechnology, Thailand (BIOTEC) for providing a scholarship for his current PhD study. Also, this work is supported by the research subsidy fund of fiscal year 2014 from King Mongkut's University of Technology Thonburi.

## References

1. Kerr, M.K., Martin, M., Churchill, G.A.: Analysis of variance for gene expression microarray data. Journal of Computational Biology 7(6), 819–837 (2000)
2. Quackenbush, J.: Computational analysis of microarray data. Nature Reviews Genetics 2(6), 418–427 (2001)
3. Rhodes, D.R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A., Chinnaiyan, A.M.: Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. Proceedings of the National Academy of Sciences of the United States of America 101(25), 9309–9314 (2004)
4. Lee, E., Chuang, H.Y., Kim, J.W., Ideker, T., Lee, D.: Inferring pathway activity toward precise disease classification. PLoS Computational Biology 4(11), e1000217 (2008)
5. Sootanan, P., Prom-on, S., Meechai, A., Chan, J.H.: Pathway-based microarray analysis for robust disease classification. Neural Computing and Applications 21(4), 649–660 (2012)
6. Engchuan, W., Chan, J.H.: Pathway activity transformation for multi-class classification of lung cancer datasets. Neurocomputing (in press, 2014)
7. Choi, J.K., Yu, U., Kim, S., Yoo, O.J.: Combining multiple microarray studies and modeling inter study variation. Bioinformatics 19, i84–i90 (2003)
8. Xu, L., Tan, A.C., Naiman, D.Q., Geman, D., Winslow, R.L.: Robust prostate cancer marker genes emerge from direct integration of inter-study microarray data. Bioinformatics 21(20), 3905–3911 (2005)
9. Benito, M., Parker, J., Du, Q., Wu, J., Xiang, D., Perou, C.M., Marron, J.S.: Adjustment of systematic microarray data biases. Bioinformatics 20(1), 105–114 (2004)
10. Chen, C., Grennan, K., Badner, J., Zhang, D., Gershon, E., Jin, L., Liu, C.: Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. PloS One 6(2), e17238 (2011)
11. Tan, P.K., Downey, T.J., Spitznagel Jr., E.L., Xu, P., Fu, D., Dimitrov, D.S., Cam, M.C.: Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Research 31(19), 5676–5684 (2003)
12. Irizarry, R.A., Warren, D., Spencer, F., Kim, I.F., Biswal, S., Frank, B.C., Yu, W.: Multiple-laboratory comparison of microarray platforms. Nature Methods 2(5), 345–350 (2005)
13. Howell, D.C.: The treatment of missing data. In: The Sage Handbook of Social Science Methodology, pp. 208–224 (2007)
14. Donders, A.R.T., van der Heijden, G.J., Stijnen, T., Moons, K.G.: Review: a gentle introduction to imputation of missing values. Journal of Clinical Epidemiology 59(10), 1087–1091 (2006)
15. Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Bostein, D., Altman, R.B.: Missing value estimation methods for DNA microarrays. Bioinformatics 17(6), 520–525 (2001)
16. Oba, S., Sato, M.A., Takemasa, I., Monden, M., Matsubara, K.I., Ishii, S.: A Bayesian missing value estimation method for gene expression profile data. Bioinformatics 19(16), 2088–2096 (2003)
17. Brock, G.N., Shaffer, J.R., Blakesley, R.E., Lotz, M.J., Tseng, G.C.: Which missing value imputation method to use in expression profiles: a comparative study and two selection schemes. BMC Bioinformatics 9(1), 12 (2008)

18. Pedreschi, R., Hertog, M.L., Carpentier, S.C., Lammertyn, J., Robben, J., Noben, J.P., Panis, B., Swennen, R., Nicolaï, B.M.: Treatment of missing values for multivariate statistical analysis of gel based proteomics data. Proteomics 8(7), 1371–1383 (2008)
19. Liew, A.W.C., Law, N.F., Yan, H.: Missing value imputation for gene expression data: computational techniques to recover missing data from available information. Briefings in Bioinformatics 12(5), 498–513 (2011)
20. Edgar, R., Domrachev, M., Lash, A.E.: Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research 30(1), 207–210 (2002)
21. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., Mesirov, J.P.: Molecular signatures database (MSigDB) 3.0. Bioinformatics 27(12), 1739–1740 (2011)
22. Stacklies, W., Redestig, H., Scholz, M., Walther, D., Selbig, J.: pcaMethods—a bioconductor package providing PCA methods for incomplete data. Bioinformatics 23(9), 1164–1167 (2007)
23. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. The Journal of Machine Learning Research 3, 1157–1182 (2003)
24. Kononenko, I.: Estimating attributes: analysis and extensions of RELIEF. In: Bergadano, F., De Raedt, L. (eds.) ECML 1994. LNCS, vol. 784, pp. 171–182. Springer, Heidelberg (1994)
25. Robnik-Šikonja, M., Kononenko, I.: Theoretical and empirical analysis of Relief-F and RRelief-F. Machine Learning 53(1-2), 23–69 (2003)
26. Wang, Y., Makedon, F.: Application of Relief-F feature filtering algorithm to selecting informative genes for cancer classification using microarray data. In: Proceedings of the 2004 IEEE Computational Systems Bioinformatics Conference, CSB 2004, pp. 497–498 (2004)
27. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. ACM SIGKDD Explorations Newsletter 11(1), 10–18 (2009)
28. Pletscher-Frankild, S., Pallejà, A., Tsafou, K., Binder, J.X., Jensen, L.J.: DISEASES: Text mining and data integration of disease–gene associations, bioRxiv, 008425 (2014)

# Text Censoring System for Filtering Malicious Content Using Approximate String Matching and Bayesian Filtering

Khairil Imran Ghauth and Muhammad Shurazi Sukhur

Faculty of Computing and Informatics,
Multimedia University, Cyberjaya, Malaysia
khairil-imran@mmu.edu.my, mshurazi@gmail.com

**Abstract.** Information obtained nowadays often contains malicious contents. These malicious contents such as profane words have to be censored as they can influence the minds of the young ones and create hate among people. In censoring the profane words, this paper introduces a hybrid text censoring method which is based on Bayesian Filtering and Approximate String Matching techniques. The Bayesian filtering technique is used to detect the malicious contents (profane words) while the Approximate String Matching technique is used to enhance the effectiveness of detecting profane words. In evaluating the performance of the proposed system, the evaluation metrics of Precision, Recall, F-measure and MAE were used. The results show that Bayesian filtering technique can be used to filter profane words.

**Keywords:** Bayesian filtering, Approximate String Matching, Text Censoring.

## 1 Introduction

With the emergence of new technologies in the Internet era, the capability to access to most of the information is within second. While good information is beneficial to human, the bad information (information pollution) is harmful. According to Wang et al (2012), the spreading of this information pollution can create problems to humans such as wasting time, physical health issues, and addiction problems.

Among information pollution available nowadays is the availability of profane words in the Internet such as in websites, blogs, social media and chat rooms. These profane words are considered as malicious content that has to be censored as they can influence the minds and behaviours of people especially children.

In this paper, malicious contents are defined as contents that contain profane words in English language. One way to filter out the malicious content is by using information filtering techniques. Information filtering techniques are used to extract useful information and block unwanted information according to certain rules [1]. The information filtering concept closely related to text censoring system because the information filtering technique deals primarily with textual information [2]. In general, information filtering is used for two purposes; first is to search for suitable

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_15

information and its related information and the results are displayed to the users. Second, is to find unwanted information and remove it from being displayed to the user. The first is more widely used and recommender system is one of the techniques that relates to it. Meanwhile, the text censoring system is more related to the second purpose of blocking information from being displayed to the users. In fact, the term 'filtering' is more suitable to be associated with removal of data as suggested by Belkin and Croft (1992) where 'filtering is often meant to imply the removal of data from an incoming stream, rather than finding data in that stream'.

The objective of this study is to propose a new text censoring system that is based on Bayesian Filtering and Approximate String Matching. To access the benefit of our approach, two other text censoring systems are chosen. They are List-Based text censoring system and Inverse Chi Square text censoring system. Empirical study shows that our proposed Bayesian filtering text censoring system is better in censoring profane words in terms of accuracy and relevance. Therefore the hybrid method described in this paper can be used as one of the method in text censoring system.

The rest of the paper is organized as follows. Section 2 describes existing works related to profanity filtering. Then we present the proposed Bayesian Filtering text censoring system in detail in Section 3. In Section 4, we discuss the experiments conducted using our own dataset and compare the results obtained by systems censoring with benchmark human censoring. Finally in Section 5, we summarize our conclusion.

## 2 Related Work

In studies conducted regarding the censorship on bad information (example: profanity words), the researchers were using few techniques such as Bayesian Network, Content Based, K Nearest Neighbour (kNN), R*Tree Based and Rules Reduction.

Qinshan et al (2011) in their paper, 'The Research of Information Filtering Technology Based on Bayesian Network' use Bayesian Network method to ensure filtering accuracy improves. The improvement is due to an adaptive function; in which it can learn new bad information and subsequently be able to censor the latest bad information.

Polpinij et al (2008) use Content Based filtering to censor malicious contents as evident in their work, 'A Web Pornography Patrol System by Content-based Analysis: In Particular Text and Image'. Here, they experiment by using Support Vector Machine (SVM) algorithm.

Du and Yi (2012) use K-Nearest Neighbour (k-NN) in their work, 'A New KNN Categorization Algorithm for Harmful Information Filtering'. In their work, they tried to solve a problem of the availability of small samples by constructing virtual samples based on the small samples available using up-sampling method. In this way, the classification effect of k-NN algorithm is also improved. As a result, the harmful text filtering performance is improved greatly.

Another method used to filter profanities is by using R*Tree based technique. R*Tree is a variant of R-Tree. This method is used by Yoon, Park and Cho (2010) in 'A Smart Filtering System for newly Coined Profanities by Using Approximate String Alignment'. In this paper, the authors suggest a smart filtering system for profanities by measuring the similarities between the profanities and the given words. In the system, the profanity filtering is done by approximate string matching using candidate finding algorithm and using matching matrix together with global alignment.

## 3 Bayesian Filtering and Approximate String Matching Text Censoring System

The framework of the proposed Text Censoring System consists of a few phases which are Tokenization, Pre-censoring and Censoring. In the first phase, the whole text from the uploaded document is divided into individual word, known as ‘token’. During tokenization, the words are separated using a naïve approach in which whitespaces are used to delimit between each word. The words are then stored in the Word database together with other information such as the frequency of the word appear in censored documents (*frequency*$_{wc}$), the frequency of the word appear in uncensored documents represented as (*frequency*$_{wu}$), and the probability the word is profane ($P_{wp}$).

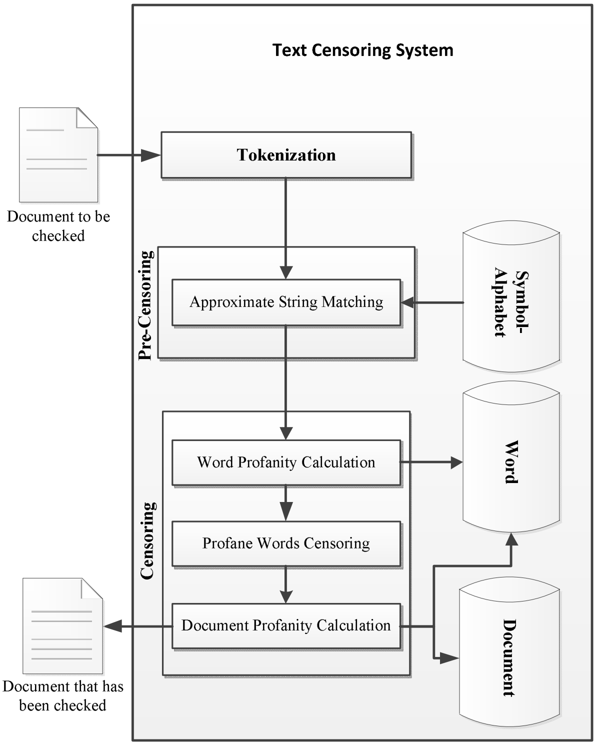


**Fig. 1.** The Proposed Framework of Text Censoring System

Next is the Pre-censoring phase. In this phase, the system will check for words that contain symbols that replace alphabetic characters. For example, '$oftw@re' can be understood as 'software' even though the word contains a combination of symbols and alphabetic characters. As such, profane words can also be constructed from the combination of letters and symbols. The Pre-censoring phase is necessary to compare words from the uploaded document with words in the word database without ignoring words that might have been mixed up with symbols. Therefore, the main function during pre-censoring is to convert the words back to their original form for better matching. In this regard, the Approximate String Matching (ASM) technique is employed to convert the symbol into a possible character. For each pair of symbol and alphabetic character, there will be a matching value that indicates how similar the characters are. In this paper, the determination of the matching values is done by performing an experiment using one of the approximate string comparison algorithms known as Jaro-Winkler Distance algorithm. The Jaro-Winkler Distance algorithm is developed based on experience of data matching conducted over many years at the US Census Bureau [6]. The Jaro-Winkler Distance algorithm stresses that the order of occurrence is essential to determine the similarity and because of this, the results obtained from Jaro-Winkler Distance algorithm is more accurate than other string comparison techniques such as Cosine Similarity [7]. Table 1 shows the example of matching values between symbol and alphabetic character.

**Table 1.** Matching Value for Each Pairing

| Symbol | Alphabetic Character | Matching Value |
|---|---|---|
| @ | a | 0.89 |
| ! | i | 0.90 |
| $ | s | 0.92 |

Based on the matching value, we calculate the similarity between the word that contains symbols ($word_{symb}$) and the suggested word ($word_{suggested}$). Comparison is done by comparing each character of the word between $word_{symb}$ and $word_{suggested}$. For an exact matching character, the score is 1.00. If the character is not an exact match, we will then follow the matching value of the pair as shown in Table 1. A comparison score between the word that contains symbol and the suggested word is then calculated by adding up the score of the characters. Then the comparison score will be compared with overall score to get a comparison percentage. Higher comparison percentage indicates higher similarity between the two words.

In order to allow $word_{symb}$ to be replaced with $word_{suggested}$, we need to determine the minimum comparison percentage. This is to prevent $word_{symb}$ from being replaced by $word_{suggested}$ even though the comparison percentage is small. In this research, we set the minimum similarity percentage to be equal to the average of the similarity percentage of all the pairs in Table 1 which is 90%. In other words, the replacement from $word_{symb}$ to $word_{suggested}$ will only be allowed if comparison percentage exceeds 90%.

The final phase of the proposed framework is the Censoring phase. Here we use the Bayesian filtering technique to censor the document. In the proposed framework, there are some values that need to be pre-determined before censoring exercise is made. It is proposed that we use the values suggested in the spam filtering for all the pre-determined values. In his article "A Plan for Spam", Graham (2002) in discussing the usage of Bayesian filtering to filter spam emails, suggested the following:

1. For initial probability value to be given to a word that never exists in word database before, the author suggested the value of 0.4 on a basis that if we never encountered a word before, it is probably fairly innocent.
2. For a value $N$ representing a top-N words that have the highest profanity probability value in a document, the author suggested 15 as the $N$ value for top-n words to be used in calculating profanity value of a document.
3. We propose a value of 0.9 for threshold probability value for a word and a document to be considered as profane. This is based on a statement by Graham (2002) in which he says "I treat mail as spam if the algorithm above gives it a probability of more than .9 of being spam".

As mentioned earlier, the value of 0.4 will be assigned to a word never exists in word database. However we propose that before a new word is assigned to the value of 0.4 we first check whether there is singular or plural form of the word in the database. If the singular or the plural form of the word exists in the database, the new word will be assigned the same $P_{wp}$ of its singular or the plural form of the word together with the values of $frequency_{wc}$ and $frequency_{wu}$. This is based on the fact that if a word is profane, its singular or plural from must also be profane.

After each word in the document has been separated into token, the Bayesian filter will compare each token with words in word database. Words in the database will either created or updated whenever this Bayesian filter learn a new document.

Whenever a token is compared in the database, Bayesian filtering will make four calculations as follows:

1. The frequency of this token found in censored document ($frequency_{wc}$).
2. The frequency of this token found in uncensored document ($frequency_{wu}$).
3. The number of documents considered as 'censored' that the filter has been trained.
4. The number of documents considered as 'not censored' that the filter has been trained.

In this exercise, a document is considered as 'censored' if the document contains a high number of profane word and 'uncensored' if the document does not contain any profane word or low profane words. The classification of 'censored' or 'uncensored' document will not have any impact on the document itself. It is only be used for calculation of profanity value of a word in future filtering. The following formula is used to calculate the probability of word [10].

$$P_{wp} = P_c/( P_u + P_c ) \tag{1}$$

where $P_c$ is the censored probability and $P_u$ is the uncensored probability. If the end value of $P_{wp}$ is 0.0 (non-profane word) or 1.0 (profane word), 0.01 and 0.99 will be assigned respectively. This is to avoid mathematical complication if 0.0 and 1.0 are used as suggested by Anderson (2006).

After the calculation of the $P_{wp}$ for each word, we now calculate the document profanity value. General Bayes' formula is represented as follows:

$$P(E|F) = \frac{P(F|E)P(E)}{\sum P(F|E)P(E)} \tag{2}$$

According to Process Software in 'Bayesian Filtering Example', if $a$ and $b$ are 2 evidentiary events, its probability $P_{ab}$ can be equalled to:

$$P_{ab} = \frac{ab}{ab+(1-a)(1-b)} \tag{3}$$

As for 3 evidentiary events $a$, $b$ and $c$; the formula expands where the probability $P_{abc}$ now is equal to:

$$P_{abc} = \frac{abc}{abc+(1-a)(1-b)(1-c)} \tag{4}$$

Therefore, any number of evidentiary events can be included by expanding the Bayes formula. Based on this fact, to determine whether the document will be classified as a 'censored document', $n$ number of tokens that appear in the document and have the highest profanity percentage are chosen for next calculation. The Bayes formula for the next calculation is as follows:

$$P_{dp} = \frac{\prod_{i=1}^{n} P_{wp}}{\prod_{i=1}^{n} P_{wp}+\prod_{i=1}^{n}(1- P_{wp})} \tag{5}$$

where $P_{dp}$ is the probability of the document profanity value. Based on the value of $P_{dp}$, we can determine whether the document is considered profane or not by comparing it with the threshold value of acceptable profanity value of a document.

# 4 Performance Evaluation

The main objective of this empirical study is to assess whether Bayesian filtering technique can be used to filter profane words as effective as it is used in spam filtering. Section 4.1 explains the experimental setup. Afterwards, Section 4.2 presents the experimental results.

## 4.1 Experimental Setup

In this section, we will look at training dataset, evaluation dataset, comparison of the proposed system with other text censoring systems and evaluation metrics used.

### Training Dataset

The proposed Text Censoring System is based on statistical feature of Bayesian filtering. Therefore it needs to be trained to allow the system to recognize the words that can be considered as profane and non-profane based on the accumulated probability of the words evaluated. The aim of the training is to set the profanity probability value of profane words to high while keeping the profanity probability value of non-profane word to low.

The training sets are divided into two categories. First, the dataset for learning profane words category. These words were taken from noswearing.com dictionary of swear (profane) words amounting to 349 words. These words were then grouped into 5 documents and each document was duplicated 60 times for a total of 300 documents that contain profane words. The documents were then executed in the text censoring system to allow the system to calculate and store the profanity probability value for each word.

Then, there is a dataset for learning non-profane words was taken from a list of most frequently used words in English language available at the website of Utah State Office of Education (http://www.schools.utah.gov). 349 words from the list were taken and these words were similarly grouped into 5 documents and each document was duplicated 60 times to create a total of 300 documents that contain non-profane words. These documents were also executed in the text censoring system in order for the system to calculate and store profanity value for each word.

The profanity probability value of profane words was increasing and the profanity probability value for non-profane words was decreasing during training session. This shows that the proposed system has been successfully trained to reach training target.

### Evaluation Dataset

We created our own dataset for evaluation since there are no publicly available data sets to test the profanity filtering systems. This dataset is taken from the Internet in various categories such as Internet forum, novels, movie quotes, song lyrics and others such as blog posting, Facebook status, online news, Twitter status and YouTube comment. There are 25 documents created. There is also a mixture of long and short documents. The longest document has 387 words while the shortest document has 20 words. The average words in the dataset documents are 167.4 words.

### Comparison with Other Techniques

After we have trained the system, we use the evaluation dataset to test the proposed technique in three text censoring systems - Bayesian filtering, List-Based and Inverse Chi Square.

The List-Based text censoring system depends on the list of profane words to evaluate and detect profane words in documents. The effectiveness of the censoring depends on the completeness of the list of profane words supplied to the system. Here, the list of profane words is taken from the profane words dictionary from noswearing.com; the same 349 profane words trained in Bayesian filtering text censoring system.

Inverse ChiSquare text censoring system, on the other hand, uses Inverse ChiSquare technique which – along with Bayesian filtering technique – is a popular statistical technique used in spam filtering. In their work, AlKhahtani, Gardner-Stephen and Goodwin (2011) classify Bayesian filtering and Inverse ChiSquare in a same category of Content Based- Machine Learning-Statistical Methods of Spam filters. Therefore, before the Inverse ChiSquare system is evaluated, it is trained using the same training dataset for Bayesian filtering text censoring system.

To ensure fairness in comparing the different techniques, all of the systems are equipped with Approximate String Matching (ASM) feature to detect words with symbols and replace them with suitable replacement words before the evaluation of profanity using their respective techniques.

### Evaluation Metrics Used

In comparing the performance of the systems and human censoring result, we use a few evaluation metrics which are Precision, Recall, F-Measure and Mean Absolute Error (MAE). The results of the evaluation will be discussed in the next section.

## 4.2 Experimental Results

In this experiment, the dataset is also examined by human censoring which acts as the ultimate censoring mechanism. Human censoring also uses the same evaluation dataset. This is to determine whether the evaluations made by systems are closely performing to the level of human evaluation. The humans involved come from various backgrounds - male and female, Malaysians and African and ages range from 20s to 40s. The results of the experiments conducted are as follows:

### Precision and Recall

Based on the results obtained from the experiments involving the three text censoring systems, the Precision and Recall values for the proposed Bayesian Filtering Text Censoring System (BF), the Inverse Chi-Square Text Censoring System (ICS) and the List-Based Text Censoring System (LB) are as follow:

**Table 2.** Precision and Recall Values for All Three Text Censoring Systems

| | **BF** | **ICS** | **LB** |
|---|---|---|---|
| **Precision** | 0.88117 | 0.89660 | 0.89722 |
| **Recall** | 0.94666 | 0.88272 | 0.89457 |

From the results, it shows that Precision value for BF is 2% lower than the Precision values of ICS and LB. However the Recall value for BF is the highest compared to the other two systems which are more than 6% higher than ICS and LB Recall values. The reason why the Precision value for BF is lower than the other techniques is that BF censoring detected some uncommon profane words that were not detected by human participants in their censoring. For example, profane word such as ‘kikes’ which is an offensive term for Jews was not detected by humans involved in the experiment but highlighted by BF censoring. Since English is not their native language, they may not aware that the word is a profane word. As a result, BF censoring received false positive for uncommon profane words which eventually affected its Precision value. On the other hand, ICS and LB are unable to detect some uncommon profane word and that saves them from receiving false positive during evaluation. It is believed that if the human censoring participants' native language is English, the results may be different in favor of the BF.

**F-measure**

After getting the Precision and Recall values from the evaluated systems, we can now calculate the F-measure percentage values for all three systems by using the values available in previous table. The F-measure values obtained are as follow:

**Table 3.** F-measure Values for All Three Text Censoring Systems

| | **BF** | **ICS** | **LB** |
|---|---|---|---|
| **F-measure** | 0.90700 | 0.88113 | 0.88798 |

The result shows that BF has the highest F-measure value. It is about 2% higher that the F-measure values for ICS and LB. The major contribution factor for high F-measure value for BF is the high Recall value received by the proposed system compare to the other two systems.

**MAE**

Based on number of document evaluated by human and systems; N = 25, the values of MAE for all three text censoring systems are as follow:

**Table 4.** MAE Values for All Three Text Censoring Systems

| | **BF** | **ICS** | **LB** |
|---|---|---|---|
| **MAE** | 0.84000 | 0.96000 | 0.92000 |

From the MAE values shown, the BF received the lowest MAE value at 0.84 compared to the other systems which achieved MAE values greater than 0.90. Since the lower MAE value indicates the better value, the BF is in this experiment achieved the closest result compared to human censoring.

## 5 Conclusion

The implemented proposed work of Bayesian filtering text censoring system is evaluated against the other text censoring systems that use Inverse Chi Square technique and the List-based technique. In measuring the performance of the proposed system with the other systems, we use the metrics of Precision and Recall, followed by F-Measure and finally the Mean Absolute Error (MAE).

In the experiment, the proposed Bayesian filtering text censoring system has performed well when it achieved the highest F-measure value compared to the values achieved by the other two text censoring systems. Moreover, the proposed system has outperformed the other systems in Recall and MAE values. This shows that the

Bayesian filtering text censoring system is a system that able to carry out the most relevant censoring with least amount of errors compared to the other two text censoring systems. In the future, we plan to involve native English speakers in human experiment to censor profane words as it may affect the result of human censoring.

**Acknowledgements.** This research was supported by the Ministry of Education Malaysia under the EXPLORATORY RESEARCH GRANT SCHEME PROJECTS (ERGS/1/2013/ICT01/MMU/02/01).

## References

1. Wang, N., Jiang, K., Meier, R., Zeng, H.: Information Filtering against Information Pollution and Crime. In: 2012 International Conference on Computing, Measurement, Control and Sensor Network (CMCSN), pp. 45–47. IEEE (July 2012)
2. Belkin, N.J., Croft, W.B.: Information filtering and information retrieval: two sides of the same coin? Communications of the ACM 35(12), 29–38 (1992)
3. Polpinij, J., Sibunruang, C., Paungpronpitag, S., Chamchong, R., Chotthanom, A.: A web pornography patrol system by content-based analysis: In particular text and image. In: IEEE International Conference on Systems, Man and Cybernetics, SMC 2008, pp. 500–505. IEEE (October 2008)
4. Du, J., Yi, Z.A.: A New Knn Categorization Algorithm for Harmful Information Filtering. In: 2012 Fifth International Symposium on Computational Intelligence and Design (ISCID), vol. 1, pp. 489–492. IEEE (October 2012)
5. Yoon, T., Park, S.Y., Cho, H.G.: A smart filtering system for newly coined profanities by using approximate string alignment. In: 2010 IEEE 10th International Conference on Computer and Information Technology (CIT), pp. 643–650. IEEE (June 2010)
6. Christen, P.: Data matching: concepts and techniques for record linkage, entity resolution, and duplicate detection. Springer (2012)
7. Clayton, R.: Calculating Similarity (Part 2): Jaccard, Sørensen and Jaro-Winkler Similarity. Gettingcirrius.com (2011), `http://www.gettingcirrius.com/2011/01/calculating-similarity-part-2-jaccard.html` (retrieved February 21, 2014)
8. Graham, P.: A plan for spam. Available from World Wide Web (2002), `http://www.paulgraham.com/spam.html`
9. Anderson, D.: Statistical Spam Filtering, EECS595 (Fall 2006)
10. Bayesian Filtering Example (n.d), `http://www.process.com/precisemail/bayesian_example.htm` (retrieved)
11. Swear Word List (n.d). NoSwearing Website, `http://www.noswearing.com/dictionary` (retrieved)
12. The 1200 Most Frequently Used Words in the English Language (n.d), Utah State Office of Education (n.d), `http://www.schools.utah.gov/CURR/langartelem/Core-Standards/Resources.aspx` (retrieved April 3, 2014)
13. Alkahtani, H.S., Gardner-Stephen, P.A.U.L., Goodwin, R.: A taxonomy of email SPAM filters. In: Proc. the 12th International Arab Conference on Information Technology (ACIT), pp. 351–356 (2011)

# Application of Optimum Binning Technique in Data Mining Approaches to Predict Students' Final Grade in a Course

Syed Tanveer Jishan, Raisul Islam Rashu, Ashfaq Mahmood,
Fahad Billah, and Rashedur Mohammad Rahman

Department of Electrical and Computer Engineering, North South University,
Plot-15, Block-B, Bashundhara, Dhaka, Bangladesh
tanveer_syed@live.com, {rashu.raisul1,ashfaq.bd92}@gmail.com,
fahadbillah@yahoo.com, rashedur@northsouth.edu

**Abstract.** In Bangladesh there is a continuous rise in demand for higher education in last decade; therefore, the need for improving the education system is imminent. Data-mining techniques could be explored on educational settings to extract useful information which could be helpful for the students as well for the instructors. In this research, we present and analyze techniques that predict students' final outcome (with respect to grade) for a particular course. We validate our method by conducting experiments on data that are related to grade for courses in North South University, one of the leading universities in higher education in Bangladesh. Our preliminary finding is encouraging. We further improve and extend our ideas through discretization of the continuous attributes by equal width binning and error minimization techniques. Experimental results demonstrate that improved technique through discretization outperforms the techniques with other forms such as probability estimation by almost 7-10 %.

**Keywords:** Educational Data Mining(EDM), classification, Naive Bayes, Decision Tree, Neural Network, Discretization, Equal Width Binning.

## 1 Introduction

Educational Data Mining is an interdisciplinary research area that focuses on the use of data mining in the educational field. Educational data can be from different sources, but generally from academic institutions. However, online learning systems are now the new environment for generating educational data which can be used to analyze and extract useful information [16].The goal of this paper is to predict the students' performance using the CGPA, Quiz, Laboratory, Midterm and Attendance marks so that the students can be alerted before the final examination regarding their grade outcome. This will not only help the students, but also the instructor who will get an insight on how the students are doing in the course.

For this work we acquired data set on a particular course from North South University. After acquiring the data we preprocessed it and then applied several classification

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems*,
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_16

algorithms such as Naïve Bayes, Decision Tree and Neural Network. We also discretized the continuous attributes using optimal equal width binning as proposed by Kaya[10] and then compared the model accuracy with the model having continuous attributes' class estimated using the probability density function. After all the models are built we compared their accuracy, precision and recall of the class labels for those models. ROC Curves for each of the models are generated and area under the curve (AUC) is also calculated and compared.

## 2 Related Works

Educational Data Mining is a vast domain which consists of different applications. Using data mining techniques it is possible to build course planning system, detecting what type of learner a student is, making group of similar types of students, predicting the performance of the students as well as helping instructors to get insight on how to commence the classes [16]. Pal and Pal [1] conducted studies at the VBS Purvanchal University, Jaunpur, India and used classification algorithms to identify the students who need special advising or counseling from the teachers. Ayers et al. [2] used several clustering algorithms such as hierarchical agglomerative clustering, K-means and model based clustering in order to understand skill levels of the students and group them based on their skill sets.

Bharadwaj and Pal [3] found that students' grade in the senior secondary exam, living location, medium of teaching, mother's qualification, family annual income, and student's family status are correlated strongly and help to predict how the students perform academically. In another study Bharadwaj and Pal [4] used students' previous semester marks, class test grade, seminar performance, assignment performance, general proficiency, attendance in class and lab work to predict the end of the semester marks. Chu et al. [5] described fuzzy logic based recommendation system for students, which will guide them what to learn before moving to the next step.

Hsia, Shie and Chen [6] conducted study at a university in Taiwan. They used Decision Tree, Link Analysis and Decision Forest to extract enrollee course preference and correlation between course categories. Rus et al. [7] showed how classification algorithms can be used to build a system which can detect student's mental model. Zhu et al. [8] explains how making a personalized learning recommendation system which will help the learner beforehand what he or she should learn before moving to the next step. Yadav et al. [9] used students' attendance, class test grade, seminar and assignment marks, lab works to predict students' performance at the end of the semester. They used the decision tree algorithms such as ID3, CART and C4.5 and made a comparative analysis. In their study, they achieved 52.08%, 56.25% and 45.83% accuracy for each of these classification techniques respectively.

Although there is handful of works on grade prediction models, our focus is to handle the continuous attributes effectively. Handling the continuous attributes effectively will result in better performance of the grade prediction models provided that most of the attributes in course mark sheets or data sets are continuous in nature.

# 3 Data Mining Process

## 3.1 Data Selection

We have collected data set of the course Numerical Analysis from North South University. The dataset contains record of five semesters, having 181 instances. Each instance is a student record. Originally the dataset had student ID, student name, five quiz marks, midterm marks, attendance, laboratory marks, final marks and final grade as attributes. We have selected the attribute which contains the percentage of marks obtained by the students in quizzes rather than taking all the quizzes into account. Final grade is considered as the class label.

## 3.2 Data Preparation

We have discarded the student ID and name attributes as these are not directly required for the data analysis purpose. Students' CGPA, which was not initially a part of the dataset, was retrieved and added as an attribute. Quiz, midterm and laboratory marks were normalized between 0 to 100 as there are fluctuation of weight assigned to each of these attributes in different semesters. All the attributes are described in details below.

**CGPA** - Cumulative Grade Point Average. CGPA of students who enrolled in the course are taken into consideration.

**Attendance Marks -** Normalization will have no effect on the attendance marks provided it had binary marking for every instances, that is either 0 or full marks. However normalization is done in order to keep consistency in terms of look with the other attributes.

**Quiz Marks -** Best four of the five quizzes are counted as per the course policy. The average of the best four quizzes are taken and then normalized between 0 to 100.

**Midterm Marks** - Midterm Examination marks are also normalized between 0 to 100. Generally, only one midterm was taken every semester except in one of them.

**Laboratory Marks -** Laboratory marks are also normalized between 0 to 100. However, very there was not much fluctuation in the weight of laboratory marks in every other different semester.

**Final Grade -** Classification techniques are used to predict the final grade. Class labels of final grades are A, B, C, D and F.

## 3.3 Probability Estimation on Continuous Data

In case of continuous data it is required to estimate the class. One of the methods is to assume a certain form of probability distribution for the continuous data and Gaussian Distribution is the most typical assumption which is stated in (1). In this equation $A_i$ is the $i^{th}$ instance of the attribute $A$ and $c_j$ is the $j^{th}$ class label. The symbol, $\mu$, stands for the population mean and the symbol, $\sigma^2$, stands for variance of the given population.

$$P(A_i \mid c_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} e^{-\frac{(A_i - \mu_{ij})^2}{2\sigma_{ij}^2}} \quad (1)$$

We used probability distribution function on the continuous attributes in the dataset for one of the three classification models,i.e., Naïve Bayes Classification, presented in the paper.

### 3.4 Discretization of Continuous Data Using Data Binning

Continuous features can be transformed into nominal features through a data preprocessing method called Data Binning. In this process, the continuous data is broken down into smaller number of bins. In this paper, we have implemented a discretization technique proposed by Fatih Kaya[10] which is based on equal width binning and error minimization. According to that paper, for a continuous attribute we will dynamically search for the bin width value until we find the optimal one. Secondly, data sets can have more than one continuous attribute. Therefore, finding optimal bin width value for all the continuous attributes in the data set will result in better overall performance.In Figure 1 a bar graph is shown which represents how different bin width values for the attribute Cumulative Grade Point Average affects the accuracy of Naive Bayes Classification model. On the *x*-axis, bin width values are provided and on the *y*-axis, accuracy of the classifier is provided. We can observe that when the bin width value is set to 4 we get the highest accuracy thus it is the optimal bin width value.

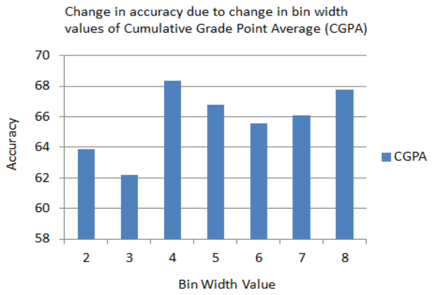


**Fig. 1.** Accuracy fluctuation due to bin width values

### 3.5 Classification Using Naive Bayes

Naïve Bayes is a method of classification in which Bayes theorem is applied. This classifier considers each feature to contribute independently to the probability and needs only a small amount of training data for classification. We used Naïve Bayes Classification to create two different models. In the first model we estimated the class labels for continuous attributes using probability distribution function (PDF) and in another model we used dynamic discretization method proposed in [10].

### 3.6 C4.5 Algorithm for Classification

C4.5 is an extension of Iterative Dichotomiser 3. In this algorithm, we need to calculate entropy of every attribute of dataset and then we have to split the dataset into subsets using the attributes of minimum entropy or maximum information gain. Some of the major extension of C4.5 from ID3 is that it accepts both continuous and discrete features, handles incomplete data points and different weights can be applied the features that comprise the training data [13]. We split the data using gain ratio and minimal size for split was set to 4. That means those nodes where the number of subsets are greater than or equal to 4 will be split.

### 3.7 Backpropagation Algorithm for Classifcation

Backpropagation is a training method of artificial neural network. It is used along with an optimization method called gradient descent. The Backpropagation algorithm is divided into two phases: propagation and weight update, more details about learning and weight update could be found in [14].

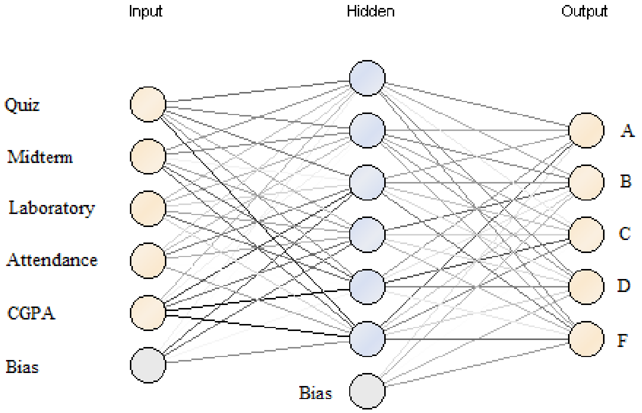


**Fig. 2.** Design of the Neural Network

In our model, which is shown in Figure 2, we used 3 layers of neurons: input layer, hidden layer and output layer. In input layer, the numbers of neurons are 6. They are basically attributes of the data set, such as Quiz, Midterm, Laboratory, Attendance, CGPA and one extra bias input. In output layers, the number of neurons are 5 and they represents are class label of the course grade. In hidden layer, the numbers of neurons are 6 with one extra input as bias neuron which makes the total number of neurons to 7. The number of neurons for hidden layer is calculated using the equation (2). The training of the Neural Network was done for 350 cycles with a learning rate of 0.1 and momentum 0.17.

$$No.ofNeurons = \frac{No.ofAttributes + No.ofClasses}{2} + 1 \tag{2}$$

### 3.8 Implementation of the Models

In order to build the models and analyze them we used the data mining tool RapidMiner 5. We import the preprocessed data set .xlsx file and preprocess in RapidMiner. During this process, we selected the grade attributes as the response variable and considered it as nominal type whereas the rest of the attributes are considered as numeric type. Then we pass the preprocessed data set into the validation system in which the classification model is built. The process is depicted in the Figure 3. At the top left corner we have the data block which is connected to validation block. Inside the validation block the data from the data block is connected to the classifier. In this case the classifier is Naive Bayes. The Apply Model block applies an already learnt trained model on the testing data set. The Performance block then calculates the accuracy of the trained model. For each block "tra" stands for training dataset, "mod" stands for model, "lab" means labeled data, likewise performance of the classifier is denoted as "per".

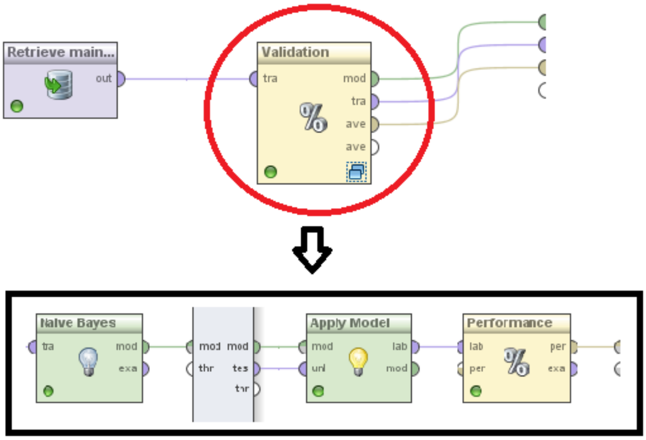
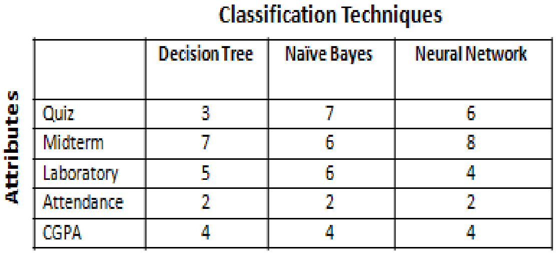


**Fig. 3.** Building a classification model in RapidMiner 5

## 4 Results

We assigned 20% of the original data set as the testing set to measure the accuracy of the models. Table 1 shows the optimal bin width value of the attributes for each of the classifiers. We used loop operation to run on each attributes to look for the optimum bin width value for each them and the set of bin width values which gives the best accuracy are chosen. Now for each of the classifiers chosen optimum bin width values actually varies other than for two attributes which are: Attendance Marks and Cumulative Grade Point Average (CGPA). Stratified Sampling which is a method of sampling is used because it gives better coverage of the whole population. The final accuracy of the model is measured by taking the average of the five iterations.

**Table 1.** Bin Width Values for the classification methods used



Classification Techniques

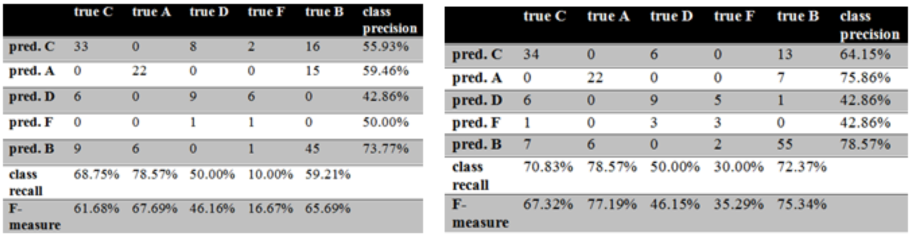
Attributes

| | Decision Tree | Naïve Bayes | Neural Network |
|---|---|---|---|
| Quiz | 3 | 7 | 6 |
| Midterm | 7 | 6 | 8 |
| Laboratory | 5 | 6 | 4 |
| Attendance | 2 | 2 | 2 |
| CGPA | 4 | 4 | 4 |

## 4.1 Naive Bayes Classification

For the model without optimal binning we estimated the class label of the continuous attributes using probability distribution function and then we build the model using Naive Bayes Classification. The accuracy of the model is 61.11%. The worst class precision comes from the prediction of class D as show in the Table 2(a). Optimal Equal Width as proposed by Fatih Kaya[10] is used to discretize every continuous attributes and then the model is build using the Naive Bayes Classification. All the optimal bin value for the continuous attributes are listed in the Table 2(b). The model has the accuracy of 68.33% which is 7.22% better than the previous Naive Bayes Model. Moreover the precision of the predicted class F is worst for this model.

**Table 2.** Detailed Analysis of the Naive Bayes Models



| | true C | true A | true D | true F | true B | class precision |
|---|---|---|---|---|---|---|
| pred. C | 33 | 0 | 8 | 2 | 16 | 55.93% |
| pred. A | 0 | 22 | 0 | 0 | 15 | 59.46% |
| pred. D | 6 | 0 | 9 | 6 | 0 | 42.86% |
| pred. F | 0 | 0 | 1 | 1 | 0 | 50.00% |
| pred. B | 9 | 6 | 0 | 1 | 45 | 73.77% |
| class recall | 68.75% | 78.57% | 50.00% | 10.00% | 59.21% | |
| F-measure | 61.68% | 67.69% | 46.16% | 16.67% | 65.69% | |

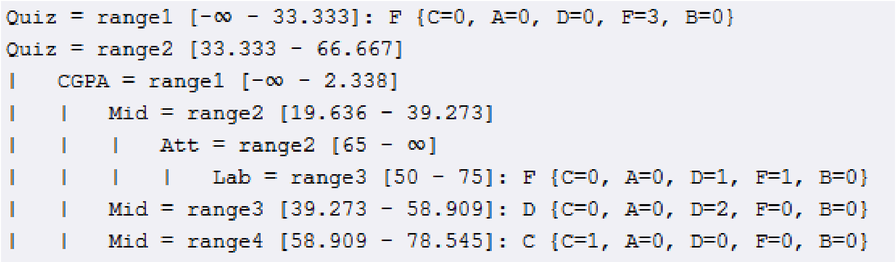
(a) Without optimal binning

| | true C | true A | true D | true F | true B | class precision |
|---|---|---|---|---|---|---|
| pred. C | 34 | 0 | 6 | 0 | 13 | 64.15% |
| pred. A | 0 | 22 | 0 | 0 | 7 | 75.86% |
| pred. D | 6 | 0 | 9 | 5 | 1 | 42.86% |
| pred. F | 1 | 0 | 3 | 3 | 0 | 42.86% |
| pred. B | 7 | 6 | 0 | 2 | 55 | 78.57% |
| class recall | 70.83% | 78.57% | 50.00% | 30.00% | 72.37% | |
| F-measure | 67.32% | 77.19% | 46.15% | 35.29% | 75.34% | |

(b) With optimal binning

## 4.2 Decision Tree Classification

For the model build using optimal equal width binning on decision tree, the accuracy is 63.89% which is a little over 7% better than the accuracy of the decision tree having all the continuous attributes discretized into 3 equal bins. The optimal bin values are listed in the Table 1. From the Table 3(b) we can also notice that the class precision of most of the class is similar to other models except for the class D which is 27.78%. This gives us a notion that due to misprediction of class D the overall accuracy for the decision tree model is affected.

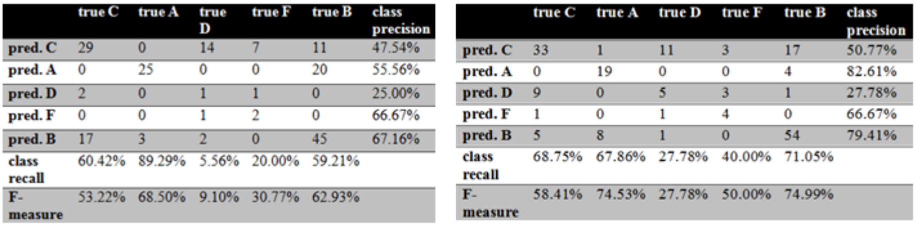


```
Quiz = range1 [-∞ - 33.333]: F {C=0, A=0, D=0, F=3, B=0}
Quiz = range2 [33.333 - 66.667]
|   CGPA = range1 [-∞ - 2.338]
|   |   Mid = range2 [19.636 - 39.273]
|   |   |   Att = range2 [65 - ∞]
|   |   |   |   Lab = range3 [50 - 75]: F {C=0, A=0, D=1, F=1, B=0}
|   |   Mid = range3 [39.273 - 58.909]: D {C=0, A=0, D=2, F=0, B=0}
|   |   Mid = range4 [58.909 - 78.545]: C {C=1, A=0, D=0, F=0, B=0}
```

**Fig. 4.** Decision Tree Model when using Optimal Equal Width Binning

Figure 4 represents a fraction of the decision tree model build after the data has been discretized using optimal equal width binning. From the figure we can derive the rules required to determine the students' grades. For Example, if a student having CGPA below 2.3, ends up getting between 33.3% to 66.7% in Quiz and less than 40% in Midterm and if the student gets between 50% to 75% in Laboratory he or she is mostly likely to get D as overall grade. Another rule that we can observe is that if the Quiz marks are below 33%, no matter what is the CGPA of the student, he or she will fail in the course. From the decision tree in Figure 4 we can understand that the attribute Quiz had the highest information gain which is then followed by CGPA.

**Table 3.** Detailed Analysis of the Decision Tree Models



| | true C | true A | true D | true F | true B | class precision |
|---|---|---|---|---|---|---|
| **pred. C** | 29 | 0 | 14 | 7 | 11 | 47.54% |
| **pred. A** | 0 | 25 | 0 | 0 | 20 | 55.56% |
| **pred. D** | 2 | 0 | 1 | 1 | 0 | 25.00% |
| **pred. F** | 0 | 0 | 1 | 2 | 0 | 66.67% |
| **pred. B** | 17 | 3 | 2 | 0 | 45 | 67.16% |
| **class recall** | 60.42% | 89.29% | 5.56% | 20.00% | 59.21% | |
| **F-measure** | 53.22% | 68.50% | 9.10% | 30.77% | 62.93% | |

(a) Without optimal binning

| | true C | true A | true D | true F | true B | class precision |
|---|---|---|---|---|---|---|
| **pred. C** | 33 | 1 | 11 | 3 | 17 | 50.77% |
| **pred. A** | 0 | 19 | 0 | 0 | 4 | 82.61% |
| **pred. D** | 9 | 0 | 5 | 3 | 1 | 27.78% |
| **pred. F** | 1 | 0 | 1 | 4 | 0 | 66.67% |
| **pred. B** | 5 | 8 | 1 | 0 | 54 | 79.41% |
| **class recall** | 68.75% | 67.86% | 27.78% | 40.00% | 71.05% | |
| **F-measure** | 58.41% | 74.53% | 27.78% | 50.00% | 74.99% | |

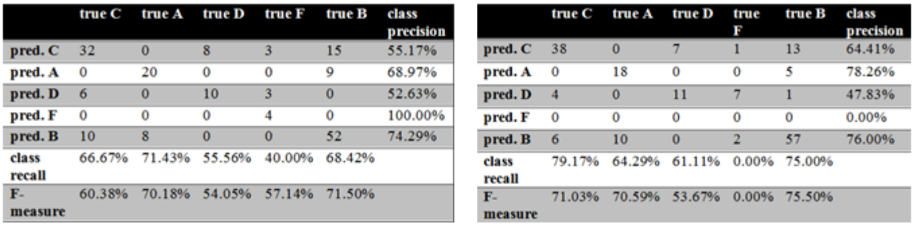
(b) With optimal binning

## 4.3 Classification Using Neural Network

The model built using the neural network is having an accuracy of 65.56%. The precision of class F is 100%. The accuracy of the neural network classification is the third best compared to the other models.

The optimal bin width values for the Neural Network Classification model is represented in the Table 1. When optimal equal width binning is used for Neural Network we get the better accuracy compared to every other models. However it is just 0.56% better than the Naive Bayes Classification for which optimal equal width binning is used. A detailed representation of results of neural network is given in the Table 4.

**Table 4.** Detailed Analysis of the Neural Network Model



| | true C | true A | true D | true F | true B | class precision |
|---|---|---|---|---|---|---|
| **pred. C** | 32 | 0 | 8 | 3 | 15 | 55.17% |
| **pred. A** | 0 | 20 | 0 | 0 | 9 | 68.97% |
| **pred. D** | 6 | 0 | 10 | 3 | 0 | 52.63% |
| **pred. F** | 0 | 0 | 0 | 4 | 0 | 100.00% |
| **pred. B** | 10 | 8 | 0 | 0 | 52 | 74.29% |
| **class recall** | 66.67% | 71.43% | 55.56% | 40.00% | 68.42% | |
| **F-measure** | 60.38% | 70.18% | 54.05% | 57.14% | 71.50% | |

(a) Without optimal binning

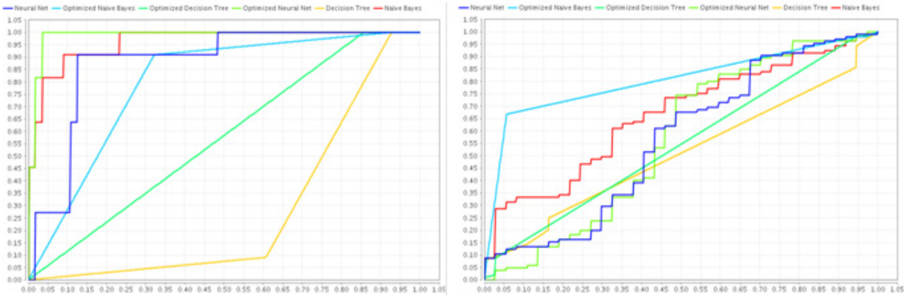
| | true C | true A | true D | true F | true B | class precision |
|---|---|---|---|---|---|---|
| **pred. C** | 38 | 0 | 7 | 1 | 13 | 64.41% |
| **pred. A** | 0 | 18 | 0 | 0 | 5 | 78.26% |
| **pred. D** | 4 | 0 | 11 | 7 | 1 | 47.83% |
| **pred. F** | 0 | 0 | 0 | 0 | 0 | 0.00% |
| **pred. B** | 6 | 10 | 0 | 2 | 57 | 76.00% |
| **class recall** | 79.17% | 64.29% | 61.11% | 0.00% | 75.00% | |
| **F-measure** | 71.03% | 70.59% | 53.67% | 0.00% | 75.50% | |

(b) With optimal binning

## 4.4 ROC Curve Comparisions

ROC curve which stands for receiver operating characteristic curve is the graphical representation of the performance of the binary classifier system for varying discrimination threshold [15]. The horizontal axis represents the fraction of false positives out of total actual negatives (FPR = False positive rate) and the vertical axis represents the fraction of true positives out of total actual positives (TPR = True positive rate). As ROC curve is a binary classifier system but we have five class labels for the grade so we are presenting five ROC curves. For each ROC curve one class is considered as True class and the rest of the classes are considered as False class.

In Figure 4(a), the ROC curve of the Neural Network Model where the continuous data is discretized using optimal equal width binning is represented along with other the curves of the other models. The area under the ROC curve is 0.985 for this model whereas the area uder the ROC curve for Naive Bayes using probability distribution function is about 0.93. This indicates the first model mentioned has better True Positive coverage for class A.

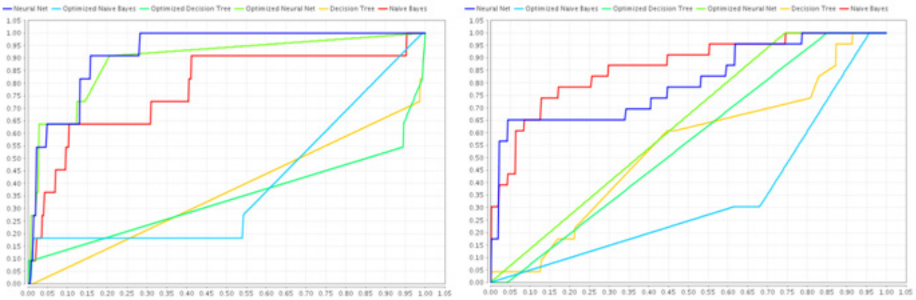


(a) ROC Curve for Class A

(b) ROC Curve for Class C

**Fig. 5.** ROC Curves of the models for the Class A and C

For the class C, Naive Bayes Classification Model where the continuous attributes are discretized using optimal equal width binning has the under the curve of around 81% coverage of the total area as shown in the Figure 5(b). Figure 6(a), represents the ROC curve of the class D and the ROC curve of Neural Network Classification Model where the continuous data are discretized using optimal equal width binning is the second best to that model for this class.



(a) ROC Curve for Class D

(b) ROC Curve for Class B

**Fig. 6.** ROC Curves of the models for the Class D and B

For Class B largest area under the curve is of the Naive Bayes Classification Model have the area under the curve which is 0.8335. The ROC curves are shown in Figure 8. ROC Curve for the class label F is represented in the Figure 7. Optimized Naive Bayes Classification Model is having area under the curve of about 85% of the whole area. Naive Bayes Classification Model is having an area of about 70% which is very close to Neural Network Classification Model having AUC of about 69%.

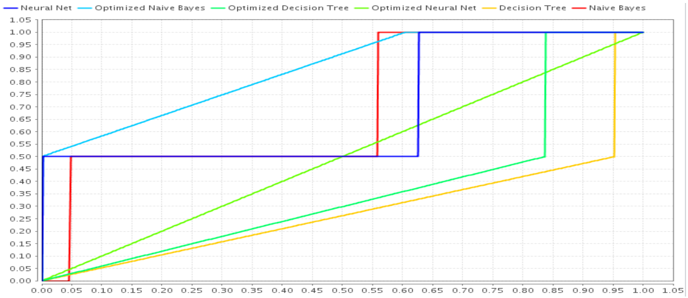
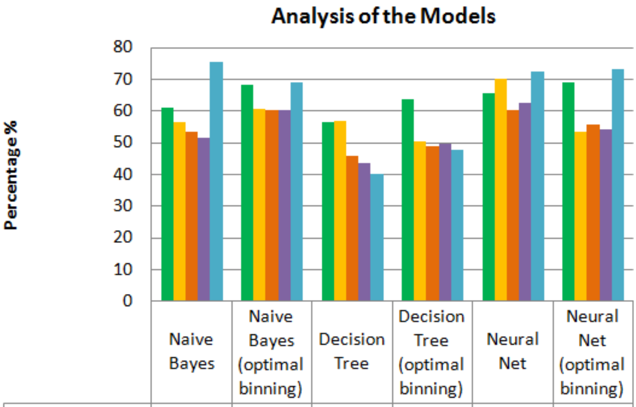


**Fig. 7.** ROC Curves of the models for the Class F

### 4.5 Summary of the Analysis

In Figure 8 an overall summary of the models are represented. On top of the figure multiple bar graphs are shown where on the *x*-axis each chunk highlights analysis of a particular model. The *y*-axis represents the percentage of the results obtained for accuracy, average precision, average recall, average F-measure and average area under the ROC curves (AUC) for each model. The table in the Figure 8 is the percentage values for which the bar graphs are built. We investigated the correlation between the classification model accuracy and the area under the ROC curve of each of the models using Pearson Correlation Coefficient [17] and the value of the correlation is 0.7567 which indicates that there is a positive correlation between accuracy and AUC.



| | Naive Bayes | Naive Bayes (optimal binning) | Decision Tree | Decision Tree (optimal binning) | Neural Net | Neural Net (optimal binning) |
|---|---|---|---|---|---|---|
| 🟩 Accuracy | 61.11 | 68.33 | 56.67 | 63.89 | 65.56 | 68.89 |
| 🟨 Avg. Precision | 56.4 | 60.86 | 56.9 | 50.56 | 70.21 | 53.3 |
| 🟧 Avg. Recall | 53.306 | 60.35 | 45.73 | 48.96 | 60.41 | 55.91 |
| 🟪 Avg. F-Measure | 51.58 | 60.26 | 43.444 | 49.54 | 62.65 | 54.16 |
| 🟦 Avg. AUC | 75.6 | 68.9 | 40.1 | 47.9 | 72.3 | 73.1 |

**Fig. 8.** Analysis of the models

## 5 Conclusion and Future Work

Our objective was to build a model that will predict the grade of the students in a particular course. Moreover we wanted to know which classifiers work better with optimum equal width binning technique when presented with continuous dataset. We have successfully built six models and presented comparative analysis between them. There was substantial misclassification error for Class D and Class F for almost every model which we believe is due to fewer numbers of instances in the data set. However, we believe when more semesters will pass by, the number of instances will be increased in the course database that could help us in building a better model.

## References

1. Pal, A.K., Pal, S.: Analysis and Mining of Educational Data for Predicting the Performance of Students. International Journal of Electronics Communication and Computer Engineering 4(5), 1560–1565 (2013)
2. Ayers, E., Nugent, R., Dean, N.: A Comparison of Student Skill Knowledge Estimates. In: International Conference on Educational Data Mining, Cordoba, Spain, pp. 1–10 (2009)
3. Bharadwaj, B.K., Pal, S.: Data Mining: A prediction for performance improvement using classification. International Journal of Computer Science and Information Security 9(4), 136–140 (2012)
4. Bharadwaj, B.K., Pal, S.: Mining Educational Data to Analyze Students' Performance. International Journal of Advance Computer Science and Applications 2(6), 63–69 (2011)

5. Chu, H.C., Hwang, G.J., Tseng, J.C.R., Hwang, G.H.: A computerized approach to diagnosing student learning problems in health education. Asian Journal of Health and Information Sciences 1(1), 43–60 (2006)
6. Hsia, T., Shie, A., Chen, L.: Course planning of extension education to meet market demand by using data mining techniques - an example of Chinkuo technology university in Taiwan. Expert System Application Journal 34(1), 596–602 (2008)
7. Rus, V., Lintean, M., Azevedo, R.: Automatic detection of student mental models during prior knowledge activation in MetaTutor. In: International Conference on Educational Data Mining, Cordoba, Spain, pp. 161–170 (2009)
8. Zhu, F., Ip, H.H.S., Fok, A.W.P., Cao, J.: PeRES: A personalized recommendation education system based on multi-agents & SCORM. In: Leung, H., Li, F., Lau, R., Li, Q. (eds.) ICWL 2007. LNCS, vol. 4823, pp. 31–42. Springer, Heidelberg (2008)
9. Yadav, S.K., Bharadwaj, B., Pal, S.: Data mining applications: A comparative study for predicting student's performance. arXiv preprint arXiv:1202.4815 (2012)
10. Kayah, F.: Discretizing Continuous Features for Naive Bayes and C4. 5 Classifiers. University of Maryland Publications (2008)
11. Tan, P., Kumar, V., Steinbach, M.: Introduction to Data Mining. Addison Wesely, NY (2006)
12. Naive Bayes Classifier: In Wikipedia (June 1, 2014), `http://en.wikipedia.org/wiki/Naive_Bayes_classifier` (retrieved)
13. Quinlan, J.R.: C4. 5: programs for machine learning. Morgan Kaufmann (1993)
14. Backpropagation: In Wikipedia (June 1, 2014), `en.wikipedia.org/wiki/Backpropagation` (retrieved)
15. Receiver Operating Characteristic. In Wikipedia (June 1, 2014), `http://en.wikipedia.org/wiki/Receiver_operating_characteristic` (retrieved)
16. Romero, C., Ventura, S.: Educational data mining: a review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 40(6), 601–618 (2010)

# Beautiful Trochoids

Noor Deenina Hj Mohd Salleh and Somnuk Phon-Amnuaisuk

Media Informatics Special Interest Group,
School of Computing & Informatics,
Institut Teknologi Brunei, Brunei Darussalam
{`deenina.salleh,somnuk.phonamnuaisuk`}`@itb.edu.bn`

**Abstract.** This paper examines the aesthetic dimensions of the patterns generated from trochoids. A virtual trochoid was implemented and its parameters were varied to generate various patterns. These patterns were evaluated by a group of 101 participants. Inspired by Birkhoff's concept of measuring aesthetics, order to complexity, we calculated order as the quality related to composition and complexity as the quality related to the intricacy of the structure. The result of this experiment suggests that the aesthetic measure was able to predict at least half of the preferred patterns by the participants and that those with art experience were more developed in considering a balanced composition. Additionally, there were certain patterns that both experienced and non-experienced participants agreed on. Our findings suggest that Birkhoff's aesthetic measure reveals useful aesthetic information of trochoid patterns.

**Keywords:** Aesthetic evaluation, Trochoids, Order, Complexity.

## 1 Introduction

Aesthetic evaluation refers to the sensory appreciation of an artifact. Aesthetic appreciation involves both cognitive and emotional process [1]. This appreciation is arguably dependent on social norms. For example, classical paintings and pop-art pieces emphasise different values. Nevertheless, one cannot deny the idea of absolute beauty since there seem to be some qualities that universally appeal to humans.

This paper investigates the aesthetic dimensions of trochoids. A trochoid is a curve traced from a fixed point $p$ on a diameter or an extended diameter of a circle $C_r$ that rolls along a fixed line. The trochoid subsumes various specific cases: hypotrochoid, epitrochoid, hypocycloid, and epicycloid. These cases are roulette-like where a fixed point $p$ of a circle $C_r$ (as described earlier) rolls along a circumference of another fixed circle $C_f$. The two circles could have any radius, but both must not be infinity[1].

The patterns generated from trochoids are generally pleasing (or at least not disturbing). Here, we would like to examine the aesthetic dimensions of the

---

[1] See `http://mathworld.wolfram.com/Trochoid.html` for more detailed explanations.

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems*, Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_17

patterns generated from trochoids. A virtual trochoid was implemented and its parameters were varied to generate various patterns. Sixty six patterns were selected from over eight hundred patterns generated and they were then grouped into 22 questions, each with three patterns. A group of 101 participants was asked to select the most pleasing pattern among the three patterns presented in each question. We then compare and contrast preferences observed from these participants with the preferences calculated using aesthetic measures derived from image processing techniques.

The rest of the materials in the paper are organized into the following sections: Section 2: Related works; Section 3: Aesthetic evaluation; Section 4: Results & Discussion; and Section 5: Conclusion and Future work.

## 2 Related Works

Could one explicitly express the evaluation criteria for beauty? There are two main viewpoints on this matter. First and foremost, Plato argues that beauty is a quality associated with harmony, symmetry and unity [2,3], which implies that humans are attracted to these qualia naturally. Secondly, according to Edmund Burke (1729-1797) [4] (p. 131), *"beauty is for the greater part, some quality in (the) bodies, acting mechanically upon the human mind by the intervention of the senses."* In Plato's view, beauty is an absolute quality and is universal. In Burke's view, beauty is dependent on each individual's perception, past experiences, knowledge and preferences.

A person's preference depends on what one considers as sensational [5]. Although there is scientific evidence that sensational reactions can be observed from chemical reactions in the brain, it is often difficult to quantitatively distinguish a person's sensational reaction to another. These reactions are also likely to change overtime, depending on the types of exposure they experience.

Plato's theory of an ideal proportion, symmetry and unity influences many later attempts to quantify beauty as some numeric parameters derived from the structure of an artifact. *Golden proportion* is one of such quantifications. In generative art, the relationships between the complexity of a computer program (used to generate an artifact) and the perceived complexity/order in the artifact, are investigated by many researchers [6,7,8].

Birkhoff [9] argued that the aesthetic value for any class of aesthetic objects may be measured using the ratio of *order O* over *complexity C.*

$$M = \frac{O}{C} \tag{1}$$

It is clear that the definition of $O$ and $C$ are class dependent. For example, a terse, imaginative expression of a poetic idea in metric form is a characteristic of high aesthetic value poem pieces [9]; in music, the classical ears emphasise the beauty of the musical form while the romantic ears emphasise expressions.

Birkhoff's aesthetic values, calculated from the perceptions of order and complexity, do not support or defy any of the two views above. Many researchers argue that it is impossible to determine the preference of a person from this formula, due to the progressive experiences involved as mentioned earlier in [10]. So how can aesthetics be measured in this experiment? Solely for the purpose of this, we define aesthetics as beauty in its dimensions rather than what is being interpreted as its content [10]. The evaluation of the dimensions will include some of the design principles which are categorized accordingly into two main components: Composition and Complexity. The component of composition will evaluate the trochoid in terms of its balance and symmetry, which refers to having unity in the overall form in terms of its shape, size and negative spaces in-between; whereas complexity is based on the intricacy of the structure created from repetitive patterns of the geometric shapes, intensity of the dots and thickness of the lines [11] (p. 101 - 122).

## 3 Aesthetic Evaluation: Beautiful Trochoids

### 3.1 A Virtual Trochoid

Figure 1 shows a rolling circle $C_r$ with radius $r_r$ (a) rolling inside and (b) outside a fixed circle $C_f$ (having radius $r_f$). The circle $C_f$ is centered at $(0, 0)$. The relationship between the angle $t$ at (0,0) and $t'$ at $(x_c, y_c)$ when $C_r$ rolls inside $C_f$ can be expressed as follows:

$$t' = \frac{r_f - r_r}{r_r} t \tag{2}$$

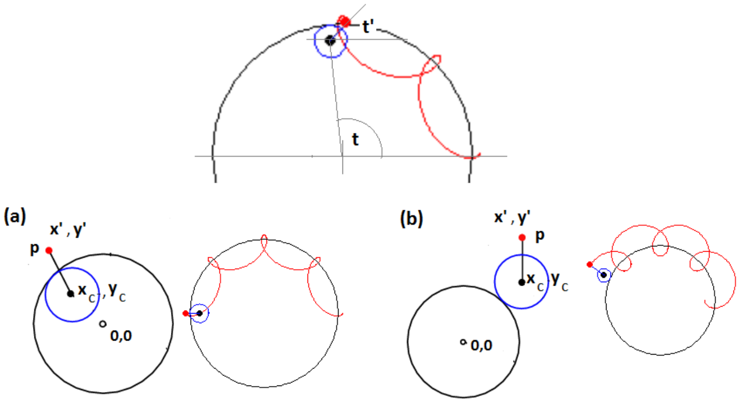


**Fig. 1.** Examples of (a) Hypotrochoid and (b) Epitrochoid

Let $(x_c, y_c)$ be the coordinate of the center of the circle $C_r$ and $(x', y')$ be the coordinate of the tracing point $p$ in the reference frame of $C_r$ i.e., $(x_c, y_c)$ is the origin of $(x', y')$ and the distance between the two points is

$\rho = \sqrt{(x' - x_c)^2 + (y' - y_c)^2}$. The locus of $p$ in the original $C_f$ coordinate reference frame is expressed as:

$$x = x_c + x'; \qquad y = y_c + y'; \tag{3}$$

$$\text{where } x_c = (r_f - r_r)cos(t); \quad y_c = (r_f - r_r)sin(t); \tag{4}$$

$$\text{and } x' = (\rho)cos(t'); \quad y' = (\rho)cos(t'); \tag{5}$$

After rearranging the variables using $k = r_r/r_f$ and $l = \rho/r_r$, the parametric equations of a trochoid are expressed as follows:

$$x(t) = r_f[(1-k)cos(t) + (lk)cos(\frac{1-k}{k}t)]; \tag{6}$$

$$y(t) = r_f[(1-k)sin(t) - (lk)sin(\frac{1-k}{k}t)]; \tag{7}$$

In the above parametric form, the parameter $r_f$ is a scaling parameter that determines the size of the drawing, the parameters $l$ and $k$ determine the structure of the trochoids.

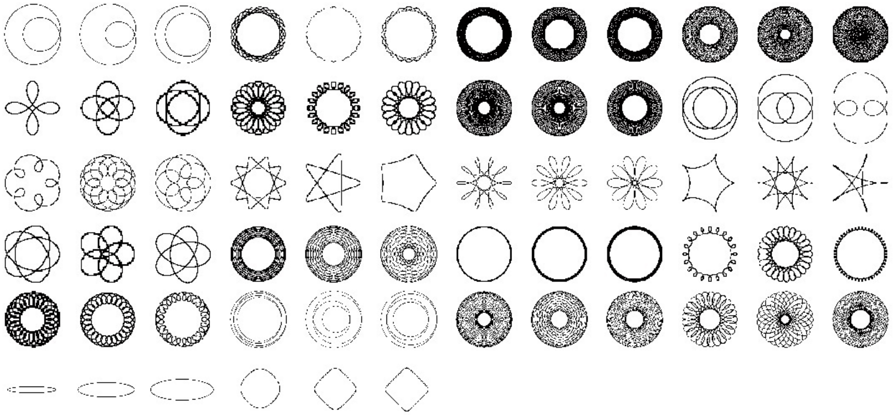


**Fig. 2.** Examples of trochoids that have been manually selected from over 800 generated patterns. These are patterns from questions 1 to 22.

### 3.2 Quantitative Aesthetic Measures

Various parameters have been suggested to objectively quantify beauty [12]. Benford's Law has been employed to describe the distribution of brightness of the pixels of an image in [13]. In [14], it was suggested that colour gradients in a visually pleasing image conform to a normal distribution.

Our approach is inspired by Birkhoff's concept of measuring aesthetics through the relationship between order and complexity, where order refers to the organisational structure of components in artifacts, and complexity refers to the complex interactions among components in the artifact that reveal various coherent interpretations. In our study, we have decided to calculate the features $f_i$ derived from balance and symmetry to quantify order $O$. It is, however, quite hard to quantify complexity $C$ in the sense described above since coherent interpretations imply semantic processing which is far beyond the scope of this work. Here, we have decided to use two different features to quantify complexity $C$: the intricacy of the structure created from repetitive patterns of the geometric shapes. Hence the aesthetic measure $M$ of each pattern can be expressed as:

$$M = \frac{\sum_i (w_i f_i)}{C} \tag{8}$$

Details of the calculation of $O$ and $C$ are as follows.

**Calculating Order.** Let $T^{m \times n}$ be an $m \times n$ binary matrix representing a trochoid pattern i.e., element $T_{ij} \in \{0, 1\}$ s.t. 1 denotes a pixel filled with black ink and 0 denotes a white pixel. Two types of balance measures are defined in the present study, (i) as the ratio of inked pixels between the left and the right parts $f_1$ and (ii) as the ratio of inked pixels between the top and the bottom parts of the pattern $f_2$.

$$\begin{aligned} a &= \textstyle\sum_{ij} T_{ij}, i \in \{1..m\}, j \in \{1..n/2\} \\ b &= \textstyle\sum_{ij} T_{ij}, i \in \{1..m\}, j \in \{(1 + n/2)..n\} \\ c &= \textstyle\sum_{ij} T_{ij}, i \in \{1..m/2\}, j \in \{1..n\} \\ d &= \textstyle\sum_{ij} T_{ij}, i \in \{(1 + m/2)..m\}, j \in \{1..n\} \\ \\ &\textbf{if } (b \geq a) \quad \textbf{then } f_1 = a/b \quad \textbf{else } f_1 = b/a \\ &\textbf{if } (d \geq c) \quad \textbf{then } f_2 = c/d \quad \textbf{else } f_2 = d/c \end{aligned} \tag{9}$$

Two types of symmetric measures are defined, (i) $f_3$ symmetry along the $x$ axis and (ii) $f_4$ symmetry along the $y$ axis.

$$\begin{aligned} g &= \textstyle\prod_{ij,kl} T_{ij} T_{kl}, i \in \{1..m\}, j \in \{1..n/2\}, k = i, l = j + n/2 \\ h &= \textstyle\prod_{ij,kl} T_{ij} T_{kl}, i \in \{1..m/2\}, j \in \{1..n\}, k = i + m/2, l = j \\ f_3 &= 2g / \textstyle\sum_{ij} T_{ij}, i \in \{1..m\}, j \in \{1..n\} \\ f_4 &= 2h / \textstyle\sum_{ij} T_{ij}, i \in \{1..m\}, j \in \{1..n\} \end{aligned} \tag{10}$$

**Calculating Complexity.** Here, we measure perceived complexity using the quality related to the activities of trochoid patterns. Let $x_p = \sum_{i'j'} T'_{i'j'}$ where a patch $T'$ was a binary matrix of size $m' \times n'$ sampled from $T$ ($T'^{m' \times n'} \subset T^{m \times n}$). The patch height and width are set at $m' = \lfloor m/10 \rfloor$ and $n' = \lfloor n/10 \rfloor$. Patches are sampled from left to right and from top to bottom without overlapping. This produces a sequence $\mathbf{x}$ that reflect the activities of the trochoid pattern.

$$C = \sum_{p=0} x_p \tag{11}$$

### 3.3 Experimental Design

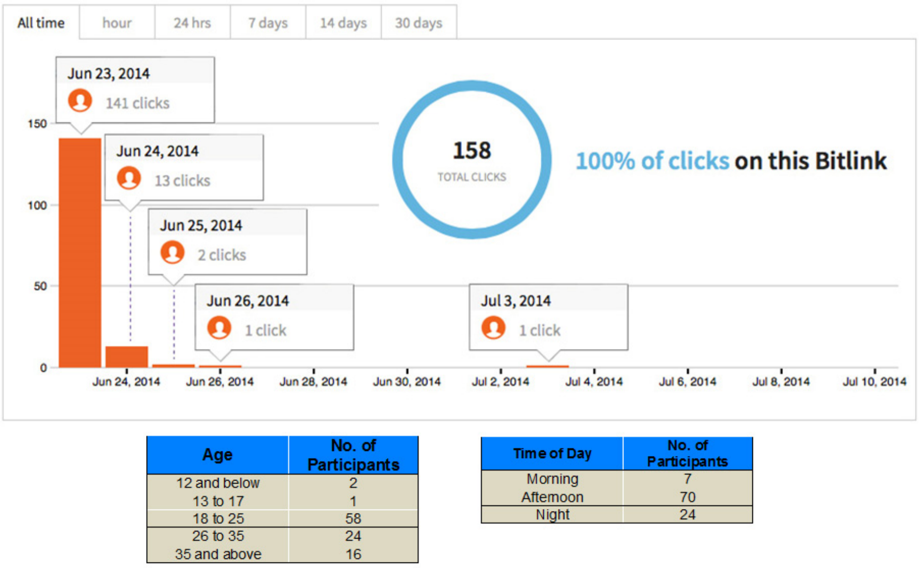
This experiment has two objectives: firstly, to compare the participants' selections among the set of given trochoids; secondly, although Hoenig argued that it is nearly impossible to conclude the universal preference of a group based purely on their aesthetic opinions [10], this experiment attempts to investigate whether or not there are predictive patterns that would account for the perception of the aesthetic dimension.

**Materials.** Eight hundred trochoid images were produced by varying the parameters $l$ and $k$ as discussed in section 3.1. This produced many similar patterns. The first stage of filtration was to choose the images based on its distinctive patterns. Once the images were selected, patterns which resembled each other were strategically grouped together in a set of three. A total of 66 patterns were chosen and grouped together into 22 sets (of three patterns).

According to Costello et al, (p. 115) [11], when images are being placed closely together, it is easier for a person to process the visual information as they are likely to perceive those images as related and thus be able to compare the overall unity of the design. Even though most of the images seemed similar, they varied in terms of the focal point of emphasis.

**Data Collection.** The study was designed as an online survey created on Google Sites. Participants were asked to fill in some information regarding themselves: name, age, sex, exposure towards any form of art education and time of the day they answered the survey, as well as selecting the most preferred image for each of the 22 questions. The web link to the online survey was shortened using the services from Bitly.com. The site revealed information regarding the number of clicks on the link on which day, from what sites and from which country. A QR code was also generated for the convenience of the participants. In order to find a range of participants, the link and QR code were disseminated through social media: Facebook, Instagram and WhatsApp.

According to the status report from Bitly.com, there was a total of 158 clicks recorded, but only one hundred and one participants actually answered the survey. Out of those participants, 68.3% were females. Fifty eight percent (58%) of the responses were between 18 to 25 years old. Fifty five percent (55%) of the responses were exposed to some form of design education; self-taught as a hobby or learnt from school. Participants also answered the survey more in the afternoon (see graphical summary in Figure 3).



| Age | No. of Participants |
|---|---|
| 12 and below | 2 |
| 13 to 17 | 1 |
| 18 to 25 | 58 |
| 26 to 35 | 24 |
| 35 and above | 16 |

| Time of Day | No. of Participants |
|---|---|
| Morning | 7 |
| Afternoon | 70 |
| Night | 24 |

**Fig. 3.** Graphical summary of the number of clicks of the Bitlink as well as participants' age groups, time and date of their participations in this study.

## 4 Results and Discussion

A total of 101 participants volunteered to evaluate these 66 patterns which were grouped into 22 questions. Their preferences were recorded and summarised in Figure 4 (bottom pane). These 66 patterns were also evaluated using the measure in Equation 3.2. A comparison between the predicted selection counts against the participants selection counts shows a 54.5% similarity (see Figure 4).

### 4.1 Data Analysis

Upon evaluation, it was found that participants preferred patterns which had a compositional weight [8] (p. 12). Also, across all of the preferred patterns,16 of them significantly had greater responses from participants who had had experience in art. This serves as plausible evidence that they agreed on what constituted a standard structure of composition. Apart from that, questions 1, 8 and 9 (respectively in Figure 2: Row 1, Column 3; Row 2, Column 10; Row 3, Column 3) had equal amount of responses for the preferred pattern from both experienced and non-experienced participants. The trochoids had a distinctive attribute where there was a considerable amount of negative space between the intricacies of the lines.

Other interesting patterns related to certain aspects of the participants' backgrounds were found and highlighted below:

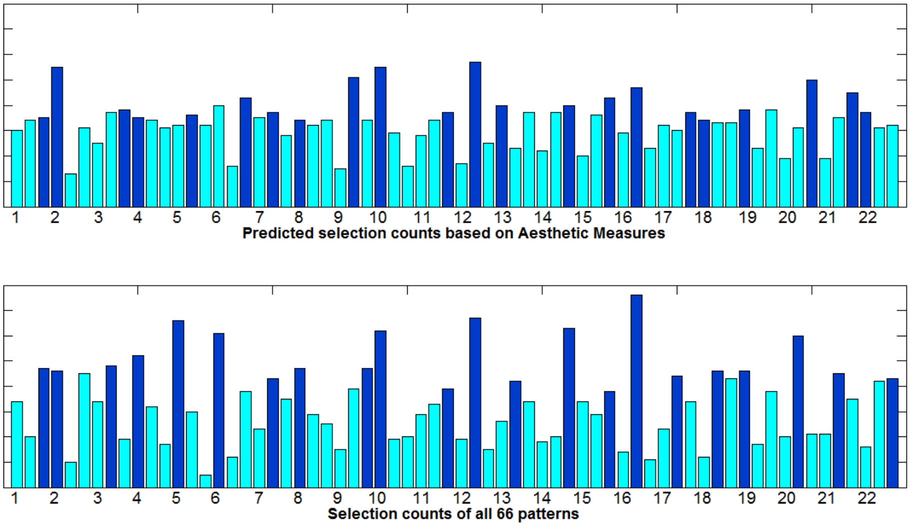


**Fig. 4.** Comparison between collective preferences from 101 participants and from our equation $M = \frac{\sum_i (w_i f_i)}{C}$. Each question number label spans over three bars denoting the three choices of patterns available in each question. The height of each bar signifies the preference counts for the choices. The most preferred choice in each question is shaded in a darker tone.

1. Participants with the age below 17 years old (both genders) had selected patterns that shows a central focal point of attraction, compensated with a balanced dimension of the lines and negative spaces (Figure 2: Row 2, Column; Row 4, Column 11).
2. Females participants, aged 18 to 25, who has had experience in art education and answered during the night, selected a pattern which has a defined central focus and reasonable negative spaces (Figure 2: Row 4, Column 6). Having such space adds a breathing room for the patterns to be processed by the viewer [11] (p. 104).
3. Participants' aged 18 to 25, who had no art education, mostly chose the predicted patterns.

Due to the fact that there were no continuous observations for the same participants to answer the same questions for the different times of the day, the gathered data concerning time is not able to give solid evidence of the preference across the different background aspects of the participants. This also applies to gender and age. However, it did not affect our experiment since the objective was to find whether there was a predictive pattern in perceiving aesthetic dimensions.

### 4.2 Discussion

From this investigation, participants preferred patterns with a central attraction as well as having visual weight as a whole. Arnheim argues that visual balance does not have to correspond to its physical balance, where size, colour and position plays an important role in determining what is considered as a unified composition [15] (p. 9). In general, having such visual weight depends firstly on the size of the shape, where a vertical shape appears to be heavier than oblique shapes [15] (p. 13). Secondly, an intensive colour is able to have central focus compared to lighter colours [11] (p. 121). Having equal negative space in between the lines is also taken into consideration in order for the viewer to easily process the pattern [11] (p. 104). Lastly, the weight given from the direction of the shape depends on how well the other factors can counteract it and therefore will show the central perceptual force [15] (p. 15).

## 5 Conclusion and Future Work

For now, we can rationally conclude that mostly those with art experience are able to define what is considered as a balanced composition through its symmetry and that as a result, the aesthetic measure used in this experiment was able to predict half of the preferred patterns. We also found that there were three similar patterns in which both the experienced and non-experienced agreed on. In spite of that, we would in the future observe from the new discovery as well as regarding compositional weight and attempt to introduce it into the formula so that it is able to calculate the weight as a whole. There were also some underlying questions regarding the preference of the participants depending on their background. Will the perception from both genders, despite their experience, be similar when the age factor has been applied? Or will answering the same question during different times and mood affect the outcome? Does the size of the trochoids matter when it is being presented to the viewer? These uncertainties might be substantial in determining the evaluation of aesthetic dimensions. In light of this, an area that is possible to venture into would include manipulating the trochoids into their three-dimensional counterparts; would aesthetic measure be able to include depth as a factor to the perception of beauty in its dimension?

**Acknowledgement.** We wish to thank anonymous reviewers for their comments, which help improve this paper. We would also like to thank GSR Institut Teknologi Brunei for their partialfinancial support given to this research.

## References

1. Leder, H., Belke, B., Oeberst, A., Augustin, D.: A model of aesthetic appreciation and aesthetic judgements. British Journal of Psychology 95, 489–508 (2004)
2. Child Jr., W.C.: Monroe Beardsley's three criteria for aesthetic value: A neglected resource in the evaluation of recent music. Journal of Aesthetic Evaluation 34(2), 49–63 (2000)
3. Feng, C.: Looking good: The psychology and biology of beauty. Journal of Young Investigators 6 (2002), `http://www.jyi.org/issue/looking-good-the-psychology-and-biology-of-beauty/` (retrieved June 28, 2014)

4. Burke, E.: The Works of the Right Honorable Edmund Burke, `http://books.google.com` (retrieved June 15, 2014)
5. Palmer, S., Shimamura, A.: The psychology of aesthetics. UC Berkeley Event (March 2, 2012), YouTube: `https://www.youtube.com/watch?v=UwsEeQpxkFw` (retrieved June 15, 2014)
6. Wolfram, S.: Cellular automata as models of complexity. Nature 311(4), 419–424 (1984)
7. Flake, G.: The Computational Beauty of Nature. The MIT Press (1998)
8. Phon-Amnuaisuk, S., Panjapornpon, J.: Controlling generative processes of generative art. Procedia Computer Science 13, 43–52 (2012)
9. Birkhoff, G.D.: Aesthetic Measure. Harvard University Press, Cambridge (1993)
10. Hoenig, F.: Defining Computational Aesthetics. In: Neumann, L., Sbert, M., Gooch, B., Purgathofer, W. (eds.) Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging (2005)
11. Costello, V., Youngblood, S.A., Youngblood, N.E.: Multimedia Foundations: Core Concepts for Digital Design. Focal Press (2013)
12. Sudweeks, F., Simoff, S.J.: Quantifying beauty: An information system for evaluating universal aesthetics. In: Proceedings of the Second Western Australian Workshop on Information Systems Research, Perth, Australia (1999)
13. den Heijer, and E., Eiben, A.: Using aesthetic measures to evolve art. In: Proceedings of the IEEE Congress on Evolutionary Computation (CEC 2010), Barcelona (2010)
14. Ross, B.J., Ralph, W., Zong, H.: Evolutionary image synthesis using a model of aesthetics. In: Proceedings of the IEEE Congress on Evolutionary Computation (CEC 2006), pp. 1087–1094 (2006)
15. Arnheim, R.: Art and Visual Perception: A Psychology of the Creative Eye. University of California Press (1954)

# The Impact of Knowledge, Skill, Attitude and Confidence in Information Communication and Technology in Teaching and Learning among Teachers in Technical School

Azizi Yahaya[1], Gooh Mo Lee[2], Halimah Ma'alip[3], Dayang Ku Zainab Pg Hj Tuah[4], Zeliha Mohammad Ali[5], Kamariah Hj Kasah[6], and Zarifa Binti Mohd Sis[2]

[1] Sultan Hassanah Bolkiah Instituit of Education, Universiti Brunei Darussalam
`profazizi_yahaya@yahoo.com`
[2] Sekolah Kebangsaan Convent Infant Jesus 1, Melaka, Malaysia
`moleegooh@gmail.com`
[3] Languange Academic, Universiti Teknologi Malaysia Johor
`m-halima@utm.my`
[4] Kolej Univerisiti Perguruan Ugama Seri Begawan, Brunei Darussalam
`dk.zainab988@yahoo.com`
[5] Maktab Sultan Omar Ali Saifuddien, Brunei Darussalam
`lihaali6.za@gmail.com`
[6] Sekolah Rendah Pantai Berakas, Brunei Darussalam
`ariel.193@hotmail.com`

**Abstract.** The aim of this study is to identify the preparation of using Information Communication and Technology (ICT) in teaching and learning from aspects of knowledge, skills, attitude and confidence among the technical teachers of technical schools in Johore, Malacca and Negeri Sembilan. A total of 92 teachers from nine technical schools in three states were randomly selected (cluster over cluster method) in this study. The findings show that the preparation of Technical teachers' knowledge and skills was at the moderate level. While the attitude and confidence in using Information Communication and Technology in teaching and learning was at high level. The findings also show that the level of preparation has s significant difference with frequency in using computer and computer ownership.

**Keywords:** knowledge, skills, attitude and confidence.

## 1 Introduction

The new generation faces the challenges and influences of Information Communication and Technology (ICT). As [1] point out, Information Communication and Technology explosion and globalizing challenges which get along more conspicuous and it should manage by judicious and increase the awareness of risk that bring. This is because Information Communication and Technology development has already made a change in human work, communication and their thinking [2]. Researchers in [3]

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_18

also believe that using Information Communication and Technology can potentially bring Malaysia's education system towards the international level. Now, whether the Malaysian residents appreciate it or not, the reality is information communication and technology has started long time ago.

Information communication and technology can be recognized as a combination of one set of technology especially microelectronic computer and communication technology, which helps in collecting, saving, processing, sending and presenting data or information through various media such as image, graphic, video, audio and texts [4]. In addition, Information Communication and Technology is also an agent of great positive changing until it influence all aspect in various sectors like sector of education, industry, business, medical, agriculture and others [5].

Intelligence in Information Communication and Technology of insistence is for everybody in the future [6]. With this, principles of education in a country always been studied to upgrade education quality and make sure it was effectiveness in execute to face the challenges of Information Communication and Technology. This can be seen from programs which organized by the Ministry of Education such as smart school, computer in education, program teaching and learning Science and Mathematics in English, Malaysian Great For Learning [7].

As [3] point out, program training in service for teachers in Information Communication and Technology field is also executed by Ministry of Education of Malaysia such as smart school teacher training, computer education fourteen weeks in service course and basic computer literacy. Until the year of 2000, almost 60,000 teachers already participate in the training program. In this situation, teachers should make an effort to use information communication and technology in teaching and learning or in education management. Meanwhile from 1996 to 2000, only 30 percent teachers just have been trained in Information Communication and Technology and only a small amount of teachers from that percentage can integrate Information Communication and Technology in their teaching and learning. From that percentage, only a few of the teachers has the knowledge and skills to develop course software's [3].

In addition, there are still many teachers who lack of effort to Information Communication and Technology in teaching and learning. They are also not expert in using Information Communication and Technology in teaching design and lack of skills in selecting, evaluating and using course software based on student needs [3] As [8] points out among 280,000 in service teachers has basic knowledge and skills in Information Communication and Technology when they studied in university.

Besides failure to master ICT which is considered to be the root of the problems by many teachers teaching Science and Mathematics in English, MOE had arranged many courses that are related in using of ICT for teachers [9]. Findings of the [10] study also shows that skill in using Information Communication and Technology among technical teachers were in moderate level. In addition, some of the teachers shows negative attitude towards Information Communication and Technology among teachers especially Science and Mathematics teachers and language teachers [11]. Most of teachers seldom use Information Communication and Technology because they are not confidence of using it [12].

Hence, this study is proposed to identify the teacher's readiness of using Information Communication and Technology in teaching and learning from aspect of knowledge, skills, attitude and confidence among technical teachers of technique school in Johore, Malacca and Negeri Sembilan.

## 2 Research Objectives and Methodology

The specific objective of this study is to explore and identify the following information among teachers in technical secondary schools in Johor, Melaka and Negeri Sembilan: (i) Identify the level of knowledge in information and communication Technology among teachers in technical secondary school in Johor, Melaka and Negeri Sembilan; (ii) Identify the level of technical skill of teachers in information and communication technology in technical secondary schools in Johor, Melaka and Negeri Sembilan; (iii) Identify the level of technical teachers' attitudes towards the use of information and communication technology in teaching and learning in technical secondary school in Johor, Malacca and Negeri Sembilan; (iv) Identify the level of confidence in the technical teachers to use information and communication technology in teaching and learning in technical secondary school in Johor, Malacca and Negeri Sembilan; (v) Identify the level of readiness of knowledge, skills, attitudes and technical teachers the confidence to use technology information and communication technology in teaching and learning in technical schools in Johor, Melaka and Negeri Sembilan; and (vi) Identify whether there is a significant difference between technical level readiness on the use of technology information and communication technology in teaching and learning by computer ownership.

### 2.1 Methodology

This study is a descriptive study which employed quantitative methods that related with preparation of using Information Communication and Technology in teaching and learning. The population of this study consists of 120 technical teachers, currently serving in nine selected technical schools from Johor, Melaka and Negeri Sembilan. Based on Morgan and Krejcie (1970), the minimum requirement for sample selection is 92 technical teachers. Simple random sampling was employed for the purposed of this study and the respondent have the equal chance to be selected. Instrument of this study was modified inventory from Information Technology Knowledge and Skills Diagnostic Tool (2003) and Computer Attitude Scale. The mean score in determining the level of knowledge, skill, attitude and confidence is based on Azizi. et.al (2007), high 3.68 – 5.00, moderate 2.34 – 3.67, low 1.00 – 2.33.

The reliability test that is used to all scale is to obtain the value of Cronbach's Alpha for internal consistency. This is done from the pilot study. Pilot test have been done to determine consistency of questionnaire [13] after analysis, the Alpha Cronbach of the questionnaire was 0.9885. This shows that the questionnaire item that was built with high reliability and can be used for actual research.

### 2.2 Procedure of Data Collection

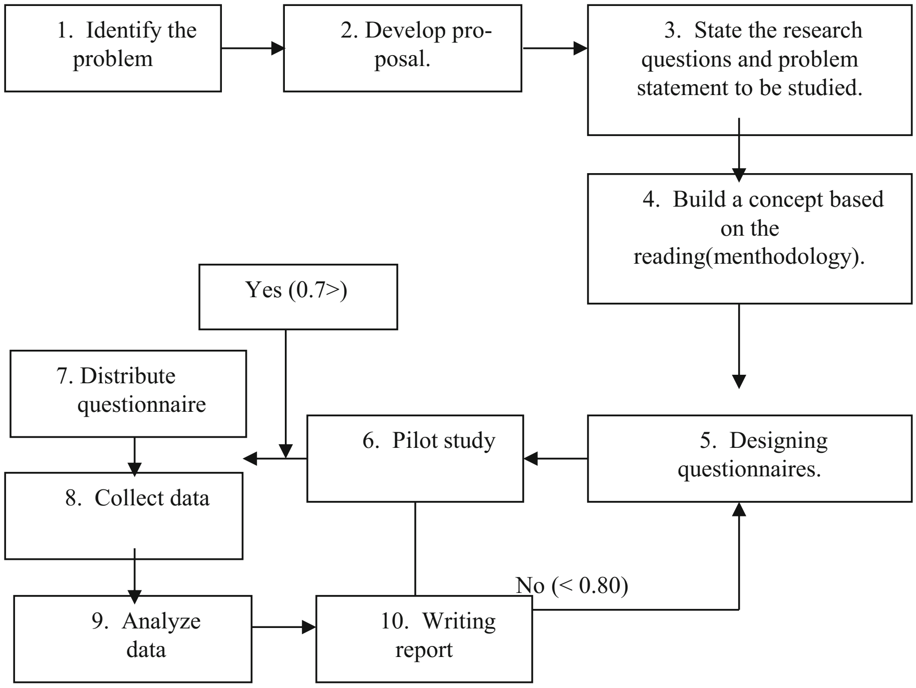


**Fig. 1.** Procedure of Data Collection

## 3 Result

### 3.1 Analysis Preparation of Respondent towards Using Information Communication and Technology in Teaching and Learning

Analysis finding of preparation of using Information Communication and Technology in teaching and learning is categorized as below.

Result in table 1 shows that teacher's knowledge in ICT in "word processing" recorded is the highest mean score of 4.1763 with standard deviation of 0.79821. For the rest of the elements, the mean score are as follows: Electronic Interface (3.5679), Database (3.3261), Presentation (3.0127), Internet (3.0514) and Communication technology (3.8891). As a whole the total mean score is 3.5039 with standard deviation of 0.76642 which has shown that the knowledge that has been mastered by the respondents is at moderate level.

**Table 1.** Respondents' Distribution according to percentage, mean and standard deviation of Information Technology and communications – Knowledge (n=92)

| **NO.** | **Elements** | **MEAN** | **SD** |
|---|---|---|---|
| 1. | Word processing | 4.1763 | 0.79821 |
| 2. | Electronic Interface | 3.5679 | 0.99784 |
| 3. | Database | 3.3261 | 1.05026 |
| 4. | Presentation | 3.0127 | 1.00676 |
| 5. | Internet | 3.0514 | 0.89655 |
| 6. | Communication technology | 3.8891 | 0.81880 |

Total mean score = 3.5039 Standard Deviation = 0.76642.

**Table 2.** Respondents' Distribution according to percentage, mean and standard deviation of Information Technology and communications – Skill

| **No.** | **Elements** | **Mean** | **Sd** |
|---|---|---|---|
| 1. | Word processing | 4.0509 | 0.73442 |
| 2. | Electronic Interface | 3.3293 | 0.99206 |
| 3. | Database | 3.0641 | 1.15021 |
| 4. | Presentation | 3.3646 | 1.08670 |
| 5. | Internet | 2.7746 | 1.06840 |
| 6. | Communication Technology | 3.6250 | 0.90197 |

**Total mean score = 3.3681 Standard Deviation = 0.86562.**

As shows in table 2, the teacher's skill in Information technology and Communications in "word processing" recorded the highest mean score of 4.0509 with standard deviation of 0.79821. For the rest of the elements, the mean score are as followed Electronic Interface (3.3293), Database (3.0641), Presentation (3.0127), Internet (2.7746) and Communication technology (3.6250). The total mean score is 3.3681 with standard deviation of 0.86562 which has shown that the knowledge that has been mastered by the respondents is at moderate level.

**Total mean score = 4.1431 Standard Deviation = 0.53142**
SD = Strongly Disagree, D = Disagree, UD = Undecided, A = Agree, SA = Strongly Agree

In table3, the attitude of teachers in Information Communication Technology (ICT) in Teaching and Learning (T&L) shows that "Information Communication and Technology can be applied in teaching" and "Information Communication and Technology can be applied in learning" recorded the highest mean score of 4.33 (43.5percent-strongly agree, 46.6 percent-agree, 8.7percent-undecided, 1.1percent- disagree and

**Table 3.** Respondents' Distribution according to percentage, mean and standard deviation of the Attitude of Information Communication Technology (ICT) in Teaching and Learning (T&L) n =92)

| **Statement** | Sd | D | Ud | A | Sa | **Mean** | **Sd** |
|---|---|---|---|---|---|---|---|
| ICT is a common thing around me. | 1.1 | 4.3 | 5.4 | 66.4 | 22.8 | 4.05 | 0.747 |
| ICT helps preventing errors in teaching and learning once implemented | 1.1 | 9.8 | 18.4 | 62.0 | 8.7 | 3.67 | 0.813 |
| ICT boost T&L processes | 0 | 2.2 | 9.8 | 53.3 | 34.8 | 4.21 | 0.704 |
| ICT can be used in teaching | 0 | 1.1 | 8.7 | 46.6 | 43.5 | 4.33 | 0.681 |
| ICT can be used in learning | 0 | 1.1 | 8.7 | 46.6 | 43.5 | 4.33 | 0.681 |
| Having a personal laptop helps me in teaching and learning | 2.2 | 2.2 | 7.6 | 42.3 | 45.7 | 4.27 | 0.866 |

**Table 4.** Respondents' Distribution according to percentage, mean and standard deviation of Confidence of the utilization of the Information Communication Technology (ICT) in Teaching and Learning (T&L) (n =92)

| **Statement** | SD | D | UD | A | SA | **Mean** | **Sd** |
|---|---|---|---|---|---|---|---|
| I am convince can do better with ICT | 0 | 4.3 | 8.7 | 57.7 | 29.3 | 4.12 | 0.739 |
| I am an individual that can utilize ICT in T&L superbly | 1.1 | 4.3 | 20.7 | 53.2 | 20.7 | 3.88 | 0.823 |
| I am convince in learning to utilize ICT in my T&L | 0 | 5.4 | 13.0 | 49.0 | 32.6 | 4.09 | 0.821 |
| I felt that the use of ICT does not burden me in my T&L | 0 | 6.5 | 12.0 | 47.8 | 33.7 | 4.09 | 0.847 |
| I am convince will be given great marks if ICT test is given to me | 3.3 | 6.5 | 30.4 | 39.1 | 20.7 | 3.67 | 0.985 |
| I am able to conduct ICT program | 3.3 | 20.7 | 22.8 | 41.2 | 12.0 | 3.38 | 1.047 |
| I am capable of using ICT in T&L | 2.2 | 2.2 | 15.2 | 59.6 | 20.8 | 3.95 | 0.803 |

none-strongly disagree). For the statement, "ICT helps preventing errors in teaching and learning once implemented" recorded the lowest mean score of 3.67 (8.7percent-strongly agree, 62percent- agree, 18.4percent-undecided, 9.8percent-disagree and 1.1percent-strongly disagree). The total mean score is 4.1431 with standard deviation of 0.53142 has shown that the respondents' mastery level towards ICT in teaching and learning is at high level.

**Total mean score = 3.8820 Standard Deviation = 0.70069**
SD = Strongly Disagree, D = Disagree, UD = Undecided, A = Agree, SA = Strongly agree

Table 4 shows the respondent's distribution according to percentage, mean and standard deviation of the confidence of the utilization of the Information Communication Technology (ICT) in Teaching and Learning (T&L). The result shows that "Iam convince to do better with the ICT" recorded the highest mean score of 4.12 (29.3percent-strongly agree, 57.7percent-agree, 8.7percent-undecided, 4.3percent-disagree and none-strongly disagree). For the statement, "ICT helps preventing errors in teaching and learning once implemented" recorded the lowest mean score of 3.67 (8.7percent-strongly agree, 62percent- agree, 18.4percent-undecided, 9.8percent-disagree and 1.1percent-strongly disagree). We can see that the total mean score is 4.1431 with standard deviation of 0.53142 has shown that the respondents' confidence level towards ICT in teaching and learning is high.

**Table 5.** Respondent's Preparation Level of Using Information Communication and Technology in Teaching and Learning

| No. | Aspect | Mean | Level |
|---|---|---|---|
| 1. | Knowledge | 3.5039 | Moderate |
| 2. | Skills | 3.3681 | Moderate |
| 3. | Attitude | 4.1431 | High |
| 4. | Confidence | 3.8820 | High |

**Overall Score Mean = 3.7243 Preparation Level = High**

Table 5 showed that the respondent's preparation level using Information Communication and Technology in teaching and learning. The most dominant aspect was attitude where the mean was 4.1431. This is followed by aspects of confidence (3.8820), knowledge (3.5039) and skills (3.3681). Overall, analysis showed that preparation level of using Information Communication and Technology was at high level where the overall score mean was 3.7243.

Table 6 shows that there is significant difference among the preparation level of using Information Communication and Technology in Teaching and Learning with respondent's computer ownership is $p< 0.001$ which less than 0.05. Hence, this hypothesis noel was rejected. Finding of the study also showed that value t was 3.608.

**Table 6.** Difference Preparation Level of Using Information Communication and Technology in Teaching and Learning with Respondent's Computer Ownership

**(n =92)**

| | **No.** | **Mean** | **SD** | **df** | **t** | **Significant** |
|---|---|---|---|---|---|---|
| Yes | 84 | 3.7897 | 0.54376 | 90 | 3.608 | 0.001 |
| No | 8 | 3.0369 | 0.76341 | | | |

* Significant at the .05 level.

## 4 Discussion

The finding of this study shows that respondent are only at moderate level in mastering knowledge and skills of Information Communication and Technology. Respondent's Information Communication and Technology knowledge and skills were at a same level because both the aspects were related to each other. It can be proved by the mean of each aspect. In aspect of attitude towards using Information Communication and Technology in teaching and learning, it is found that respondents have positive attitude. Respondents also have high level of confidence towards using Information Communication and Technology in teaching and learning. [14] Commented, mastering in knowledge of Information Communication and Technology already became necessities in future for teachers. Individual who acquired knowledge in computer and practices positive attitude will be a responsible teachers in using compute. This teachers will show effort to evaluate, elect and using various computer application; able to use hardware and software with computer application; and expert in computer operating through the programming [15].

Even though knowledge and skills in Information Communication and Technology of the respondents is moderate level, but it can be improved through education programs and training in future. This is because positive attitude and confidence were the aspect which can influence preparation of individual towards using Information Communication and Technology in teaching and learning. Overall, the preparation of technical teachers towards using Information Communication and Technology in teaching and learning is at high level. Attitude was important element in education field. Teachers who have positive attitude can make teaching and learning process more effective. This is because they can motivate students learning in the class. As [16] comment, positive attitude was essential to motivate someone to learn something. If the teachers have negative attitude towards using Information Communication and Technology in teaching and learning, it will become barrier to teaching and learning process.

As [17] pointed out, self-confidence was expecting which may be will achieve by someone in certain situation. Positive attitude and self-confidence as a basic support need to achieve objective [18] As [13] commented, confidence of the teachers also will influence them in using Information Communication and Technology in teaching and learning.

The findings of this study also showed that the preparation level of using Information Communication and Technology in teaching and learning have significant difference with frequency in using computer and computer ownership. However, there were no significant difference among the level of preparation with computer ownership in using computer and participating Information Communication and Technology course.

## 5 Conclusion

The finding of this study showed that respondents have high level in preparation of using Information Communication and Technology in teaching and learning. In opinion of researcher, it was normal for the technical teachers that should have high level in preparation. This is because teachers were the main media in planning teaching and learning activities with using Information Communication and Technology to improve effectiveness of teaching and learning process. Hence, teachers should always prepare themselves to integrate Information Communication and Technology in teaching and learning to make their students intelligent in Information Communication and Technology.

## References

1. Kerajaan Malaysia. Perkhidmatan Awam Malaysia Menuju Alaf Baru. Percetakan Nasional Malaysia Berhad, Kuala Lumpur (1999)
2. Ding, C.M.: Zaman Maklumat. Dewan Budaya 7(20), 42–43 (1998)
3. Pusat Perkembangan Kurikulum Kementerian Pelajaran Malaysia. Penggunaan Teknologi Maklumat dan Komunikasi (ICT) dalam Pengajaran dan Pembelajaran. Pusat Perkembangan Kurikulum, Kuala Lumpur (2001b)
4. Hamidah Baba Peranan Guru Di Dalam Era Teknologi Maklumat Dan Komunikasi ICT). In: Saifuddeen, (ed.) Memahami Islam Melalui Teknologi Maklumat & Komunikasi, pp. 91–105. Institut Kefahaman Islam Malaysia, Kuala Lumpur (2000)
5. Razak, A., Zakree, M., Faidzul, M., Hafiz, Zakaria, M.S., Yazrina, Zamri, M.: Teknologi Maklumat & Komunikasi. Edisi Kedua. McGraw Hill Sdn. Bhd., Kuala Lumpur (2004)
6. Bakar, A.: Medium Persaudaraan Islam Alaf Baru. In: Saifuddeen, S.M., Salleh, S.M. (eds.) Memahami Islam Melalui Teknologi Maklumat & Komunikasi, pp. 107–120. Institut Kefahaman Islam Malaysia, Kuala Lumpur (2000)
7. Laporan Parlimen Kesebelas: Sejauh Manakah Kejayaan Program Meng"IT"kan Sekolah-Sekolah (2005), `http://www.malaysia-today.net/parlimen/2005_05_06_MT_parlimen_archieve.htm`
8. Gan, S.L.: IT & Education In Malaysia: Problems, Issues and Challenges. Longman, Kuala Lumpur (2000)
9. Rohani Kesediaan Guru Sains dan Matematik Terhadap Perlaksanaan Pengajaran Dalam Bahasa Inggeris. Institut Penyelidikan Matematik, Universiti Putra Malaysia, Serdang (2003)
10. Mokhtar, M.: Aplikasi Teknologi Maklumat dan Komunikas (ICT) Dalam Meningkat-kan Pembelajaran dan Pengajaran Oleh Guru Teknikal (Awam, Elektrik dan Jentera) Seko-lah Menengah Teknik Di Negeri Selangor. Tesis Sarjana Muda. Universiti Teknologi Malaysia (2004)
11. Cox, M., Webb, M., Abbott, C., Blakeley, B., Beauchamp, T., Rhodes, V.: ICT and Pedagogy. DfES Publications, London (2003)

12. Williams, D., Wilson, K., Richardson, A., Tuson, J., Coles, L.: Teachers' ICT Skills and Knowledge Needs. The Robert Gordon University, Scotland (1998)
13. Yahaya, A., Hashim, S., Ramli, J., Boon, Y., Hamdan, A.R.: Menguasai Penyelidikan Dalam pendidikan; teori, analisis dan Interpretasi data. PTS Profesional Publishing Sdn. Bhd., Kuala Lumpur (2007)
14. Mohamad, K.A.: Pendekatan Pengukuhan Ukhwah Melalui Teknologi Maklumat dan Komunikasi (ICT). In: Saifuddeen, S.M. (ed.) Memahami Islam Melalui Teknologi Maklumat & Komunikasi, pp. 1–6. Institut Kefahaman Islam Malaysia, Kuala Lumpur (2000)
15. Simonson, M.R., Maurer, M., Montag-Torardi, M., Whitaker, M.: Development of a Standardized Test of Computer Literacy and a Complex Anxiety Index. Journal of Educational Computing Research 3(2), 231–247 (1987)
16. Zulkifli dan Raja Maznah Attitudes, Knowledge and Previous Computer Experience of Teacher Trainees in the Diploma of Education Programme at the University of Malaya. Jurnal Pendidikan, 161–180 (1994)
17. Tom, G.S.: You Can Choose To Be Happy: "Rise Above" Anxiety, Anger and Depression. California State University, Long Beach (1999)
18. Robiah, Juhana, dan Nor Sakinah: Pembudayaan Teknologi Maklumat dan Komunikasi Di Kalangan Pelajar Melayu Dalam Arus Globalisasi. Prosiding Seminar Pembinaan Keupayaan Melayu dalam Arus Perdana II. Universiti Kebangsaan Malaysia, Bangi (2003)

# The Impact of Social Networking on Behaviour Development among Secondary School Students

AziziYahaya[1], Gooh Mo Lee[2], Halimah Ma'alip[3], Dayang Ku Zainab Pg Hj Tuah[4], Zeliha Mohammad Ali[5], Kamariah Hj Kasah[6], and Zarifa Binti Mohd Sis[2]

[1] Sultan Hassanah Bolkiah Instituit of Education, Universiti Brunei Darussalam
`profazizi_yahaya@yahoo.com`
[2] Sekolah Kebangsaan Convent Infant Jesus 1, Melaka, Malaysia
`moleegooh@gmail.com`
[3] Languange Academic, Universiti Teknologi Malaysia Johor
`m-halima@utm.my`
[4] Kolej Univerisiti Perguruan Ugama Seri Begawan, Brunei Darussalam
`dk.zainab988@yahoo.com`
[5] Maktab Sultan Omar Ali Saifuddien, Brunei Darussalam
`lihaali6.za@gmail.com`
[6] Sekolah Rendah Pantai Berakas, Brunei Darussalam
`ariel.193@hotmail.com`

**Abstract.** The aim of this study is to identify the impact of social networking on behavior development. This study also focuses on four factors; loneliness, social anxiety, social influences and addiction. The dependent variables of this study are the self-esteem, social skills and learning behaviours. Respondents were 220 students from seven schools in Johor Bahru district. The data were analyzed using SPSS version 17. Descriptive statistics show that social influence is the dominant factor. The study also found that the dominant effect is self-esteem. Instant Messenger is an online social site that contributes to self-esteem, social skills and learning behaviours. Inferential statistics show that there is a significant impact between social networking sites ($\beta$= .50, p <.05) and loneliness ($\beta$= .31, p<.05)towards the effects of social websites. In addition, there is a significant impact between loneliness ($\beta$= .37, p <.05) towards self-esteem. There is also a significant impact of social networking sites($\beta$= .33 ,p <.05) and addiction ($\beta$= -.23, p <.05) towards social skills. Social networking sites ($\beta$= .55, p <.05) and loneliness ($\beta$= .23, p <.05) also brings a significant impact on learning behaviour. Apart from that, addiction ($\beta$= -.30 ,p<.05), social influence($\beta$= -.30 p <.05) and loneliness ($\beta$= .16 , p <.05) gives impact towards online social networking.

**Keywords:** Social Networking, Self-Esteem, Social Skills and Learning Behaviour.

## 1 Introduction

Social Networking has gathered many Internet users and by March 2009, the number of Facebook users in Malaysia has surpassed the one million mark [1]. This shows

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_19

that social network users in Malaysia is increasing and thus will be increasing more in subsequent years. The advantages of social network have gathered people from around the world to be a part of it. As being part of one's daily routine, there are factors that influence people to use the online social network. Factors that will be discussed are loneliness, social anxiety, social influence and addiction.

Loneliness is the unpleasant experience that occurs when a person's network of social relations is deficient in some important way, either quantitatively or qualitatively [2] This definition have three general points to be noticed, first, loneliness results from deficiencies in the person's social relations; second, loneliness is a subjective phenomenon (it is not necessarily synonymous with objective isolation, so that people can be alone without being lonely); third, loneliness is unpleasant and distressing. Further evidence for the link between loneliness and social problems has been reported by [3] surveys of American adolescents, where they found that loneliness was associated with poor grades, expulsion from school, running away from home, and engaging in delinquent acts such as theft, gambling and vandalism.

Social anxiety is one of the common psychological problems that could be faced by human being. Social anxiety disorder (SAD) is considered to be the third most prevalent psychiatric disorder [4] Though it can start at any age, some individuals with this disorder is reported of being very shy throughout their life, but many says that the problem was developed at the early of adolescence or in adulthood [5].

The effects that will be discussed in this study are self-esteem, social skills and learning behaviour. Based on [6], healthy self-esteem or positive self-regard is about feeling competent and feeling of love or being 'appereciated'. It involves the evaluation of the self-concept and is often unrelated to our true abilities. A new study suggests that spending time online with hundreds of friends, posting 'witty' status updates and all the unflattering photos of oneself might help increase self-esteem [7].

The social network has been a routine for most who have easy access to the Internet. Besides studying, social networking has become a part of their daily routine. The questions to be addressed are regarding how students divide their time and how well do they learn from the Internet?

## 2 Objectives and Methodology

The specific objectives of this study is to explore and identify the following information: (i) to identify the most dominant factors on social anxiety, social influence, loneliness and addiction that contributes to online social networking; (ii) to identify the most dominant effect on self-esteem and learning behaviour towards online social networking; (iii) to identify the most dominant social networking sites that contributes to self-esteem and learning behavior; and (iv) to identify the impact of loneliness, social anxiety, influence and addiction towards social networking sites

### 2.1 Method

Methodology may be referred to a scientific method for the summary of approaches and methods applied in a particular study. Quantitative method were employed in this study to ensure all the items in the questionnaire will be easily analyzed by using SPSS

Population is the set of all the individuals of interest in a particular study. The population for this research consits of students from seven secondary school in Johor Bahru district. The term "population" is used in statistics to represent all possible measurements or outcomes that are of interest to us in a particular study. A sample is a part of the population of interest, a sub-collection selected from a population.

The population determines the studied problem area and determines the data and information should be collected and analyzed [8]. Meanwhile, according to [8], the sample is a small group of elements taken from a sampling frame. The sample is said to represent the population in any vote and is important to make inferences about the population based on the information contained in it. The research populations in question are secondary four students which are selected from five schools in Johor Bahru.

## 3 Research Sample

The population is defined as a set of elements or elements that become the object of research [8]. Elements of this population are usually the unit of analysis. The population can also be interpreted as a whole unit of analysis characteristics that will be tested. The unit of analysis is the unit to be studied. Sample is part of a population (or a representative portion of the population studied). The samples were part of the population is taken as a source of data and can represent the whole population. Samples are some of the characteristics possessed by the population. The advantage in using a sample that is: facilitate researchers, research more efficient, more thorough and meticulous in data collection, as well as more effective research

Based on [8], sample size is important as it symbolizes the strength of the result of the research. Respondent who are chosen is selected randomly from seven schools with a total of 220 respondents.

Determination of sample size is based on a method applied according to the chart by Krejcie and Morgan (1970). The type of sample selected is based on simple random sampling [9].

The subsample population of this study is the 220 students from the seven secondary schools in Johor Bahru, Malaysia.

### 3.1 Instruments

The instrument used in this research is in the form of a questionnaire. Some of the item used in this study are adapted from items in the Internet Addiction Test, Social Interaction Anxiety Scale, R-UCLA Loneliness Scale, Rosenberg's Self-Esteem Scale (1965) while the rest of the items used were developed by the researcher [10].

### 3.2 Reliability

A pilot study was held in SMK Seri Perling involving 30 students. The students are asked to answer the questionnaire provided. Data obtained from this pilot study were analyzed using Statistical Package for Social Science (SPSS) Version 17.0 for Windows. Alpha coefficient value obtained is 0.875. According to [9] the value between 0.71 to 0.99 is the optimal value. This shows that the questionnaire item that was built have a high reliability and can be used for real study.

## 4 Results

This section will discuss the results and analysis of the data obtained from respondents of the selected schools in Johor Bahru. The analysis of data is explained through descriptive and inferential statistics which was analyzed using *Statistical Packages for Social Science (SPSS) Version 17.0 for Windows*. The statistical methods employed were mean, standard deviation and multiple regressions.

### 4.1 Objective (i): Most Dominant Online Social Networking Factors

The factors involved consist of loneliness, social anxiety, social influence and addiction. The findings are shown in Table 1. Eight questions are applicable for loneliness. Both social anxiety and social influence have six questions and ten questions for addiction.

**Table 1.** Distribution of Mean and Standard Deviation for Overall Factors towards Online Social Networking

| **Factors Towards Online Social Networking** | **Mean** | **Standard Deviation** |
|---|---|---|
| Loneliness | 2.4431 | 1.1099 |
| Social Anxiety | 2.5237 | 1.0967 |
| Social Influence | 3.0739 | 1.2590 |
| Addiction | 2.9063 | 1.2230 |
| **Overall** | **2.7367** | **1.1721** |

Table 1 shows the distribution of mean and standard deviation for the overall factors towards online social networking. For the factor loneliness, the acquired mean is 2.4431 (SD= 1.1099) while the recorded mean for social anxiety is 2.5237 (SD=1.0967). For the factor of addiction the mean is 2.9063 (SD=1.2230). As shown in Table 1 above, social influence generates the highest mean value at 3.0739 (SD=1.2590). This indicates that this statement is the most dominant factors towards online social networking.

### 4.2 Objective (ii): Most Dominant Online Social Networking Effects

The effects of this study consist of self-esteem, social skills and learning behaviour. The findings are shown in Table 2 and Table 3.

Self-esteem, social skills and learning behavior consist of ten items. The items were measured by Likert scale; strongly agree, agree, not sure, disagree and strongly disagree. Only means and standard deviation were employed to explain the result of this variables (self-esteem, social skills and learning behavior).

**Table 2.** Distribution of Mean and Standard Deviation for Overall Online Social Networking Effects

| **Online Social Networking Effects** | **Mean** | **Standard Deviation** |
|---|---|---|
| Self-Esteem | 2.8076 | 1.0277 |
| Social Skills | 2.7116 | 1.0643 |
| Learning Behaviour | 2.0604 | 1.1160 |
| **Overall** | **2.5265** | **1.0693** |

Table 2 above shows the distribution of mean and standard deviation for online social networking effects. It was discovered that the mean for self-esteem is the highest with 2.8076 and standard deviation of 1.0277. This is a valid indicator and that self-esteem is the most dominant effects of online social networking. Meanwhile, mean for learning behaviour is the lowest, which is 2.0604 and standard deviation of1.1160. Apart from that, social skills generate a mean of 2.7116 and standard deviation of 1.0643. As a result of this, the overall mean for online social networking effects is 2.5265 and standard deviation of 1.107.

### 4.3 Objective (iii): Most Dominant Online Social Networking Sites

The networking sites in this question are Facebook, Blog, Twitter, Instant Messenger and YouTube. Six items are given for each site. For each item, a 5-Likert scale point which was used namely strongly agree, agree, not sure, disagree and strongly disagree to quantify the respondent's answers.

Table 3 shows the distribution of mean and standard deviation for overall social networking sites that contribute towards self-esteem, social skills and learning behaviour. Facebook generates the lowest mean at 2.6398 (SD=1.1142). Meanwhile, Twitter produces a mean score of 3.0620 (SD=1.1808), Blogging with a mean of 3.1454 (SD=1.1514) and YouTube with a mean of 2.7495 (SD=1.1220). Instant Messenger generates the highest mean with 3.2206 (SD=1.0891). This indicates that Instant Messenger is the most dominant factor that contribute towards self-esteem, social skills and learning behaviour.

**Table 3.** Distribution of Mean and Standard Deviation for Overall Social Networking Sites that Contributes Towards Self-Esteem, Social Skills and Learning Behaviour.

| **Overall Social networking** | **Mean** | **Standard Deviation** |
|---|---|---|
| Facebook | 2.6398 | 1.1142 |
| Blogging | 3.1454 | 1.1514 |
| Twitter | 3.0620 | 1.1808 |
| Instant Messenger | 3.2206 | 1.0891 |
| YouTube | 2.7495 | 1.1220 |
| **Overall** | **2.9634** | **1.1315** |

## 4.4 Objective (viii) Impact of Loneliness, Social Anxiety, Social Influence and Addiction towards Social Networking Sites

The result from Table 4 is used to answer the research question regarding "What is the impact of loneliness, social anxiety, influence, and addiction towards social networking sites?" The results from the analysis shows that for the population (n=175), three variables; addiction, social influence and loneliness are the predictors for social networking sites. The other variables are not predicting factors for social networking sites.

The correlation between criterion variable and regressor variable social influence is 0.30, correlation between criterion variable and linear combination between addiction and social influence is 0.34 while correlation between criterion variable and linear combination between addiction, social influence and loneliness is 0.37. The ANOVA results shows that there is a significant relationship between the criterion variable and regressor variable with p<.05. Apart from that, Table 5 shows that the Beta value for addiction is -0.237, social influence with -0.190 and loneliness with 0.159.

**Table 4.** Results of Linear Regression Analysis for the impact of loneliness, social anxiety, Social Influence and addiction towards social networking sites.

| **Model** | **R** | **R Square** | **Standard Error** | **R Square Change** | **F** | **Sig.** | **Standard Coefficients Beta** | **t** | **Sig** |
|---|---|---|---|---|---|---|---|---|---|
| 1 | .299a | .089 | 2.171 | .089 | 16.938 | .000a | -.299 | 16.387<br>-4.116 | .000<br>.000 |
| 2 | .335b | .112 | 2.149 | .023 | 10.842 | .000b | -.206<br>-.177 | 15.839<br>-2.442<br>-2.100 | .000<br>.016<br>.037 |
| 3 | .368c | .136 | 2.127 | .024 | 8.946 | .000c | -.237<br>-.190<br>.159 | 13.898<br>-2.779<br>-2.268<br>2.166 | .000<br>.006<br>.025<br>.032 |

a.Predictors: (Constant), addiction
b.Predictors: (Constant), addiction, social influence
c Predictors: (Constant), addiction, social influence, loneliness
d.Dependent variable: social networking sites

Model 1=[(1,173)]=16.938 , p<0.005
Model 2 =[(2 ,172 ) = 10.842, p<.05 ].
Model 3= [(3,171) =8.946, p<.05

**Table 5.** Results of Predictors and Beta for the impact of loneliness, social anxiety, Social Influence and addiction towards social networking sites

| Predictors | Beta (β) | p |
|---|---|---|
| Addiction | -.237 | p=0.006 |
| Social Influence | -.190 | p=0.025 |
| Loneliness | .159 | p=0.032 |

As shows in table 5, Addiction significantly, [(1, 173) = 16.938, p<. 05] contributes to 8.9 percent of the variance ($R^2$=. 089) to social networking sites. In other word, when addiction factors increase by one unit, then social networking sites score will increase by 8.9 percent. It can be seen that addiction is the main factor in social networking sites. The combination of addiction (β= -.30 ,p <.05) and social influence (β= -.18, p <.05) add (11.2 – 8.9) = 2.3 percent to the variance ($R^2$=.112) in the criterion variable social skills with significance of [(2 ,172 ) = 10.842, p<.05 ].The combination of social influence, (β= -.30 p <.05) addiction (β= -.21, p <.05) and loneliness, (β= .16 , p <.05) add (11.2 – 13.6) = 2.4 percent to the variance ($R^2$=.112) in the criterion variable social networking sites with significance of [(3,171) =8.946, p<.05 ]. Therefore addiction, social influence and loneliness are the predictor variables for the social networking sites.

## 5 Discussion

Social influence has the greatest impact especially towards the adolescent. Social Influence may come from peers, family or the media. Influence may not always be negative, however there are also influence that are positive. Family however, are the one whom should show good behaviour towards their family member especially to the growing ups and also adolescent. Unfortunately, there are recent report showing that the children's home and family lives have long been considered a primary environmental context influencing their psychological development and behaviour[11].

Research from the scientists at Cornell University in Ithaca, New York show that spending time online with the hundreds of friends, posting "witty" status updates and all the unflattering photos of oneself might help increase self-esteem [12]. This also applies to the field of psychology whereby American psychologist Abraham Maslow, included self-esteem in his hierarchy of needs. This is where online social networking is a platform helps some people whom consider themselves being invaluable and unworthy of being loved find their self-esteem.

Internet Messenger may give impact to the users. Internet Messenger users perceive Internet Messenger use to be easier, more useful and favorable. Furthermore, they have greater intent to use Internet Messenger and a greater need for affiliation and subjective norms [12].

Online social networking services have recently emerged as popular ways to share personal information and communicate with friends. The study has found out that loneliness gives the major impact from the entire five variables that were tested; effect, self-esteem, social skills, learning behaviour and social networking sites. Loneliness gives impact towards the four variables, which are effects, social skills, learning behaviour and social networking sites. As such, this shows that loneliness is one of the effects of online social networking. In addition, it also gives impact towards social networking. In fact, previous research [13] has shown that, although Internet use is positively associated with initial loneliness, overtime it leads to a decrease in loneliness.

Apart from that, social networking sites also give an impact to three variables, such as effects, social skills and learning behaviour. It also gives some additional impact to addiction, social influence and loneliness. Thus this proves that, students' social networking, especially when the networking increasingly shifts to online, is more likely to be self-initiated learning, in which individuals create a system of information and support by building and nurturing personal links [14].

## 6 Conclusion

Online Social Networking has become a trend among for every student. It is a must for adolescent nowadays to have at least a social networking account. As the source of the internet is widely provided, namely at house, school, on mobile phone, cafés and in addition to cheap access to the Internet, it is no doubt that all everyone are able to access to the World Wide Web.

This study is conducted in the hopes that the findings would be beneficial for the society, starting from family up to the ministry. Parents should not leave their children unattended when their children are on the World Wide Web. It is important to control the children's activities so that the online social networking may contribute towards a positive learning behaviour.

Teachers and school may implement social networking as part of the student's learning environment. Teachers should be more creative in adapting the social network into the classroom. Conventional learning is no longer attracting students to learn but teachers should instead use new method in teaching such as disseminating learning materials through social networking sites in a creative and more appealing way.

## References

1. Hui, L.Y.: Snapshot of Social Networking in Malaysia (July 28, 2009), `http://www.greyreview.com/2009/07/28/snapshot-of-social-networking-in-Malaysia/`
2. Peplau, Perlman: Toward a Social Psychology of Loneliness Personal Relation-ships in Disorder. Academic Press, London (1982)
3. Brenna, Auslander, N.: Adolescent Loneliness: An Explotary Study of Social and Psychological Pre-disposition and Theory. Behavioral Research Institute, Boulder (1979)
4. Brunello, N., Den Boer, J.A., Judd, L.L., Kasper, S., Kelsey, J.E., Lader, M., Lecrubier, Y., Lepine, J.P., Lydiard, R.B., Mendlewicz, J., Montgomery, S.A., Racagni, G., Stein, M.B., Wittchen, H.U.: Social Phobia: diagnosis and epidemiology, neurobiology and pharmacology, comorbidity and treatment. Journal of Affective Disordes 60, 61–74 (2000)
5. Psychatric Malaysia. Social Anxiety Disorder (July 6, 2006), http://www.psychiatry-malaysia.org/article.php?aid=86 (retrieved November 28, 2011)
6. Plummer, D.: Helping adolescents and adults to build self-esteem. Jessica Kingsley Publishers (2005)
7. CNN. Study: Facebook Helps Your Self-Esteem (March 1, 2011), `http://www.aarp.org/technology/social-media/news-03-2011/study_facebook_helps_your_self_esteem.html` (retrieved June 5, 2011)
8. Yahaya, A., Hashim, S., Ramli, J., Boon, Y., Hamdan, A.R.: Menguasai Penyelidikan Dalam Pendidikan: Teori, Analisis& Interpretasi Data. PTS Professional Publishing Sdn. Bhd, Selangor (2007)
9. Konting, M.M.: Kaedah Penyelidikan Pendidikan, c.2. Dewan Bahasa dan Pustaka, Kuala Lumpur (1993)
10. Rosenberg, M.: Society and the adolescent self-image. Princeton University Press, Princeton (1965)
11. Belsky, J.: Classroom Composition, Child-Care History and Social Development: Are Child-Care Effects Disappearing or Spreading? Social Development 18, 230–238 (2008)
12. Rusell, J.: Is Malaysia world's biggest social network addict? (October 10, 2011), `http://asiancorrespondent.com/41329/is-malaysia-the-worlds-biggest-social-networker/` (retrieved on December 5, 2011)
13. Donghun, C., SooNum, C.: An analysis of the variables predicting instant messenger use. New Media Society 9 (2007)
14. Kraut, R., Kiesler, S., Boneva, B., Cummings, J.N., Helgeson, V., Crawford, A.M.: Internet paradox revisited. Journal of Social Issues 58, 49–74 (2002)

# KIHECT$^{©}$: Reliability of Hand-Eye Coordination among Rugby Players Using Consumer Depth Camera

Mohd Razmi Rozan[1], Mohd Kufaisal Mohd Sidik[2],
Mohd Shahrizal Sunar[2], and Abdul Hafidz Omar[1]

[1] Sport Innovation & Technology Centre (SITC),
Universiti Teknologi Malaysia, Malaysia
`mohdrazmidaniel89@gmail.com, aho@utm.my`
[2] Media and Game Innovation Centre of Excellence (MaGIC-X),
Universiti Teknologi Malaysia, Malaysia
`kufaisal@magicx.my, shahrizal@utm.my`

**Abstract.** Rugby is a growing sport in the whole world. For athletes who want to achieve a high level of potential, such as hand coordination skills and speed of reaction (sport) both are expected to be in good performance. The main problem in many sports is that there is no test for the reliability of the developed test tools. Thus, the aim of this study is to determine the reliability of the tests, so the research was conducted twice with an interval of 6-7 days (test-retest). A total of 33 rugby players (16-18 years of age) have participated in the study. The analysis includes one motor tasks performed only with the use of KIHECT$^{©}$. Thus, this study was conducted on a differentiated time and score through two methods of evaluation test: automatically and manually in order to obtain the level of reliability of the tools. The hardware and software that were used in this research are Kinect camera 360, windows SDK and computer or monitor use by CPU systems. The average values of the resultant p-values ($<0.005$) showed that the instrument is reliable and practical to be used in research area. A test on KIHECT$^{©}$ displayed sufficient reliability to see the difference in score data using automatic method (evaluation score using KIHECT$^{©}$) and manual method (evaluation score of the two observers). Therefore, it can be implemented as a tool to train the hand-eye coordination in rugby games.

**Keywords:** Hand-eye coordination, reliability, reaction time, KIHECT$^{©}$

## 1 Introduction

Rugby union is a full contact team sport which is characterized by recurrent high-intensity dashes and a high grade of physical contact. Claimed rugby union is a high-intensity field based contact team sport which is considered as a New Zealand's national game [1]. Besides that, rugby is becoming a completely skilled sport in 1996, as the expectation from the society and media for the accomplishments of both the team and individual players has been equally overwhelming. Therefore, rugby games demand a high body performance and are definitely competitive, and a lot of research has been extended in rugby union [2].

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_20

In rugby, physical characteristic is an important factor to the improvement of the overall performance. Clearly, the physical character of the player is crucial for the player`s performance, as it will be improved in the games. Specific criteria are expected in order to obtain success is highly dependent on the physical characteristic, particularly the role of a certain position in the rugby [3]. In particular, the essential physical characteristics to be successful will be primarily determined by the role of the specific player position. For instance, outside backs must be able to overcome the opponent's defensive attack in open play, hence are compulsory to be fast and agile [4, 5]. On the other hand, players who have good physical fitness are capable of accomplishing their tasks much more easily during the games. Hence, where the players that have advanced levels of physical fitness may have an edge in the performance of responsibilities in competition [6].

There are a few website providing a free online application for hand-eye coordination test; The Hand-Eye Coordination Test [7] Hand-Eye Coordination Testing [8], both are not using any real equipment's except computer mouse. Thus, there are a few devices blending both application (software) and hardware; Batax by Quotronics [9], The Visagraph II [10], The Wayne Saccadic Fixator [11], and The Acuvision 1000 [12]. Vision in game is paramount to settle on a decision making and reaction in every task. The basis of sport vision incorporates three handling stages comprised of recognition, decision making, and reaction which are the usage of development [13, 14]. In fact, four components are recognized and are perceived to have an incredible effect on an execution to all games are ability execution, fixation, reaction time and decision making [15].

However, vision in the context of this article actually refers to the real player's physical ability to make certain task inside the functionality of the eye as a muscle [16]. Thus, sport vision is one of the critical components in rugby because most of the trainings intend to improve and use the vision skill in terms of passing, catching, tackling, side stepping, player marking, and any other skills in rugby. Sports vision is indeed vital as it does not focus on an exact aspect of the body, but on the body as a whole [17].

## 2 Related Works

KIHECT© or the Kinect Hand Eye Coordination Trainer was created based on "Natural Interaction on Human Body Tracking for Low-Cost Motion Capture for Real-Time Computer Graphics Applications". KIHECT© is a tool that helps the human hand-eye coordination. The results of this study showed that KIHECT© helps in improving human concentration and agility. A number of constraints have been identified in KIHECT©. We use the method of markerless to automatically identify a human hand but the hand position is less accurate due to the noise from the camera and so on. Due to this situation, it should be improved in terms of accuracy in mapping the human hand because what the athletes need is a measure of agility. These things are taken into consideration in making KIHECT© prototype for athletes.

These devices are widely used by researchers because of the reliability of the devices. Nonetheless, there are still disadvantages; heavy, needs an expert to operate and calibrate the device, and high cost that only research center can afford to buy. Therefore, this study will develop a new innovation in technology-based device which is lightweight, can be operated by anybody, less time to calibrate, and low cost. The problems that were mentioned will be solved by the principle of human-computer interaction (HCI). The new way of HCI brought to us is a natural-user interface (NUI). For interaction of NUI, we call it natural interaction (NI) [18, 19] because this interaction is a new concept that goes back to the natural basis. NI is an emerging computer interaction methodology which focuses on human abilities such as touch, vision, voice, motion and higher cognitive functions such as expression, perception and recall. Previously, BATAK LITE was used to measure both hand-eye coordination and visual reaction skill.

Design innovation is known as an idea to a design or project to be developed for specific purposes according to the problems faced [20]. Thus, the advantages of technologies in sports research are acquired to improve skills in sport performance. Hence, the technological growths in notational study have unavoidably insulated those in the practical computing technology environment [21]. The function of the technologies in sport research is how the contribution of the research enhances performance.

BATAK LITE is still lacking in scientific of study to measure the hand-eye coordination skill and reaction time in research yet, because this tool is only for market commercialization while at the same time improving the hand-eye coordination and reaction time. Hence, it is important to investigate a new approach and create more reliable instruments to prove the new findings. Hence, stated that a reliable instrument for research would generate same data from the same subject [22]. Reliability is a measure of consistency over time and over the same samples [23]. Therefore KIHECT© was developed with the purpose that this new instrument is created based on sound research which did not happen in BATAK LITE. KIHECT© is the new tools designed to train the hand-eye coordination (HEC) and visual reaction time (VRT) while it would enhance the performance of athlete's skills. Innovation is known as an idea to a design or project to be developed for specific purposes according to the problems [20].

## 3 Data Acquisition

In this section, we will describe a method for 3D data acquisition from the camera to the data acquisition process and the type of face 3D camera utilized for face point cloud data and color images.

### 3.1 Input Device

In this study, we use the 3D camera that has several components that are included in the 3D camera. 3D camera is known as the Microsoft Kinect (Kinect). Kinect has 5 components: RGB camera (normal color camera), depth camera (3D camera), infrared (IR) laser projector, multi-array microphone and also motorized tilt. All components are shown in Figure 1. IR laser projector is used to emit infrared rays to the scene and reflected back by objects on the beam to the 3D camera. Kinect reading speed is 30Hz or 30 times per second and Kinect is 640x480 pixels resolution for both RGB and 3D camera. Kinect has the ability to read a range of up to ± 9.9 m starting at a distance of ± 500 cm but with Kinect drivers supplied by Microsoft's own Kinect set can only be read between the ranges 1.2 - 3.5m. Kinect was used in this study because it is one of the 3D cameras that low-cost and able to afford by users.

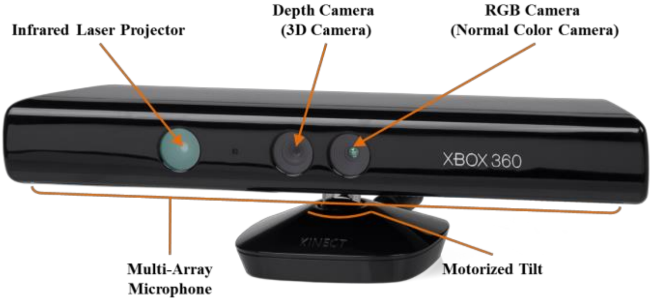


**Fig. 1.** Microsoft Kinect

### 3.2 RGB and Depth Data

For this study, the RGB data is very important to use to detect the human face as RGB color data contains information that can be used to get the current position of the face in the scene camera.

Depth data is also becoming an important element in getting the point cloud data for the process to form a human face avatar is to use the data. It is also important to position parallel with similar images in RGB camera.

Fig. 2 shows the images obtained from the RGB Kinect depth data and data from the same scene. RGB and depth data have no similarities in terms of pixel position due to its 3D camera and the RGB camera which has different position and it is repaired with a little calibration of both the camera.

**Fig. 2.** Depth data (left) and RGB data (right)

### 3.3 KIHECT$^{©}$

KiHECT$^{©}$ is one of the products that use hand and eye coordination to interact with the application. Natural User Interface (NUI) that KiHECT$^{©}$ use enables user's full body involvement to respond with the developed application. The features of KIHECT$^{©}$ is shown in Table 1.

**Table 1.** KIHECT© Features

| **Hardware** | **Software** |
|---|---|
| Camera mirrored both users' hands detection for the user-application interaction. | Able to set training time, movement pattern and difficulties level. |
| Voice recognition for Start and Stop procedure. | Automatically save training results, and can view personal records by data numbers and graphs. |
| Text to speech compatibility to give command and result. | Intelligent detection to recognize and records specific hand either left or right. |
| Motorized tilt setting for adjusting level of the camera. | Interactive and user friendly interface. |

KIHECT$^{©}$'s original idea came out from the problem of sport science fields that concentrate to enhance athlete's performance for specific parts of the body (hand-eye coordination) for training purpose. It is developed for specific population, i.e.: elite athletes, rehabilitation centers, training centers etc. KIHECT$^{©}$ also uses fewer raw materials as components such as steel, plastic, rubber and others.

Benefits of KIHECT$^{©}$ are listed below:

- Non reflective body marker technology that detects both hands.
- Solving contemporary product issues: high cost product, high maintenance cost, durability issues, bulky, heavy and time consuming for calibration.
- Improving drivers focus, alertness and efficiency for high speed driving and casual driving (i.e.: Formula 1 drivers, daily commuters).
- Enhancing hand-eye coordination for racquet sports athletes.
- Enhancing children cognitive developments in reading ability.
- As treatment method to overcome dyslexia among kids.
- As tools in rehabilitation program for users who are losing an ability to move and coordination between their hands and eyes.

All players were informed about how the power 2000 New Test device works and how the measurements be made. In order to measure learning, test measurements were conducted on all of the members or study group. Measurement environment was quiet, airy and there were no distracters. During tests, players successfully carried out the test's directives and were encouraged to have the maximum attention. Hence, the reliability test was done in one week.

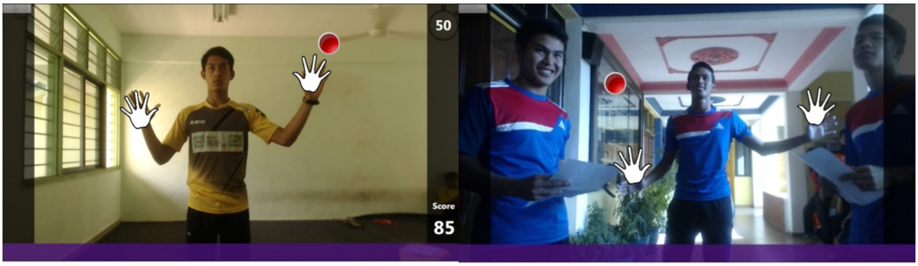


**Fig. 3.** KIHECT© using automatic system (left) and KIHECT$^{©}$ using manual system and helped by other 2 players (right)

## 4 Results

In this study, the participants were recruited from the Sekolah Sukan Tunku Mahkota Ismail (SSTMI) to measure the reliability test. 33 rugby players participated in this experiment to test the consistency of hand eye coordination using KIHECT$^{©}$. Furthermore, KIHECT$^{©}$ developed by combining the hardware and software technology such as consumer depth camera, KIHECT$^{©}$ system and display system.

The validity of the motor task 1 was determined on the comparison between manual and automatic test score (KIHECT$^{©}$). The test-retest procedures are to prove the reliability of KIHECT$^{©}$. Reliability test indicates methods of two routines: for manual system is by evaluating left and right hand scores while programmed system is to survey score from left and right hands. In this device, the red dot as stimulus will appear randomly on the software and subject must react with a good hand-eye coordination to complete the task. The investigational and review models of research, this would mean that if a test and then a retest were conducted within an appropriate time span, the parallel results would be obtained [22].

In fact, hand-eye coordination is important to track the movement of the hands with the eyes, hence enabling the eyes to send the compulsory signals to the brain about hand movement while poor hand-eye coordination can greatly compromise the ability to exercise or even move, which may also affect the everyday basic tasks such as writing [24]. Hence, the score forms were provided to all players to fulfill the requirement accordingly. This study consists of one task only and the reliability study on the BATAK LITE has all 6 motor tasks [25]. Technically, in this study task 1 has been used to measure and train the weakness of hand-eye coordination among sport school athletes. Thus, condition 1 (automatic count) the time given is 30 seconds for the subject to complete the task by touching the red dot as a stimulus randomly, after that it will evaluate the total score of both hands during the test.

In addition to that, the automatic test recorded in the system that was completed in 30 seconds should measure the reaction speed with good hand-eye coordination depending on the total right and left hand touches. After the task is completed, the system will illustrate/demonstrate the result which indicates how many touches of the right hand, left hand and both hands. The data will then be recorded in the score forms (see Fig. 3 (left)). Meanwhile, for condition 2 (manual count) the total number of touches by the right hand and left hand, the time and the total number of both hand touches are required to be counted manually by 2 observers. One observer counts the right-hand touches and another one counts left-hand touches. After the specific time has completed, the sum of the total of both hand touches is then recorded in the score forms (see Fig. 3 (right)).

Data of this research are analyzed with SPSS 16.0. Comparison of the research sampling in terms of reaction hand-eye coordination, reaction times, perception speed and decision making according to gender and variables is made with paired sample t-tests analysis and significant level is determined to be 0.005.

**Table 2.** Comparison paired sample T-test between the total for auto test and retest of left/ right hand scores to measure the hand eye-coordination

| **Variable** | **N** | **Mean** | **Standard Deviation** | **t** | **df** | **Significance** |
|---|---|---|---|---|---|---|
| Total test auto (left/right hand) | 33 | 32.97 | 5.187 | -13.457 | 32 | .000 |
| Total retest auto (left/right hand) | 33 | 39.55 | 3.930 | -13.457 | 32 | .000 |

The results of paired sample t-test, as depicted in table 2, indicate that the comparison between auto (KIHECT$^{©}$) test and retest evaluated by left/right hands touches scores. The p value at 0.000 which is lower than t value at -13.457 at the 0.005 level of significance established. This is means, the mean and standard deviation auto test shows (32.97±39.55) and it is clearly significance. This conclude that using the paired sample t-test using innovation devices as a KIHECT$^{©}$ test score by left/right is reliable to use as a training tools and able to measure the hand eye coordination.

**Table 3.** Paired Samples Correlations for test-retest auto (KIHECT$^{©}$) and manual (count by Manual)

| | | **N** | **Correlation** | **Sig.** |
|---|---|---|---|---|
| Pair 1 | test_auto & retest_auto | 33 | .839* | .000 |
| Pair 1 | test_manual & retest_manual | 33 | .587 | .000 |

As shown at the Table 2, the comparison correlation between test-retest using auto method (KIHECT$^{©}$) and manual (count by manual). Hence, the auto (KIHECT$^{©}$) test-retest show the highest reliability was observed in pair 1, where the correlation coefficient was r = 0.83, while the lowest reliability was noticed in pair 2, where r = 0.58. This is meant the KIHECT$^{©}$ was reliable to use as a training tools.

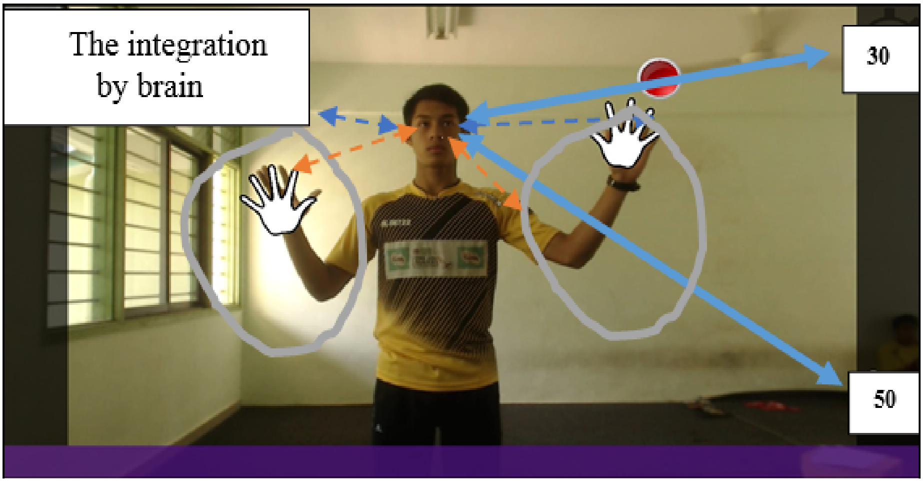


**Fig. 4.** Descriptive in term sport science and the process of Hand-Eye Coordination (HEC) and Reaction Time Speed

## 5 Conclusions and Discussions

KIHECT© developed is intended for use in testing and training will be able to measure and train the hand-eye coordination (HEC) and visual reaction speed (VRS). In addition, tests were performed on these tools by using experimental methods in test-retest. KIHECT© as training tools using high technology is to test the level of hand-eye coordination as well as to assess the level of visual reaction speed. Furthermore, in terms of sports science view, HEC is included the physical fitness component of sport vision. Hence, visual skill is one of the important components to give info to the info for hand to complete the task. In fact, visual is a learned skill that implies an appropriate interpretation of what is seen and interpreted [26]. So, KIHECT© is reliable to be a training tool to improve the visual skill while it is enhance the hand-eye coordination in sport performance. Refer to the figure 5 below; this is how the system and the test initiated.

Refer to the figure above, the skill showed by the subject to test the hand-eye coordination in which it tried to touch stimuli to react as soon as possible. This means that, based on the picture above shows the movement of the subject or skill points to make air visual contact of the stimulus and it is shown that the good combination between eye and hand to touch it. Based on the paired sample t-test, indicate that the mean or the average of the auto (KIHECT©) test is sufficiently better.

This is because; in a given time is 30 seconds, to touch the red dot as a stimulus as much as quickly as possible to get that many in a short time. Motor task 1 from BATAK LITE were applied the test in 30 seconds to touch the stimuli randomly as much as possible [25]. Moreover, KIHECT© usefulness to improve and training for vision in pass and catch technique.

**Acknowledgements.** This research was funded by Prototype Development Research Grant Scheme (PRGS) Grant with title Hand-Eye Coordination Trainer using 3D Sensor Camera for Athlete (R.J130000.7828.4L623) and also supported by Virtual, Visualization and Vision Laboratory (UTM ViCubeLab) part of Media and Games Innovation Center of Excellence (MAGIC-X), Faculty of Computing, University Teknologi Malaysia. Thanks to sports athlete from Sekolah Sukan Tunku Mahkota Ismail (SSTMI), Bandar Penawar, and Johor for the collection of data. Also thanks to members of Sports & Innovation Technology Centre (SITC) and Media and Game Innovation Centre of Excellence (MaGIC-X).

## References

1. Quarrie, K.L., Handcock, P., Toomey, M.J., Waller, A.E.: The New Zealand Rugby Injury and Performance Project. IV. Anthropometric and physical performance comparisons between positional categories of senior A rugby players. British Journal of Sports Medicine 30(1), 53–56 (1996)
2. Smart, D.J.: The reliability of a rugby union specific test: with reference to the relationship between performance and the concentration of testosterone and cortisol (Unpublished Masters Thesis). WINTEC, Hamilton, New Zealand (2005)
3. Gabbett, T.J., King, T., Jenkins, D.: Applied physiology of rugby league. Sports Medicine 38(2), 119–138 (2008d)
4. Holway, F.E., Garavaglia, R.: Kinanthropometry of Group I rugby players in Buenos Aires, Argentina. Journal of Sports Sciences 27(11), 1211–1220 (2009)
5. Quarrie, K.L., Handcock, P., Waller, A.E., Chalmers, D.J., Toomey, M.J., Wilson, B.D.: The New Zealand Rugby Injury and Performance Project. III. Anthropometric and Physical performance characteristics of players. British Journal of Sports Medicine 29(4), 263–270 (1995)
6. Smart, D.J.: Physical Profiling of Rugby Union Players: Implications for Talent Development. Auckland University of Technology Faculty of Health and Environmental Sciences (2011)
7. Ebaumsworld (2012), `http://www.ebaumsworld.com` (retrieved June 30, 2012)
8. Break (November 15, 2011), `http://www.break.com` (retrieved June 30, 2012)
9. Batax (April 1, 2011), `http://www.batax.com` (retrieved February 24, 2012)
10. Okumura, T., Laukkanen, H.: Use of the Visagraph II to Evaluate Eye Movements During Reading of Japanese Text. Journal of Behavioral Optometry 22, 117–125 (2011)
11. Wayne Engineering (July 27, 2011), `http://www.wayneengineering.com/SaccadicFixator` (retrieved December 4, 2012)
12. Beckerman, S., Fornes, A.M.: Effects of changes in lighting level on performance with the AcuVision 1000. Journal of American Optometric Association 68(4), 243–249 (1997)
13. Silverthorn, D.U.: Sensory physiology. In: Berriman, L., Reid, A.A., Dekel, Z. (eds.) Human Physiology - An Integrated Approach, p. 314. Pearson Education, San Francisco (2004)
14. Barrett, B.T.: A critical evaluation of the evidence supporting the practice of behavioural vision therapy. Ophthalmic and Physiological Optics 29(1), 4–25 (2009)
15. Ludeke, A.: Training of Visual Skills and Transferability to Overall Rugby Performance Improvement. Faculty of Health Sciene at the University of Johannesburg (2010)

16. Liam Middleton opinion, Visual Skills For Rugby Union Players. Coachwise Ltd Chelsea close off Amberley Road Armley Leeds LS12 4HP (2011), `http://Rugbyedgetherfu.com`
17. De Toit, P.J., Nerves, R., de Wet, K.B., Kruger, E.: Exercise, performance and sport vision testing. African Journal for Physical, Health Education, Recreation and Dance (suppl.), 140–149 (2007)
18. Valli, A.: Notes on Natural Interaction (2005)
19. Calle, J., Martínez, P., et al.: Towards the Achievement of Natural Interaction. In: Engineering the User Interface, Springer London, pp. 1–19. Springer, London (2009)
20. Zainal, Z.B.: Reka Cipta Alat Mudah Alih Bagi Mencari Bakat Melalui Ujian Ketangkasan. Ijazah Sarjana Muda, Universiti Teknologi Malaysia, Skudai (2009)
21. Bulchanadi, S.: The Science of Cycling: Aerodynamics & Drafting (2012), `http://www.exploratorium.edu/cycling/aerodynamics.html`
22. Oluwatayom, Ayodele, J.: Validity and Reliability Issues in Educational Research. Institute of Education, Ekiti State University, Ado–Ekiti, Nigeria (2012)
23. Cohen, L., Manion, L., Morrison, K.: Research methods in education, 6th edn., pp. 133–164. Routledge Taylor & Francis Group, London & New York (2008)
24. Gonzalez, C.: Eye and Hand Coordination: comparing Effects of Age on Performance. Brunel University London (August 2, 2011)
25. Gierczuk: Reliability and Accuracy of Batak Lite Tests Used For Assessing Coordination Motor Abilitie In Wrestlers (2014)
26. Abel, R.: The eye care revolution, p. 239. Kensington Books, New York (1999)

# 3D Object Detection for Reconstructed Scene Using Multi-layer Growing Neural Gas

Parham Nooralishahi and Chu Kiong Loo

Faculty of Computer Science and Information Technology,
University of Malaya, Kuala Lumpur, Malaysia
{parham.nooralishahi,ckloo}@um.edu.my

**Abstract.** In this paper, we tackle problems of indoor dynamic reconstructed scene created using multiple static Kinect sensors; toward this goal, we propose a dynamic object detection algorithm based on multi-layer growing neural gas for reconstructed scenes and creating a dynamic hierarchical structured space; in fact the proposed technique creates a multi-layer structure for representing a point cloud as points, fragments, objects/groups, and environment. Moreover the proposed algorithm uses statistical outlier removal technique and a down-sampling algorithm based on growing neural gas in order to remove edge and shadow noises being very common in reconstructed scene created using multiple static Kinect sensors. With the proposed algorithm time complexity of object recognition, object tracking algorithms can be decreased. Experimental results demonstrate that the proposed algorithm achieves substantial improvement over the state-of-the-art.

**Keywords:** 3-dimensional object detection, range image segmentation, growing neural gas, hierarchical data representation, depth integration.

## 1 Introduction

Object detection is the process of labeling object-like segments in a point cloud making systems able to separate objects from reconstructed models and using them to post-processes like object recognition, object tracking, or etc.

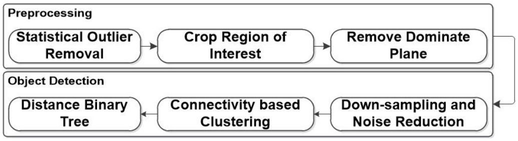


**Fig. 1.** System Architecture

Most of object detection methods are based on curvature and other higher level derivatives lead to over segmentation needing a lot of manual editing or manual parameterizing [1,2,3]. There are many varieties of object detection and point cloud segmentation are available, some of these techniques are edge-based segmentations

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_21

having two general steps: (1) edge detection which followed by (2) a grouping technique for the points inside the boundaries [4,5]; edge-based segmentations are very sensitive to noise. Surface based segmentation algorithm uses local surface properties to merge those together [6]; surface-based segmentation techniques are less sensitive to noises compares to edge-based segmentation techniques. Some object detection and point cloud segmentation algorithm are based on Scanline processing each row in range images to segment the data [7].

Objects can become fragmented in reconstructed scenes created by multiple static Kinect sensors; therefore a method is needed to detect these fragments and to make the system able to integrate them as an object. In this paper, there are two general steps, (1) pre-processing steps, and (2) object detection steps. In pre-processing steps, a point cloud is fetched to the algorithm, then "Statistical Outlier Removal" method is utilized to remove noises, and finally a dominate plane detection algorithm is applied to remove ground surface (Ground-Zero) making objects separated. After that the proposed multi-layer object detection is applied to detect objects; in first layer, a growing neural gas is utilized to down-sample the point cloud, Second layer starts working with first layer simultaneously to cluster neurons according to their connectivity; after that fragments (segments) are determined but there is not any dynamic structure for these segmentations, therefore third layer is applied; in this layer a binary tree is created using distance factor. This iterative process creates a hierarchical structure representing following layers: environment, groups/objects, segments, and points. Although a dynamic hierarchical data structure is created by applying to the third layer.

## 2 Preprocessing

In this paper, there are some preprocessing steps are needed to detect objects accurately using the multi-layer object detection algorithm proposed here.

### 2.1 Statistical Outlier Removal

Edge noises are so common in point clouds, because of relatively low resolution when a sharp discontinuity in depth data occurs, for instance the edge of a book or a chair, edge noises are expected [8].

Most of these irregularities can be omitted by using a simple statistical analysis on each point's neighborhood, and remove those which cannot fulfil certain criteria. This outlier removal is based on the computation of the distribution of point to neighbors' distances in the input dataset. Mean distance from each point to all its neighbors are calculated. In this paper, the number of neighbors considering for each point as neighbors is set to 60, and the standard deviation is set to 9 mm meaning that all points having distance longer than 9 mm standard deviation of the mean distance to the query point will be marked as outlier and removed.

### 2.2 Crop Region-Of-Interest

Two general problems in point cloud processing can increase runtime and complex the procedure, (a) point cloud's size and (b) useless part of point clouds. Therefore by decreasing size of region of interest, runtime will be decreased. In this paper, a predefined boundary is used to crop a specific region of point cloud.

### 2.3 Dominate Plane Detection

One of the most important steps in most of these sort of applications are to recover geometrical structures for scene understanding. By assuming that objects in reconstructed scene defined in this paper are on a surface named as "ground-zero", removing the surface is an elementary step in detecting objects. Dominate planes are detected using Random Sample Consensus (RANSAC) method because the RANSAC robust regression algorithm is the most efficient and popular technique for extracting individual planar patches specially from noisy point clouds having multiple surfaces [9].

## 3 Three Dimensional Object Detection

In this step, a hierarchical dynamic data space is created determining objects using multi-layers growing neural network. As shown in figure 2, firstly a growing neural gas is used for down-sampling the point cloud; while first layer of growing neural gas is continuing to form, second layer is clustering first layer's neurons according to their connectivity. After first and second layers are done, third layer starts forming a binary tree of objects using their distances.

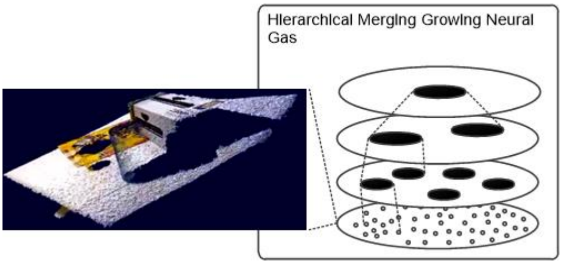


**Fig. 2.** Hierarchical Data Representation

### 3.1 FIRST LAYER : Down-Sampling and Noise Reduction

3-dimensional sensors provide point clouds with huge number of points which probably may include noisy data making it impossible to apply key-point detection and feature extraction algorithms; moreover defining a point cloud with less number of points can reduce algorithm's runtime.

A growing neural gas is used for down-sampling and noise reduction of point clouds in [10], which growing neural gas is applied with a predefined neuron per signal insertion rate value to down-sample the point cloud. Growing neural gas starts working with only two neurons and after certain number of received signals, the algorithm inserts a new neuron to reduce local error of specific input data space's part [11], so down-sampling rate can be calculated as follows,

$$r_d = \frac{1}{s} \quad \text{and } n_{down} \cong n_p \times r_d \tag{1}$$

In equation (1), $r_d$ is down-sampling rate, s is the number of signals for neuron insertion, and $n_{down}$ is number of down-sampled point cloud. Classical filtering and noise reduction techniques like median and mean filters have been used in many researches to filter noises in point clouds collected from 3-dimensional sensor. the median filter is one of simplest and popular filters that has used because it is simple to implement and more or less efficient, but the median filter can remove noise only if noisy pixels occupy less than half of the pixel's neighborhood area; although the median filter removes noisy pixels, details of dataset are also may be removed [10].

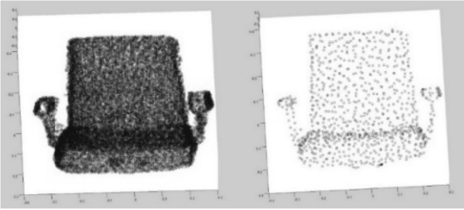


**Fig. 3.** Down-sampling algorithm is applied on the left point cloud and the result is shown on the right point cloud

By applying multiple experiments comparing the results of RANSAC down-sampling method and the result of growing neural gas is shown that by applying growing neural gas as down-sampling technique the results will be extremely improved [10].

### 3.2 SECOND LAYER : Clustering

As mentioned before, first layer extracts key-feature points, therefore the result is a structured point cloud having information about their neighbors, and so second layer uses the connectivity information for each point to extract fragments. Clustering process gets all resulted particles then selects one of the particles and inserts it into a new layer, then it sets the selected particle as a used particle, after that the same process is applied to each of them and inserts them to the layer as the selected particle were until there will be no targeted particles.

### 3.3 THIRD LAYER : Dynamic Hierarchical Structured Data

In third layer, the algorithm starts to form a hierarchical data structure of detected fragments. The structure presents the environment as root node having the highest abstraction level and the particles creating objects as lowest abstraction level.

As mentioned before, Microsoft Kinect sensor creates noisy distance data including shadow and edge noises; shadow noises cause object fragments in Kinect depth data meaning that an object can be divided into separated fragments; moreover reconstructed scans using multiple static Kinect sensors cannot match the distance data completely so an object may represented as multiple segments.

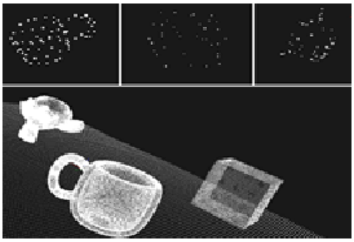


**Fig. 4.** The generated point cloud and detected objects.

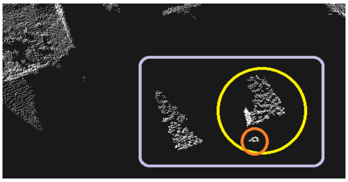


**Fig. 5.** Fragments in a point cloud are shown in this figure

A dynamic hierarchical structure is offered in this paper for solving fragmentation problem in Kinect distance data. Fragments are grouped according to their distance using a binary tree structure. First a distance matrix is created showing the distance of each fragment from other fragments.

$$D_{N\times N}=\begin{bmatrix} 0 & d_{o_1\to o_2} & d_{o_1\to o_3} \\ d_{o_2\to o_1} & 0 & d_{o_2\to o_3} \\ d_{o_3\to o_1} & d_{o_3\to o_2} & 0 \end{bmatrix},\ D(o_i,o_j)=\sqrt{\sum_{k=1}^{n_{dimensions}}\left(o_i(k)-o_j(k)\right)^2} \qquad (2)$$

All fragments are inserted into a tree as leaves then two closest fragments are combined under a new node, furthermore the two selected nodes are replaced with the created parent node in the distance matrix and all distances are recalculated; this process repeats itself until there is only one node remains in distance matrix.

```
[Point : (-0.03, -1.17,  0.11) : 91]
  ⊞ [Point : ( 0.18, -0.62,  0.15) : 17]
  ⊟ [Point : (-0.08, -1.29,  0.10) : 74]
      ⊞ [Point : ( 0.02, -1.72,  0.07) : 7]
      ⊞ [Point : (-0.09, -1.25,  0.10) : 67]
```

**Fig. 6.** Object detection results of a sample point cloud

## 4 Results and Discussions

Three dataset are used in this study to evaluate proposed 3-dimensional object detection, (a) generated point cloud dataset, (b) MPII Multi-Kinect dataset, and (c) reconstructed scene using Microsoft KinectFusion software. As seen before, the proposed algorithm is used various techniques to reduce noises – edge and shadow noises – and finally a multi-layer object detection technique to extract objects; so to evaluate the detection subsystem first an ideal and isolated data set can be used. Because these point clouds are generated artificially, there isn't any unexpected data or noise. Seven 3-dimensional scene is generated using Blender software and saved in XYZ file format.

In order to evaluate the algorithm with a standard dataset, MPII Multi-Kinect dataset proposed in [12] is used. This novel RGBD dataset is created in Max Planck Institute Informatik (MPII) using four static Kinects. The dataset includes 9 different object classes aggregated in various challenging scenes. This dataset is used by some studies for evaluating object detection and other algorithms [13, 14] making us able to compare the proposed algorithm to other available comparable algorithms.

This study is designed specifically for reconstructed scene. A dataset is also created using Microsoft KinectFusion software to let the algorithm be tested in reconstructed scene created by multiple Kinects; it includes five different objects in two different scenes further the scenes are indoor and none of the objects are transparent.

As mentioned before, growing neural gas is used to down-sample point cloud, so growing neural gas's parameters influence the result of proposed algorithm directly. There are two important factors related to growing neural gas, (a) down-sampling rate, and (b) objects' distances. Down-sampling rate is formulated in this study as follows,

$$P' = \left[\frac{P-2}{\lambda}\right] + 2,\ dsr \approx \frac{\lambda}{P} \tag{3}$$

$P'$ is number of points in down-sampled point cloud, P is number of original point cloud and $\lambda$ is number of signals for inserting new neuron. "dsr" is down-sampling rate showing number of needed signals to insert a new neuron; this down-sampling rate is used to generalize the testing process to find strengths and limitations. Finding the best down-sampling rate helps us to define a generalized parameters for the algorithm which can be defined manually or automatically.

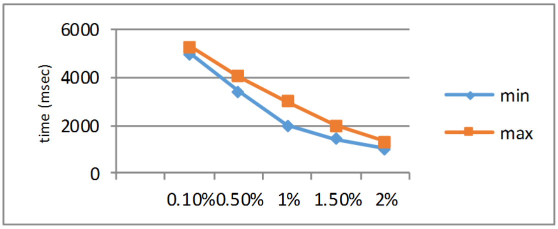


**Fig. 7.** Chart of down-sampling results using growing neural gas

The results are shown above is measured on an Intel Core i7 3.4 GHz, RAM 12.0 GB (DELL XPS 8500) using twenty different point clouds with different sizes. As it is shown above, whenever down-sampling rate is increased, proposed algorithm is faced a significant reduction in its runtime, moreover size of point cloud also influence upon the down-sampling algorithm apart from down-sampling rate.

**Fig. 8.** Results of applying GNG algorithm on different point clouds with different sizes

The result of applying different size of point cloud shows that growing neural gas's learning time is being increased exponentially faced with big point clouds and the point cloud's quality is begin decreased so a balanced value should be chosen.

**Fig. 9.** Down-sampling results, the left photo is for 2% down-sampling, central photo is for 1% down-sampling, and the right photo is for 0.1% down-sampling

Experimental results are shown that more than 0.6% down-sampling rate makes object recognition and labeling processes more difficult and also as you can see, for example results with 2% down-sampling rate is unformed.

The factors considered in this evaluation are number of objects, noise rate, position of all objects toward each other, and distance of objects from each other, because in this algorithm objects with very short distance may be considered as one object.

**Table 1.** Evaluation results for the proposed algorithm

| PC Size | Objects | Detected | noise rate | PC Size | Objects | Detected | noise rate |
|---|---|---|---|---|---|---|---|
| 241407 | 3 | 4 | 13.07% | 307634 | 2 | 2 | 0% |
| 91886 | 3 | 3 | ~ 0% | 491084 | 5 | 4 | 1.1% |
| 325699 | 2 | 2 | ~ 0% | 1816500 | 2 | 5 | 56% |
| 44049 | 1 | 2 | 21.27% | 5414531 | 4 | 4 | 31% |
| 412311 | 5 | 7 | 32.12% | | | | |

The experiments show that the algorithm introduced in this paper can reduce noises in a point cloud and especially it is so effective in the face of shadow noises, however this algorithm also has its own limitations; according to the experiments, point clouds with more that 31% can make the results questionable, however the hierarchical data representation proposed in this paper, makes the system able to improve the results.

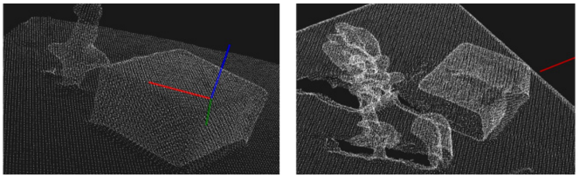


**Fig. 10.** Examples of point clouds used in this paper

## 5 Conclusion and Future Work

In this paper, we have presented a 3-dimensional multilayer object detection based on growing neural gas for reconstructed scene which it detects objects from reconstructed scene created by a multi static Kinects reconstruction technique. The main idea is first to remove noisy data then to detect ground-zero in the point cloud and finally to detect objects in the point cloud using a multi-layer 3-dimensional object detection technique; moreover the proposed algorithm creates a dynamic multilayer data representation structure to assist object recognition techniques by avoiding fragments and prioritrize the segments.

In future work, we would like to improve the proposed technique in order to reduce its time complexity and to modify the algorithm for omitting the fragments directly and maximizing the detection accuracy by avoiding detection of objects as fragments using a filter based on probability techniques.

**Acknowledgement.** This research is supported by High Impact Research MoE Grant UM.C/625/1/HIR/MoE/FCSIT/10 from the Ministry of Education Malaysia.

## References

1. Rabbani, T., van den Heuvel, F., Vosselman, G.: Segmentation of point clouds using smoothness constraint. In: ISPRS Commission V Symposium Image Engineering and Vision Metrology (2006)
2. Douillard, B., Underwood, J., Kuntz, N., Vlaskine, V., Quadros, A., Morton, P., Frenkel, A.: On the segmentation of 3D LIDAR point clouds. In: 2011 IEEE International Conference on Robotics and Automation. IEEE Press, Shanghai (2011), doi:10.1109/ICRA.2011.5979818
3. Golovinskiy, A., Funkhouser, T.: Min-cut based segmentation of point clouds. In: IEEE Workshop on Search in 3D and Video, S3DV (2009)

4. Bihanu, B., Lee, S., Ho, C.C., Henderson, T.: Range data processing: representation of surfaces by edges. In: Proceedings of Eighth International Conference on Pattern Recognition, pp. 236–238 (1986)
5. Wani, M.A., And Arabnia, H.R.: Parallel edge region-based segmentation algorithm targeted at reconfigurable multiring network. Journal of Supercomputing 25(1), 43–62 (2003)
6. Xiang, R., Wang, R.: Range image segmentation based on split-merge clustering. In: 17th ICPR, pp. 614–617 (2004)
7. Sithole, G., Vosselman, G.: Automatic structure detection in a point cloud of an urban landscape. In: 2nd Joint Workshop on Remote Sensing and Data Fusion Over Urban Aereas (2003)
8. Natale, D.J., Baran, M.S., Tutwiler, R.L.: Point cloud processing strategies for noise filtering, structural segmentation, and meshing of ground-based 3D Flash LIDAR images. In: IEEE 39th Applied Imagery Pattern Recognition Wrokshop (AIPR). IEEE Press, Washington, DC (2010)
9. Gallo, O., Manduchi, R., Rafii, A.: CC-RANSAC Fitting planes in the presence of multiple surface in range data. Pattern Recognition Letters 32(3), 403–410 (2010)
10. Garcia-Rodriguez, J., Cazorla, M., Orts-Escolano, S., Morell, V.: Improving 3D keypoint detection from noisy data using Growing Neural Gas. In: Rojas, I., Joya, G., Cabestany, J. (eds.) IWANN 2013, Part II. LNCS, vol. 7903, pp. 480–487. Springer, Heidelberg (2013)
11. Fritzke, B.: A Growing Neural Gas Network Learns Topologies. In: Advances in Neural Information Processing System 7, pp. 625–632. MIT Press (1995)
12. Susanto, W., Rohrbach, M., Schiele, B.: 3D object detection with multiple Kinects. In: 2nd Workshop on Consumer Depth Cameras for Computer Vision, ECCV Workshop (2012)
13. Zhang, X., Wang, X., Jia, Y.: The Visual Internet of Things System Based on Depth Camera. In: Sun, Z., Deng, Z. (eds.) Proceedings of 2013 Chinese Intelligent Automation Conference. LNEE, vol. 255, pp. 447–455. Springer, Heidelberg (2013)
14. Yang, Z., Xiong, Z., Zhang, Y., Wang, J., Wu, F.: Depth Acquisition from Density Modulated Binary Patterns. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 25–32. IEEE Press, Protland (2013)

# Modified K-means Combined with Artificial Bee Colony Algorithm and Differential Evolution for Color Image Segmentation

Mohammad Babrdel Bonab[1,*], Siti Zaiton Mohd Hashim[1],
Ahmed Khalaf Zager Alsaedi[2], and Ummi Raba'ah Hashim[1]

[1] Faculty of computing, Universiti Teknologi Malaysia, Johor Bahru, Malaysia
`bbmohammad2@live.utm.my, sitizaiton@utm.my, ummi@utem.edu.my`
[2] Faculty of Computer sciences, Universiti Tun Hussein Onn Malaysia, Malaysia
`ahmedkhalafalsady@yahoo.com`

**Abstract.** Clustering is one of most commonly used approach in the literature of Pattern recognition and Machine Learning. K-means clustering algorithm is a fast and simple method in the clustering approaches. However, due to random selection of center of clusters and the adherence to preliminary results of center of clusters, the risk of trapping to a local minimum ever exist.in this study, we have taken help of effective hybrid of optimization algorithms, artificial bee colony (ABC) and differential evolution (DE), is proposed as a method to mentioned problems. The proposed method consists of two main steps. In first step, Seed Cluster Center Algorithm employed to best initial cluster centers. The combined evolutionary algorithm explores the solution space to find global solution. The performance of proposed method evaluated with standard data set. The evaluation results of the proposed algorithm and its comparison with other alternative algorithms in literature confirms its superior performance and higher efficiency.

## 1 Introduction

Data Clustering is one of the most important knowledge discovery techniques to extract structures from data set and is widely used in data mining, machine learning, statistical data analysis, vector quantization and pattern recognition. The aim of clustering is to partition data into k cluster, so that each cluster contains data, which has most similarity and maximum dissimilarity with the other clusters. Clustering algorithms can be comprehensively classified into hierarchical, partitioning, model-based, grid-based and concentration-based clustering algorithms [1, 2].

The k-means algorithm is one of the partitional clustering algorithm and one of the most popular algorithms, used in many domains. The k-means algorithm implementation is easy and often practical. However, results of k-means algorithm considerably depend on initial values. In other words, its efficiency highly depends on the first

---

[*] Corresponding author.

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_22

initial cluster centers and parameters. The main purpose of k-means clustering algorithm is to minimize the diversity of all objects in a cluster from their cluster centers. The initialization problem of k-means algorithm is considered by heuristic algorithms but it still risks being trapped in local optimality. Therefore, for achieving a better cluster algorithm we should find a solution for overcoming the problem of trap into local optimum.

There are many studies to overcome this problem. For instance, Niknam et al have proposed a hybrid approach based on combining partial swarm optimization and ant colony optimization with kmeans algorithm for data clustering [3], and Cao et al have proposed a combination technique based on the hybrid of k-means, genetic algorithm and maximization of logarithmic regression expectation [4]. Kao et al have presented a combination algorithm according to hybrid of partial swarm optimization, Nelder-Mead simplex search and genetic algorithm [5]. The main contribution of this article is introduction of a new novel combination of evolutionary algorithm according to artificial bee colony algorithm and differential evolution to overcome mentioned problem, hybrid with initial cluster centers algorithm. In the rest of article, in section two and section three the data clustering and the k-means technique are reviewed. In section four and five, artificial bee colony algorithm and differential evolution algorithm are introduced. In section six, the proposed algorithm is described. In section seven, experimental results of proposed algorithm are shown and compared with alternative algorithms on benchmark data and in the second experimental, the industrial images has been used to image segmentation and finally section eight is included the conclusions.

## 2 Data Clustering

Clustering revolves around grouping samples or observations inside classes of similar objects. A cluster is a collection of samples similar to each other and diverse from the samples of other clusters. In clustering, data is divided into clusters with maximum similarity among the data inside a cluster and minimum similarity among different clusters . In classic definition, the hard type-clustering algorithm of $X = \{X_1, X_2, \ldots, X_k\}$ is a data set of $n$ members that are to be grouped or clustered into $k$ clusters as $C = \{C_1, C_2, \ldots, C_k\}$ so that each cluster satisfies the following:

1. $\bigcup_{i=1}^{k} \quad C_i = X$
2. $C_i \neq \emptyset \quad i = 1 \ldots k$
3. $C_i \cap C_j \neq \emptyset, \quad i \neq j, \quad i, j = 1 \ldots k$

Hence the number of different states for clustering $n$ objects into $k$ cluster will be:

$$NW(n,k) = \frac{1}{k!} \sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n \tag{1}$$

In many methods, the number of clusters is defined manually, so in Equation (1), it implies that even with a given k, finding the best scenario for clustering is difficult. Also clustering states from n observations to k clusters increases by (k^n )/k!. Due to this problem, the best way of clustering n observations into k clusters is an NP-Complete problem, which should be solved optimally by engaging some techniques.

## 3 K-means Clustering

K-means clustering algorithm is a fast and simple method, which is generally used because of its ease of implementation and little iteration. k-means algorithm find cluster centers $(C_1, C_2, \dots, C_k)$ by minimizing the squared sum of the distance of each point $x_i$ from the nearest cluster $C_j$. The initial selection method for cluster centers influences the performance of k-means algorithm. The following stages are suggested for clustering data in the k-means clustering algorithm.

1. Randomly choose *k* data objects from $X = \{X_1, X_2, \dots, X_k\}$ as cluster centers of $(m_1, m_2, \dots, m_k)$
2. Based on Equation (2) add $x_i$ object from $X = \{X_1, X_2, \dots, X_k\}$ set to $Cj$ cluster using Equation (2) provided that:

$$\|xi - mj\| < \|xi - mp\| \; 1 \le p \le k, j \neq p \tag{2}$$

3. Based on the clustering of previous stage the new cluster centers are calculated by using Equation (3) in a way that $n_i$ is the number of objects in the cluster:

$$m_i^* = \frac{1}{n_i} \sum_{x_j \in C_i} x_j \quad 1 \le i \le k \tag{3}$$

4. If the cluster centers were changed, repeat the algorithm from stage 2, otherwise do the clustering based on the resulted centers.

The performance of the k-means clustering algorithm relies on initial centers and this is a major challenge in this algorithm. Random selection of initial cluster centers makes this algorithm yield out different results for different runs over the same data sets, which is considered as one of potential disadvantages of this algorithm [5].

### 3.1 Initial Cluster Centers Algorithm

In this study, with regards to efficiency purposes, first all data objects are clustered using k-means algorithm to find the initial cluster centers to be used in the solutions based on all their attributes. Based on the generated clusters the pattern for an object is produced from each attribute at any stage. Objects with the same patterns are located in one cluster and hence all objects are clustered. The obtained clusters in this stage will be more than the original number of clusters. For this reason, refer to paper [6]. In this paper, clustering is completed in two stages. The first stage is performed as discussed above and in the second stage; similar clusters are integrated with each other until achieving a given number of clusters. Algorithm 1 shows the proposed approach for initial clustering of data objects and the achieved cluster centers are called seed points.

Algorithm 1. ***Proposed Initial Cluster Center Algorithm***

```
1. Input: Data SET ($X = \{x1, x2, \dots, xn\}$), Attribute Set ( $A = \{A1, A2, \dots, Aq\}$), Cluster Number (K),
2. Output: Clusters Seed Set ($SC = \{sc1, sc2, \dots, scH\}$)
3. Begin
4. While ($\forall Aj \in A$)
        4.1. Standard Deviation ($\sigma_j$ ) and Compute Mean ($\mu_j$)
        4.2. Compute Cluster Center ($e$= 1,2,..., $k$ )
             $X_e = Z_e * \sigma_j + \mu_j \quad Z_e = \frac{2*e-1}{2*k}$
        4.3. Perform k-means on this attribute
        4.4. Allocate labels to each cluster obtained from step 4.3 to every data pattern
5. Find unique patterns ($H \geq k$ ) and clustering each data with obtained patterns.
6. Return SC
7. End;
```

### 3.2 Fitness Function

To calculate the fitness of each solution, the distance between the centers of clusters and each data will be used. To do this, firstly a set of cluster centers will be generated randomly and then clustering of the numerator will be conducted based on equation 2**Error! Reference source not found.**. Now according to centers obtained in the interaction step, the new centers of the clusters, and fitness of solutions based on equation 3 will be calculated.

$$\mathrm{Fitness(C)} = \sum_{i=1}^{k} \sum_{x_j \in C_i} \|x_j - m_i^*\| \tag{4}$$

## 4 Artificial Bee Colony Algorithm

Artificial bee colony algorithm is one of the most recent optimization algorithms that are designed according food search process of honeybees. [2, 7]. ABC algorithm is consists of three parts of bees: employed bees, scout bees and onlooker bees. In this algorithm, a food resource is one possible solution for optimization problem. Nectar amount of a food source shows the quality of solution by food source. In the early stages of bees are divided into two classes, the employed bees and onlooker bees. First employed bees starts to seek around hive without any knowledge about solution space. At this stage, the number of food source is equal with number of employed bees, in the other words, for each food source in one place is equal with the number of employed bees to search them [2]. The amount of food source nectar at this stage are determined, in other words, quality grade of initial solution is calculated. Generally each employed bee flies to find new food source in around neighborhood of previous food source and then is evaluated the nectar of them. [8]. The general flowchart of ACB algorithm is shown in the Figure 1.

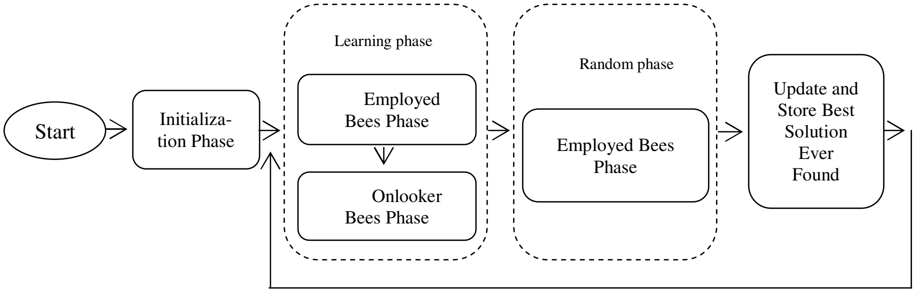


**Fig. 1.** Artificial Bee Colony Algorithm

## 5 Differential Evolution

Differential evolution is a type of standard genetic algorithm. Differential evolution algorithm is evaluates the initial population by using probability motion and observation models and population evolution is performed by using evolution operators [9]. The main idea in the differential evolution algorithm is to generate a new solution for each solution by using one constant member and two random members. In each generation, the best member of population is selected and then the difference between each member of population and the best member is calculated. Two random members are then selected and the difference between them is calculated. Coefficient of this difference is added to ith member and thus a new member is created. The cost of each new member is calculated and if the cost value of the new member is less, the ith member is replaced instead of ith member otherwise the previous value be kept in the next generation . Generally, the DE algorithm steps are as follows:

**Algorithm 3**. *Pseudo-code of differential evolution*

```
1. Define algorithm parameter
2. Generate and evaluate initial population or solutions
3. For all members of population per form the following steps
    a. With mutation operator create a new trial solution
    b. By using the crossover generate new solution and evaluate them
    c. Replace new solution with current solution if new solution is better than current solution otherwise, the current solution is retained.
4. Return to step three if termination condition is not achieved.
```

## 6 Proposed Algorithm

As noted in the former sections, studies conducted on the differential evolution and artificial bee colony algorithm have shown these algorithms can be a powerful approach with enough performance to handle different types of nonlinear problems in various fields. However, it can be possibly trapped into local optimum. To increase the diversity of algorithm, we have used the differential evolution algorithm to find better food resources. In the DE algorithm, all members have a possibility to win the global optimum and move to that side. The ability of the best particle to local search also depends on the other particles by selecting the two other particles and calculating the difference between them. This situation may lead to local convergence.

In this algorithm, each employed bee decides based on three sources that it gives more power to this algorithm and increases the exploitation. In this algorithm, for each employed-bee are applied differential evolution operators. To escape from random selecting of the global best particle we used competency selection for choosing the global best particle. If particle is better than the other solutions, then the probability of being selected is greater.

The basic idea behind the proposed algorithm is that our solutions are grouped based on the artificial bee colony because of its power and ability. In onlooker phase, for onlooker bees, fitness value and selection probability is calculated previously. In this phase, first a source site is selected based on the selection probability and the selection probability is based on the fitness value then for selected site behavior like employed bees (differential evolution operators). To increase the exploration in this algorithm for each scout bee is considered a counter. In the scout bees phase, if counter for each scout bee is more than MaxCount then it will replace with new solution where MaxCount is a positive number. This effectively helps us to prevent from trapping to local optimum. In the following algorithm, the entire process is shown.

Algorithm 4. ***Pseudo-code of proposed CCIA-ABC-K algorithm***

```
Input: Data SET ($X = \{x1, x2, ..., xn\}$), Attribute Set ( $A = \{A1, A2, ..., Aq\}$), Cluster Number (K),
Output: Clusters Seed Set ($SC = \{sc1, sc2, ..., scH\}$)
Begin
Find –seed –cluster–centers (); (preprocessing phase);
Initialize Population();
While (r < MaxIt)
    I.    For each Employed Bees
              % Mutation
        1.    Choose target site and base site from this group of bees
        2.    Random choice of two sites from this group of bees
        3.    Calculate weighted difference site
        4.    Add to base selected site
              % Crossover
        5.    Perform crossover operation with Crossover Probability
        6.    Evaluate the trial site that is generated
                  % update site
        7.    If trial site is less than target site
                      Select trial site instead of target site
                else
                      Select target site
                End if
    II.   Calculate Fitness Value and Selection Probability
    III.  For each  Onlooker Bees
        1.    Select Source Site randomly
              % Mutation
        2.    Choose target site and base site from this group of bees
        3.    Random choice of two sites from this group of bees
        4.    Calculate weighted difference site
```

```
    5.    Add to base selected site
          % Crossover
    6.    Perform crossover operation with Crossover Probability
    7.    Evaluate the trial site that is generated
              % update site
    8.    If  trial site is less than target site
                  Select trial site instead of target site
            else
                  Select target site
            End if
IV.     For each Scout  Bees
    1.    If  counter  is less than MaxCount
                  Select position randomly
                      Evaluate new position
                     Counter=0;
V.      Sort and update best site ever found
End of while
End
```

## 7 Experimental Results

To evaluate the accuracy and efficiency of the proposed algorithm, experiments have been performed on four standard datasets from which to determine the correctness of clustering algorithms will be used. This collection includes Iris, Glass, Wine, and CMC data sets that have been chosen from standard UCI data set and color images in deferent light condition. The suggested algorithm is coded by an appropriate programming language and is run in a i5 computer with 2.60 GHz microprocessor speed and 4GB main memory. For measuring the efficiency of the proposed algorithm, the standard data items of Table (1) are used.

**Table 1.** Type Styles

| **Name of Data set** | **Data Set attribute** | | |
|---|---|---|---|
| | ***Size of data set*** | ***No. of Cluster*** | ***No. Of attribute*** |
| Iris | 150(50,50,50) | 3 | 4 |
| Wine | 178(59,71,48) | 3 | 13 |
| CMC | 1473 (629, 334, 510) | 3 | 9 |
| Glass | 214 (70, 17, 76, 13, 9, 29) | 6 | 9 |

### 7.1 Standard Data Using Proposed Algorithm

For better study and analysis of the proposed approach, , the execution results of the proposed approach along with ACO, TS, GA, PSO-SA, HBMO, ACO-SA, PSO, K-Means, PSO-ACO-K and SA results which are reported in Ref [3] are tabulated in the Tables (2) to (5). In Tables (2) to (5) best, worst and average results are reported for 100 runs respectively. The resulted tables illustrate the distance of each data object from cluster center is presented in section 3.2. As simply observed in the figures, the proposed algorithm generates acceptable solutions according to execution time.

**Table 2.** The Results Of Implementing The Algorithms Over Iris Test Data For 100 Runs

| **Method** | ***Result*** | | | ***CPU Time (S)*** |
|---|---|---|---|---|
| | ***Best*** | ***Average*** | ***Worst*** | |
| PSO-ACO-K | 96.650 | 96.650 | 96.650 | ~16 |
| PSO–ACO | 96.654 | 96.654 | 96.674 | ~17 |
| PSO | 96.8942 | 97.232 | 97.897 | ~30 |
| SA | 97.457 | 99.957 | 102.01 | ~32 |
| TS | 97.365 | 97.868 | 98.569 | ~135 |
| GA | 113.986 | 125.197 | 139.778 | ~140 |
| ACO | 97.100 | 97.171 | 97.808 | ~75 |
| HBMO | 96.752 | 96.953 | 97.757 | ~82 |
| PSO_SA | 96.66 | 96.67 | 96.678 | ~17 |
| ACO-SA | 96.660 | 96.731 | 96.863 | ~25 |
| k-Means | 97.333 | 106.05 | 120.45 | 0.4 |
| Proposed ALG. | 96.5411 | 96.5484 | 96.5672 | ~12 |

**Table 3.** The Results Of Implementing The Algorithms Over wine Test Data For 100 Runs

| **Method** | ***Result*** | | | ***CPU Time (S)*** |
|---|---|---|---|---|
| | ***Best*** | ***Average*** | ***Worst*** | |
| PSO–ACO–K | 16,295.3 | 16,295.3 | 16,295.3 | ~30 |
| PSO–ACO | 16,295.3 | 16,295.9 | 16,297.9 | ~33 |
| PSO | 16,345.9 | 16,417.4 | 16,562.3 | ~123 |
| SA | 16,473.4 | 17,521.0 | 18,083.2 | ~129 |
| TS | 16,666.2 | 16,785.4 | 16,837.5 | ~140 |
| GA | 16,530.5 | 16,530.5 | 16,530.5 | ~170 |
| ACO | 16,530.5 | 16,530.5 | 16,530.5 | ~121 |
| HBMO | 16,357.2 | 16,357.2 | 16,357.2 | ~40 |
| PSO_SA | 16,295.8 | 16,296.0 | 16,296.1 | ~38 |
| ACO-SA | 16,298.6 | 16,310.2 | 16,322.4 | ~84 |
| k-Means | 16,555.6 | 18,061.0 | 18,563.1 | 0.7 |
| Proposed ALG. | 16,292.45 | 16,292.81 | 16,293.23 | ~24 |

**Table 4.** The Results Of Implementing The Algorithms Over CMC Test Data For 100 Runs

| **Method** | ***Result*** | | | ***CPU Time (S)*** |
|---|---|---|---|---|
| | ***Best*** | ***Average*** | ***Worst*** | |
| PSO–ACO–K | 5,694.28 | 5,694.28 | 5,694.28 | ~31 |
| PSO–ACO | 5,694.51 | 5,694.92 | 5,697.42 | ~135 |
| PSO | 5,700.98 | 5,820.96 | 5,923.24 | ~131 |
| SA | 5,849.03 | 5,893.48 | 5,966.94 | ~150 |
| TS | 5,885.06 | 5,993.59 | 5,999.80 | ~155 |
| GA | 5,705,63 | 5,756.59 | 5,812.64 | ~160 |
| ACO | 5,701.92 | 5,819.13 | 5,912.43 | ~127 |
| HBMO | 5,699.26 | 5,713.98 | 5,725.35 | ~123 |
| PSO_SA | 5,696.05 | 5,698.69 | 5,701.81 | ~73 |
| ACO-SA | 5,696.60 | 5,698.26 | 5,700.26 | ~89 |
| k-Means | 5,842.20 | 5,893.60 | 5,934.43 | 0.5 |
| Proposed ALG. | 5,532.318 | 5,532.345 | 5,533.901 | ~71 |

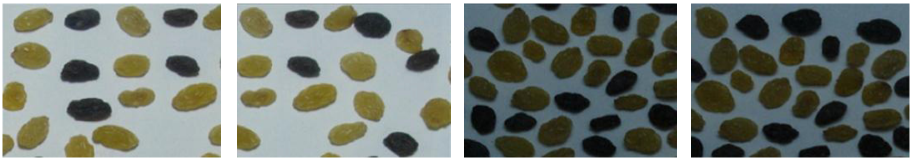
**Table 5.** The Results Of Implementing The Algorithms Over wine Glass Data For 100 Runs

| **Method** | ***Result*** | | | ***CPU Time (S)*** |
|---|---|---|---|---|
| | ***Best*** | ***Average*** | ***Worst*** | |
| PSO–ACO–K | 199.53 | 199.53 | 199.53 | ~31 |
| PSO–ACO | 199.57 | 199.61 | 200.01 | ~35 |
| PSO | 270.57 | 275.71 | 283.52 | ~400 |
| SA | 275.16 | 282.19 | 287.18 | ~410 |
| TS | 279.87 | 283.79 | 286.47 | ~410 |
| GA | 278.37 | 282.32 | 286.77 | ~410 |
| ACO | 269.72 | 273.46 | 280.08 | ~395 |
| HBMO | 245.73 | 247.71 | 249.54 | ~390 |
| PSO_SA | 200.14 | 201.45 | 202.45 | ~38 |
| ACO-SA | 200.71 | 201.89 | 202.76 | ~49 |
| k-Means | 215.74 | 235.5 | 255.38 | ~1 |
| Proposed ALG. | 211.163 | 212.5705 | 13.7614 | ~95 |

According to reported results in tables two to five the proposed method over Iris, CMC and Wine Data Sets provides the best results in comparison with other mentioned algorithms. According to tables five the suggested algorithm over Glass Dataset provides acceptable results rather than the alternative algorithms. The reason for this behavior is justified by the fact that as data objects increase in number the efficiency of the alternative algorithms decrease while the deficiency of the suggested algorithm highlights more.

## 7.2 Image Segmentation Using Proposed Algorithm

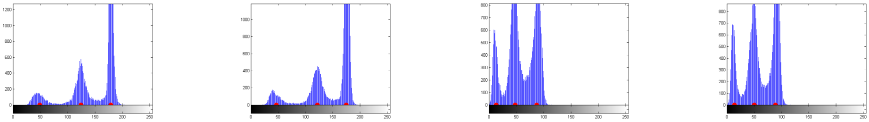
Digital image segmentation is a complex and challenging issue that can be different depending on the nature of inherent of image. There are many methods for image segmentation issues. In recent years, much research has been done on the evolutionary algorithm inspired from nature and human societies. In this paper digital image segmentation though modified k-means algorithm combined with artificial bee colony algorithm is performed. Image histogram is a chart that is made by the number of pixels on image that is determined based on brightness level. To obtain a histogram of image is enough to scrolling in whole pixel of image and calculate the number of pixels for each brightness level. Figure 2 shows the image samples that are used in this paper for image segmentation [10]. To evaluate the performance of proposed method are used the histograms of images to find the peak points of images that number of peak points is equal with desired number of segment and number of segment is equal with number of clusters. To find the best cluster centers and optimum cluster centers are used the modified k-means algorithm to find minimum distance between each data object with each cluster centers. Proposed algorithm should be able to separate the segments of images in different light condition with high accuracy and precision.



(a) (b) (c) (d)

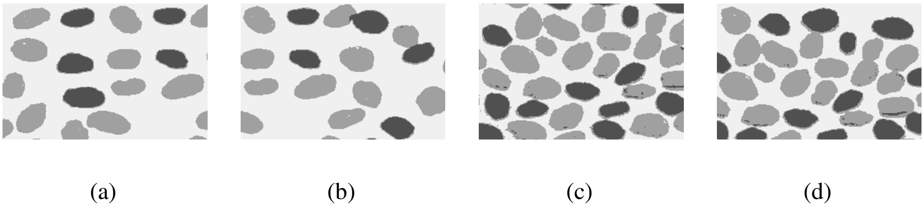
**Fig. 2.** Used sample images for image segmentation in the different light condition

By specifying, the location of peaks in the histogram can obtain the optimal threshold for image segmentation. Figure 4 shows segmented images on the different light condition is shown. Figure 4 shows segmented images on the different light condition is shown. In Figure 3, the peak points illustrated with red points as best cluster centers.



(a) (b) (c) (d)

**Fig. 3.** Histogram charts with optimum cluster centers and peak locations



(a) (b) (c) (d)

**Fig. 4.** Segmented image on the different light condition

According to the reported results in figure 4, the suggested algorithm over images on different light condition provides the best segmentations and do not depend on situations and behavior of data. It can be work on the different condition and different behavior. In figure 3, we can see the histograms and peak points that have been found with proposed algorithm on the different light condition or different behavior of data. The reason of this behavior is justified by the fact that suggested algorithm not sensitive to behavior of data and can find the global optimum with this effective hybrid algorithm.

## 8 Conclusions

In this paper, modified k-means algorithm with an effective combined artificial bee colony and differential evolutionary algorithm was used for optimal solutions of objects in datasets and images. Due to the high capability of convergence and accuracy in artificial bee colony, to achieve global optimum point in the optimization problem were expected by using this algorithm can be separated the different part of image with high speed and accuracy. To increase the exploitation of proposed algorithm the differential evolution algorithm is used. In this article, used method to determine the best cluster centers and peaks in the histogram of image with modified hard k-means clustering. The optimal cluster centers obtained by combination of artificial bee colony and modified k-means clustering algorithm. The results of comparison with alternatives algorithm show that proposed algorithm with optimum clusters is able to segmentation in any light conditions.

**Acknowledgement.** The authors would like to express their cordial thanks to University Technology Malaysia (UTM) for research university grant with number of (Q.J130000.2528.06H90). The authors are also grateful to thanks and soft computing research group (SCRG) for their support and incisive comments in making this study a success.

## References

1. Gan, G., Ma, C., Wu, J.: Data clustering: theory, algorithms, and applications, vol. 20. SIAM (2007)
2. Karaboga, D., Basturk, B.: A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. Journal of Global Optimization 39(3), 459–471 (2007)

3. Niknam, T., Amiri, B.: An efficient hybrid approach based on PSO, ACO and k-means for cluster analysis. Applied Soft Computing 10(1), 183–197 (2010)
4. Nguyen, C.D., Cios, K.J.: GAKREM: A novel hybrid clustering algorithm. Information Sciences 178(22), 4205–4227 (2008)
5. Kao, Y.-T., Zahara, E., Kao, I.W.: A hybridized approach to data clustering. Expert Systems with Applications 34(3), 1754–1762 (2008)
6. Khan, S.S., Ahmad, A.: Cluster center initialization algorithm for K-means clustering. Pattern Recognition Letters 25(11), 1293–1302 (2004)
7. Karaboga, D., Basturk, B.: On the performance of artificial bee colony (ABC) algorithm. Applied Soft Computing 8(1), 687–697 (2008)
8. Zhang, C., Ouyang, D., Ning, J.: An artificial bee colony approach for clustering. Expert Systems with Applications 37(7), 4761–4767 (2010)
9. Rocca, P., Oliveri, G., Massa, A.: Differential Evolution as Applied to Electromagnetics. IEEE Antennas and Propagation Magazine 53(1), 38–49 (2011)
10. Abbasgholipour, M., et al.: Color image segmentation with genetic algorithm in a raisin sorting system based on machine vision in variable conditions. Expert Systems with Applications 38(4), 3671–3678 (2011)

# Computer-Aided Dental Caries Detection System from X-Ray Images

Abdolvahab Ehsani Rad, Ismail Bin Mat Amin,
Mohd Shafry Mohd Rahim[*], and Hoshang Kolivand

MaGIC-X (Media and Games Innovation Centre of Excellence) UTM-IRDA Digital
Media Centre Universiti Teknologi Malaysia 81310 Skudai Johor MALAYSIA
`shafry@utm.my`

**Abstract.** The early detection of disease is one of important matter of diagnostic imaging. In this paper we developed a system to analysis the dental x-ray images and diagnosis the tooth which has abnormalities of caries. Enhancement applied to improve the quality of x-ray images and Thresholding method performed to simplify the images. Segmentation has been done by applying the integral projection technique to extract the individual tooth and therefore feature map of tooth surface generated to analysis and detection process. Nevertheless, experiments show the accurate segmentation and caries detection with proposed method which achieves high accuracy and promising result.

**Keywords:** caries detection, segmentation, feature extraction.

## 1 Introduction

Detection of dental caries is important for the diagnosis and treatment planning of the dental disease, which has been affecting a very large population throughout the globe [5]. Dental x-rays are a valuable imaging exam that can provide your dentist with detailed evaluations of your teeth and gums for diagnostic and prevention purposes [10]. Detection of dental diseases is decided on the basis of certain criteria, such as based on whether the lesion is within the enamel, dentin or whether it touches the pulp.

Development of computer-aided caries detection and diagnosis algorithm has become a priority to help the dentists and specialist for make better and faster decision on diagnosis and furthermore treatment. Using image for Diagnostic needs electronic technology and also medical equipment to capture pictures that exhibit inside of human body. Most common types of medical imaging tests include x-rays, CT scans, scan, digital mammography, MRI and ultrasounds. The early detection of disease is one of important matter of diagnostic imaging. Transferring technique from analog to digital be entitled many advantages to x-ray imaging. Like improvement of image contrast or reduce the noises and other facts of image quality by use of image

---

[*] Corresponding author.

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_23

processing. Radiologists could use computers more effectively to help with diagnosis. Some of the medical imaging applications are such as: Computer-aided diagnosis systems, tumor imaging and treatment, angiography, bone strength and osteoporosis and also tortuosity [4]. In dental imaging and diagnosis applications, x-rays are recommended because it shows the size, location and condition of teeth and because they are able to detect the presence of cavities before they appear on the tooth's surface. The x-rays are cheaper than other imaging techniques and easy to access [2].

Dentists are now capable to detect and treat caries early, by the use of preventative, non-invasive measures such as cavitation, texture and roughness, opacification and discoloration of the tooth surface [17]. In particular, visual inspection or visual-tactile inspection has very low sensitivity rate; that is, human inspection alone misses a high percentage of caries [13]. Also detecting of some cavities in early stages can't be seen by visual examination and even difficult to detect by human vision system in x-ray images. Dental caries can be classified by location, etiology, rate of progression, and affected hard tissues [20]. Developing an algorithm for detection of caries and measurement of tooth damage has been majority of studies to focus on it [9]. These systems can show the demineralization, which cannot be observe visually. It has recommended that such systems must be used to enhance the visual or visual-tactile examination, which has done by a dentist.

Due to similarity of features in some areas of teeth like root, which is visible in x-ray images and make suspicion to detect the caries of tooth. Thus, one of the important and challenging problems of this research is to find the proper feature extraction or developed method to extract the features of area of tooth which contain the caries and non caries for find the exact area of caries in tooth and eliminate the misjudgments. The process of analysis of such dental images is important in order to help dentist procedures. This process comprises many different steps of image processing such as: image enhancement, segmentation, feature extraction, detection and classification. Figure 1 demonstrates the framework of this dental caries detection algorithm.

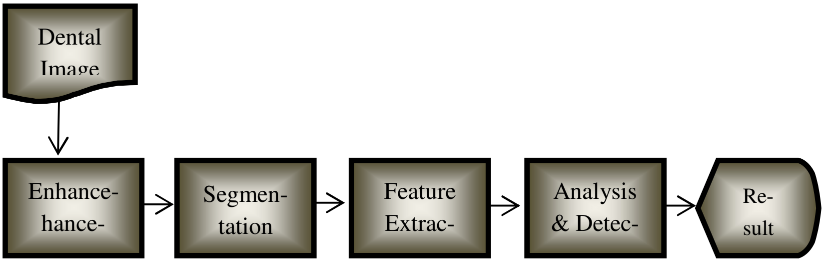
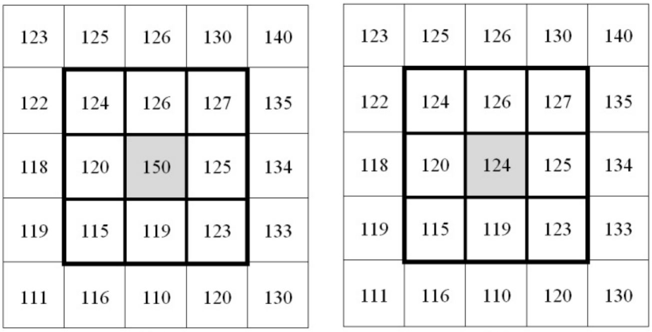


**Fig. 1.** Dental caries detection algorithm framework

## 2 Enhancement

Image enhancement algorithms are used to reduce image noise and increase contrast of structure of interest [8]. In image where the distinction between normal and abnormal tissue is subtle, accurate interpretation may become difficult if noise levels are relatively high [1]. Mostly the dental X-ray images have low contrast and noisy which will affect the segmentation process. Gum and teeth area has same similarity intensity. To distinguish the boundaries between gum and teeth the enhancement by utilizing histogram equalization performed. Furthermore, to remove the image noises causes by device the median filtering applied to achieve the better quality of images.



| | | | | |
|---|---|---|---|---|
| 123 | 125 | 126 | 130 | 140 |
| 122 | 124 | 126 | 127 | 135 |
| 118 | 120 | 150 | 125 | 134 |
| 119 | 115 | 119 | 123 | 133 |
| 111 | 116 | 110 | 120 | 130 |

(a)

| | | | | |
|---|---|---|---|---|
| 123 | 125 | 126 | 130 | 140 |
| 122 | 124 | 126 | 127 | 135 |
| 118 | 120 | 124 | 125 | 134 |
| 119 | 115 | 119 | 123 | 133 |
| 111 | 116 | 110 | 120 | 130 |

(b)

**Fig. 2.** Calculating the median value of neighborhood pixels, a) original pixel values, b) pixel values after one median filtering for current pixel

Neighborhood values in sorted numerical order:

115, 119, 120, 123, 124, 125, 126, 127, 150
Median value: 124

Imaging difficulties are more common in medical imaging due to device limitations such as producing low contrast images or noisy image and most of the applications suffer from imaging difficulties. Chan [3] introduced the salt and paper noise removal using median filtering. The median filter examines each pixel in image and scans to its nearest neighbors to change the pixel value base on surroundings. The pixel simply replaced with median of neighboring pixel values.

Figure 2 demonstrates an example on median filter calculation with considering pixel value 150 as current pixel place. The median pixel value here is 124, therefore, the pixel value 150 will be replaced with 124. The current pixel value is 150 and it is represented in numerical order along neighboring pixel values. Here square neighborhood 3×3 is considered which can be different based on image and algorithm. Figure 3 represents the dental x-ray image after applying the median filtering noise removal with neighborhood value 4×4 which shows the result is smoother than the original image.

**Fig. 3.** Median filtering for image enhancement

## 3 Segmentation

Segmentation of dental image analysis means, extracting teeth or particular tooth form the image background it my inclusive the gum and jaw. Each tooth or object extracted from image represents region of interest (ROI) that encompass important data used for later steps [15].

Jain [6] proposed a semi-automatic contour extraction method for tooth segmentation by using integral projection and Bayes rule, in which the integral projection is semi-automatically applied for tooth isolation since an initial valley gap point is required. Zhou and Abdel-Mottaleb [21] presented a segmentation method that consists of three steps: image enhancement, region of interest localization, and tooth segmentation by using morphological operations and Snake method. Nomir, and Abdel-Mottaleb [11] developed a fully automated approach based on iterative thresholding and adaptive thresholding for dental X-ray image segmentation. Keshtkar and Gueaieb [7] introduced a swarm-intelligence based and a cellular-automata model approach for segmenting dental radiographs. Said et al. [18] offered a mathematical morphology approach to the problem of teeth segmentation, which used a series of morphology filtering operations to improve the segmentation, and then analyzed the connected components to obtain the desired region of interests (ROIs). Shuo et al. [19] proposed a semi-automatic lesion detection framework by using two coupled level set functions in which initial contour are derived from a trained support vector machine to detect areas of lesions from dental X-ray images.

Dental radiographs images which known as X-ray images have many variations of intensity distribution in regions from image to another image due to difficulties which mentioned before. There are three different regions according to intensity distribution: the lowest intensity regions, corresponds to the background or soft tissues. The average intensity regions, corresponds to bone area and the highest intensity regions which corresponds to the teeth area. There are regions with over exposes intensity on teeth area which corresponds to the filling tooth with artifacts. But in some images the intensity of the teeth area is quite close to the bone area, which makes the segmentation process with more difficulty.

Thresholding is the simplest and fastest pixel-based method. There are many techniques in thresholding [12, 16]. The simplest technique in thresholding is to partitioning the image histogram into two areas and assigning the single global threshold "T". In this case Otsu's thresholding method performed which chooses the threshold to minimize the intraclass variance of the black and white pixels and labeling each pixel as the foreground or background. Labeling is based on value of pixel in gray level, whether it is greater than "T" value or lesser than it [14].

$$g(x) = \begin{cases} r_0 & if\ f(x) \geq T \\ r_1 & otherwise \end{cases}$$

Thresholding produces the binary image which simplifies the image analysis for next step.

As suggested by Jain and Chen [6], individual teeth can be isolated by the integrated intensity value as shown in Figure 4 the integrated intensity values accumulate the intensity values of pixels along the vertical direction. Since the teeth usually yield higher intensity values than the jaws and other tissues, the gap of teeth will have a very low value on the integrated intensity value profile. However, unlike a dental forensic X-ray analysis, which can be assumed to have certain orientation, the clinical dental X-rays used to detect lesions, etc., could have any orientation.

The different of intensities in regions on radiographs in most of the images gives the advantage of separating the individual region of each tooth. The horizontal line can separate the region of tooth from top to bottom. This can be achieves by using the horizontal projection as follows:

Let g(i,j) be the m×n (dimension) binary image obtained from the thresholding step, the horizontal integral projection is:

$$H(i) = \sum_{j=1}^{n} g(i,j)$$

To separate the upper and lower regions in the teeth by straight line, the minimum value of accumulated points in horizontal direction of image considered as top of the tooth and the maximum value of accumulated points in horizontal direction of image considered as the bottom line of teeth.

Separation of individual tooth regions has been done using same integral projection process. The aim is to find the lines that separate the adjacent tooth. This achieved by using the vertical integral projection method. If g(i,j) be the m×n (dimension) binary image obtained from the thresholding step, the vertical integral projection is:

$$V(j) = \sum_{i=1}^{n} g(i,j)$$

Locating the area between teeth is possible because of different intensity value of teeth area and other area. To separate the vertical lines between the each tooth the image divided into the 5 areas and process will be done by finding the valleys in the result of vertical projection. The reason of dividing the image in 5 areas is the type of radiographs which is periapical dental radiograph and mostly the numbers of teeth in this type are under 4 teeth. The extraction of each individual tooth has some conditions such as the a) width of each region to eliminate the extraction of non-completed tooth or teeth which located in the contours of radiographs and not clear or not completed as one tooth. b) Eliminate the extraction of missing tooth places. Figure 3 shows the individual tooth separation lines in vertical and horizontal status.

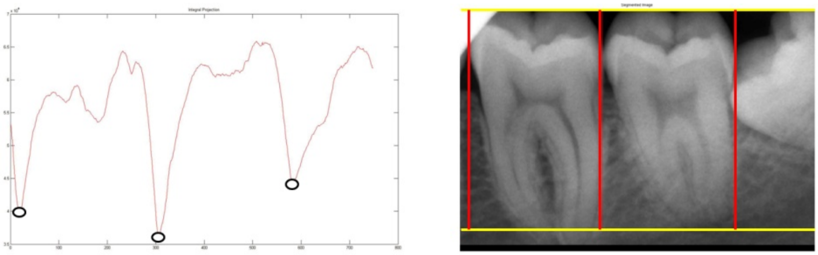


**Fig. 4.** Vertical integral projection with selected points for vertical lines (left). Straight lines to separate each tooth (right)

## 4 Feature Extraction

Process of computing the texture feature is known as feature extraction. Different features are chosen to describe different properties of the image. These features can be utilized for later applications such as identification systems or diagnosis systems. Each isolated tooth will be considered as one object for detection process. Features of selected region for detection will be calculated in order to find that whether the ROI has caries of not. The features here extracted from ROI are: intensity average, entropy and etc. extracted features will be compared with the neighboring regions and according to the caries features, condition of selected region for examine will be monitored.

**Fig. 5.** Selected region on tooth surface No. 1 which contains caries No. 2 region without caries

Figure 6 demonstrates the tooth image with selected region which contains caries and region without caries. Furthermore, Table 1 show the difference between features of caries area and non-caries are on tooth surface.

**Table 1.** Some extracted features of caries area and non caries area

| Feature | Caries Area | Non-Caries Area |
|---|---|---|
| **Intensity** | Min intensity= 48<br>Max intensity=81 | Min intensity= 16<br>Max intensity= 255 |
| **Mean** | 9.0380 | 1.5611 |
| **Entropy** | 6.206 | 5.411 |
| **Perimeter** | Mean of Perimeter= 1.5824 | Mean of Perimeter= 9.0380 |
| **Energy** | 0.165512908 | 0.402948120 |

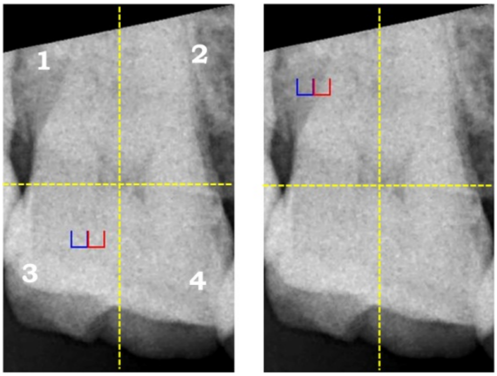


**Fig. 6.** Selected region on tooth surface, which image on left without caries and image on right contains caries

## 5 Analysis and Detection

After extracting the features of tooth surface area, selected region for examine will be compare with the nearest region of tooth surface according to feature map. In this section the tooth surface will be divided to four areas to determine the position of examination region. If the region of examine located in areas 1 or 3 the nearest area for comparison will be the right side of region and if the region of examine located in areas 2 or 3 the nearest area for comparison will be the left side of region. According to the features of caries area which determined by dental expertise the regions will be examined for caries detection process. Furthermore, based on similarity of features in caries area and tooth root, if the selected region located in the middle of tooth, it will consider as pulp or root of tooth. Seriousness or type of caries also can be determine according to the region of detected caries in each area, whether it is near to the gum or near to the tooth enamel. Figure 6 shows selected area and nearest area selected for comparison in order to detect the caries.

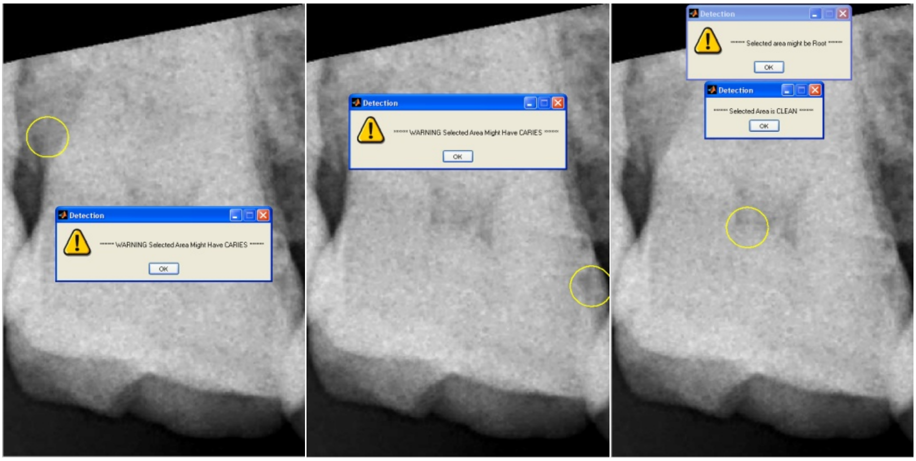
## 6 Result and Discussion

The images for this study are periapical or bitewing dental x-rays, which are the close-up views of a few individual teeth including the root and surrounding bone. Periapical X-rays is very useful in diagnosing a caries, abscess, impacted tooth or bone loss due to periodontal disease. The data images collected from UTM Health Centre, Dental Clinic.

The proposed method have been applied on the dataset of 32 gray level digital periapical dental radiograph which collected form medical center of Universiti Teknologi Malaysia (UTM). Each radiograph images is in dimension of 512×748, the last 12 pixel in vertical direction contains the details of each image which has been removed during the processing steps. The Otsu's Thresholding method result is quite satisfactory to eliminate the background of images to use in extraction of individual tooth. But some images which contain more noises or contain over exposes will appear white in result of thresholding even it belongs to the background. However the thresholding result is fair enough to next step which is for extraction of tooth regions. The extractions of individual tooth in horizontal and vertical directions have been done. For locating the vertical lines the image divided into 5 equal column sections (150 pixels) and scanning process applied within these sections. In some cases the radiographs have many problems which mentioned before and due to these difficulties the tooth extraction may give wrong result. Figure 7 shows the wrong results of some false positive tooth extractions.

The extracted individual tooth considered as input for the level set method, level set method with predefined initial contour (20 pixels space from bottom and top and 5 pixels space from left and right) applied. Because that the images in this step are extracted tooth with different dimensions there is not exact boundary for initial contour. The experiment of proposed method shows that the segmentation of extracted tooth is quite robust. The problem may occur during segmentation with level set when the selected initial contour is wrong or radiograph image has more noise. Figure 8 demonstrates the final result of caries detection system.

Fig. 7. Some wrong false positive segmentations



****** WARNING Selected Area Might Have CARIES ******

****** WARNING Selected Area Might Have CARIES ******

****** Selected area might be Root ******

****** Selected Area is CLEAN ******

(a) (b) (c)

**Fig. 8.** Result of caries detection system. (a) and (b) caries detected. c) detected region is clear and its root area

## 7 Conclusion

In this paper we present a method for automatic segmentation and feature extraction for dental x-ray images. The proposed method has been implemented using traditional image processing techniques, by using clustering k-mean method for segmentation, after image enhancement and illustrate contour for teeth to complete the segmentation step. Furthermore, we extracted some features of dental x-ray images using texture statistics techniques by gray-level co-occurrence matrix. The experimental results show that it is a promising technique for segmentation, but need improvements. Extracted data can be perform to obtain the teeth measurements for automatic dental applications such human identification or dental diagnosis systems. Preparatory experiments show the significance of the proposed method to extract teeth from an x-ray image.

For the future work a better solution to distinguish each particular tooth in segmentation step and evaluation of segmentation methods are expected. However, this paper's method doesn't need to separate the jaws to find the teeth. In addition, it was developed a procedure to recognize the tooth boundary and eliminate other tissues.

**Acknowledgement**. This research is supported and collaboration with Research Management Center (RMC), Universiti Teknologi Malaysia (UTM). This paper is financial supported by Ministry of Higher Education (MOHE),Vot R.J130000.7828.4F491.

## References

1. Bankman, I.: Handbook of medical image processing and analysis. Academic Press (2008)
2. Carroll, S., Frojdh, C.: Dental x-ray imaging system. Google Patents (June 11, 2002)
3. Chan, R.H., Ho, C.-W., Nikolova, M.: Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. IEEE Transactions on Image Processing 14(10), 1479–1485 (2005)
4. Dougherty, G.: Digital Image Processing for Medical Applications. Cambridge University Press, Cambridge (2009)
5. Ikeguchi, E., Nagase, Y., Omura, I., Shibatani, K., Yamauchi, J.: Dental caries detection. Google Patents (August 31, 1982)
6. Jain, A.K., Chen, H.: Matching of dental X-ray images for human identification. Pattern Recognition 37(7), 1519–1532 (2004)
7. Keshtkar, F., Gueaieb, W.: Segmentation of dental radiographs using a swarm intelligence approach. In: Canadian Conference on Electrical and Computer Engineering, CCECE 2006, pp. 328–331 (2006)
8. Kolivand, H., Shahrizal Sunar, M.: A Survey of Shadow Volume Algorithms in Computer Graphics. IETE Technical Review 30(1) (2013)
9. Kutsch, V.K.: Caries Detection, Inside Dentistry. AEGIS Communications (2011)
10. Molteni, R.: Direct digital dental x-ray imaging with Visualix/VIXA. Oral surgery, Oral Medicine, Oral Pathology 76(2), 235–243 (1993)
11. Nomir, O., Abdel-Mottaleb, M.: Human identification from dental X-ray images based on the shape and appearance of the teeth. IEEE Transactions on Information Forensics and Security 2(2), 188–197 (2007)
12. Norouzi, A., Rahim, M.S., Altameem, A., Saba, T., Rad, A.E., Rehman, A., et al.: Medical Image Segmentation Methods, Algorithms, and Applications. IETE Technical Review 31(3), 199–213 (2014)
13. Olsen, G.F., Brilliant, S.S., Primeaux, D., van Najarian, K.: An image-processing enabled dental caries detection system. In: ICME International Conference on Complex Medical Engineering, CME 2009, pp. 1–8 (2009)
14. Otsu, N.: A threshold selection method from gray-level histograms. Automatica 11(285-296), 23–27 (1975)
15. Rad, A., Mohd Rahim, M., Rehman, A., Altameem, A., Saba, T.: Evaluation of Current Dental Radiographs Segmentation Approaches in Computer-aided Applications. IETE Technical Review 30(3), 210 (2013)
16. Rahim, M.S., Norouzi, A., Amin, I.B., Rad, A.E.: Current Methods in Medical Image Segmentation and Its Application on Knee Bone. Archives Des Sciences 65(9) (2012)

17. Roberson, T., Heymann, H.O., Swift Jr., E.J.: Sturdevant's art and science of operative dentistry. Elsevier Health Sciences (2006)
18. Said, E.H., Nassar, D.E., Fahmy, G., Ammar, H.H.: Teeth segmentation in digitized dental X-ray films using mathematical morphology. IEEE Transactions on Information Forensics and Security 1(2), 178–189 (2006)
19. Shuo, L., Fevens, T., Krzyzak, A., Jin, C., Li, S.: Semi-automatic computer aided lesion detection in dental X-rays using variational level set. Pattern Recognition 40(10), 2861–2873 (2007)
20. Sonis, S.T.: Dental Secrets. Hanley & Belfus (2003)
21. Zhou, J., Abdel-Mottaleb, M.: A content-based system for human identification based on bitewing dental X-ray images. Pattern Recognition 38(11), 2132–2142 (2005)

# Multimodal Fusion: Gesture and Speech Input in Augmented Reality Environment

Ajune Wanis Ismail and Mohd Shahrizal Sunar

UTM-IRDA Digital Media Centre
MaGIC-X (Media and Games Innovation Centre of Excellence)
Universiti Teknologi Malaysia, 81310 Skudai Johor, Malaysia
`{Ajune,shahrizal}@utm.my`

**Abstract.** Augmented Reality (AR) has the capability to interact with the virtual objects and physical objects simultaneously since it combines the real world with virtual world seamlessly. However, most AR interface applies conventional Virtual Reality (VR) interaction techniques without modification. In this paper we explore the multimodal fusion for AR with speech and hand gesture input. Multimodal fusion enables users to interact with computers through various input modalities like speech, gesture, and eye gaze. At the first stage to propose the multimodal interaction, the input modalities are decided to be selected before be integrated in an interface. The paper presents several related works about to recap the multimodal approaches until it recently has been one of the research trends in AR. It presents the assorted existing works in multimodal for VR and AR. In AR, multimodal considers as the solution to improve the interaction between the virtual and physical entities. It is an ideal interaction technique for AR applications since AR supports interactions in real and virtual worlds in the real-time. This paper describes the recent studies in AR developments that appeal gesture and speech inputs. It looks into multimodal fusion and its developments, followed by the conclusion.This paper will give a guideline on multimodal fusion on how to integrate the gesture and speech inputs in AR environment.

**Keywords:** Augmented Reality, Multimodal Fusion, Gesture and Speech Input, User Interaction.

## 1 Introduction

Augmented reality (AR) environment is when the real world and virtual world objects are presented together on a single display [1]. Recently, the AR applications have shown that AR interfaces can enable a person to interact with the real world in ways never before possible [2]. Recently, interaction is a crucial key inVirtual Reality (VR) and AR research area. Traditionally, keyboards and mice are common intermediary between human and machine, in most of interfaces. However, the bottleneck occurs rely on user interaction due to the unnaturalness of the interaction [3]. Many interaction methods and technologies have been proposed towards attempting to

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_24

eliminate this bottleneck. By improving the ways of interacting with computers naturally and intuitively, people started to explore the human forms such speech, and gesture recognition [4]. Human gestures come in many forms, such as, hand gestures, general body gestures and facial expressions [5].The human factors need to be addressed before moving to integrate the modalities [6]. That is motivating people to study and explore multimodal interaction [7]. When it comes to unimodal, however, we usually use only one interface device at a time like typing, clicking the mouse button, speaking, or pointing with a magnetic wand. The ease with which this unimodal interaction allows us to convey our intent to the computer is far from satisfactory. The practical reason can lead to consider the use of multimodal interaction [8]. The task can be more practical and convenient with multimodal inputs. The interaction techniques that combine hand gestures provide a separate complementary modality to speech [9]. Successful embodiment of these modalities into an interface noticeable with the advances in computing and communication has the potential of easing the bottleneck in either VR or AR interfaces [8]. It has also become increasingly evident that the difficulties encountered in the analysis and interpretation of individual modalities may be overcome by integrating them into a multimodal interface. Modalities such as speech, vision-based gesture recognition, eye and facial recognition.

Another drawback of current advanced unimodal is that it lacks robustness and accuracy. Whether they use a stylus or a glove or are vision based, they are still constrained to the recognition of few predefined hand movements and are burdened by cables or strict requirements on background and camera placement [10]. However, concurrent use of two or more interaction modalities may loosen the strict restrictions needed for accurate and robust interaction with the individual modes. For instance, spoken words can affirm gestural commands, and gestures can disambiguate noisy speech. Gestures that complement speech, on the other hand, carry a complete communicational message only if they are interpreted together with speech and, possibly, gaze. The use of such multimodal messages can help reduce the complexity and increase the naturalness of the multimodal interface [11].

In the wide studies in AR area, at the early stage people however pay less attention on porting these modalities into AR. One of the most important research areas in AR is creating appropriate interaction techniques for AR applications to allow users to seamlessly interact with virtual content [3]. Many different interaction methods have been explored including natural gestures [8] and they started to look thoroughly into multimodal fusion. In multimodal interaction, users invite the hand gesture and speech input to imitate manipulation tasks in the real world either direct or indirect ways [10]. Thus, in recent years, there has been a tremendous interest in introducing various gesture and speech input into AR that will potentially resolve the user interaction limitation in AR environment. In AR, multimodal considers as the solution to improve the interaction between the virtual and physical entities [9]. It is an ideal interaction technique for AR applications since AR supports interactions in physical and virtual worlds in the real-time. Therefore, it has recently given rise to a number of novel interaction modalities. The multimodal fusion relies on unobtrusive input modalities and natural user interactions. It focuses on providing an intuitive environment, which supports natural interaction with virtual objects while sustaining accessible real tasks and interaction mechanisms. Therefore this paper will discuss the progresses in multimodal fusion in AR involves with gesture and speech input for interaction. The paper presents a few sections to detail out the related works about to recap the multimodal approaches until it recently has been one of the research trends in AR. It describes the recent studies in AR developments that appeal gesture and speech inputs for multimodal.

## 2 Multimodal: VR vs. AR

One of the first multimodal HCI systems can be accredited to Bolt [11], the fusion spoken input and magnetically tracked 3D hand gestures during the integration architecture. The system was used for simple management of a limited set of virtual objects such as selection of objects, modification of object properties, and object relocation. Even though the naturalness of the interaction was hindered by the limitations of the technology at that time, "*Put-That-There*" has remained the inspiration of all modern multimodal interfaces. The rest of this section focuses on some of its descendants. *QuickSet* [12] is a multimodal interface for control of military simulations using handheld PDA's. It incorporates voice and pen gestures as the modes of interaction. This interface belongs to the class of decision level fusers. It follows the [13] with recognition of pen gestures sensed through the PDA is conducted by the gesture agent.

In the past, multimodal interaction has been used not only for 2D user interfaces but also for interacting with 3D virtual contents. Chu et al. [14] showed how multimodal input can be used in VR applications to interact with virtual objects while Krum et al. [15] used it to navigate a virtual world. Laviola et al. [16] developed a prototype multimodal tool for scientific visualization in an immersive virtual environment; a user could not only interact with virtual objects but also navigate through the VR scene by using gesture input from the pinch gloves and triggering corresponding speech input. Wang [17] proposed a multimodal interface with gaze, 3D controller, voicekey and keyboard to select and manipulate the virtual object in the desktop VR environment. As shown in Fig. 2, the Fröhlich [18] meant to chain the multimodal interaction and immersive CAD systems to produce a generic demo for virtual prototyping based on VR technology. Multimodal interaction is concerned with the gesture hand recognition and speech input to drive the modifications of a 3Dvisualization scene. Meanwhile, Immersive CAD is more concerned with the design, exploration and assessment of virtual prototypes using VR simulations. The virtual prototypes are displayed in realistic size like a CAVE with multiple projection system that simplifies on both manipulative gestures for interaction with close and distant virtual parts.

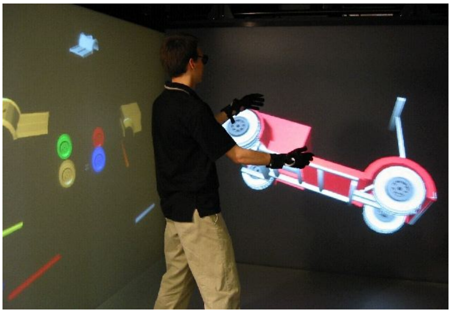


**Fig. 1.** Multimodal in VR prototyping using gesture and speech [18]

Interactive systems featuring multimodal interfaces are becoming widespread. They now cover many different application domains and support a wide variety of users in the performance of their tasks including in AR. Previously AR interface uses general VR interaction techniques, for example a data glove, without modifications. Adopting VR interaction techniques yield gaps between the virtual environment and real world because they only consider interaction techniques useful in VR environments. The functions the hand recognition interface provides limited and users have to wear a marker or to have a fixed hand posture [19]. The rise of AR interfaces development with these issues lead the experts to explore multimodal interaction in AR. Using speech to provide an additional input modality to the gesture in AR interface overcomes the limitations of gesture input alone.There has been some earlier work in applying multimodal in AR applications. Kaiser et al. created *SenseShape*[20] as shown in Fig.3, a multimodal AR interface in which hand to provide visual information about interaction with augmented or virtual objects and speech to provide information where the user wanted to move the object, by using words such as "*this*" or "*that*". However Kaiser et al. [20] also did not conduct user studies to measure the effectiveness of their system. A user must wear a data glove to detect the hand gestures for interaction with objects.

**Fig. 2.** SenseShape a multimodal AR interface in which hand to provide visual information [20]

Heidemann et al. [21] developed a multimodal interface for information retrieval in AR. Hand gestures with speech are adopted to move between menu options. However, the speech input was used to select the menu item that the user wanted to choose, in the same way a mouse did; thus, the system did not use multimodal input fully. Irawati et al., [22] have extended the VOMAR [23] project into multimodal interaction and they verified that combined multimodal speech and paddle gesture input was more accurate than using one modality alone. However, the system could not provide a natural gesture interface for users, and required the use of a paddle with computer vision tracking patterns on it.Lee et al. [24] developed AR multimodal interface with 3D hand vision-based recognition to precise the hand gesture recognition for interaction in AR. They develop an AR multimodal system that allows us to combine gesture and speech input with a multimodal fusion architecture that merges the two different input modalities in a natural way [22]. As presented in Neumann [25] has developed the multimodal AR interface able to remove, manipulate or add communicatively relevant multimodal information in real-time. By using an AR interception proposed by [26] and methodology explored by [27]. Pitsch et al. [28] developed a tool for linguistic studies. They built on the psycholinguistic tradition of experimenting with communicational parameters. Dierker et al. [26] proposed a prototype AR as a novel methodology to investigate human to human interaction within collaborative tasks as shown in Fig. 4. Their goal is to facilitate the recording and analysis of multimodal interaction.

**Fig. 3.** AR Prototype developed to analysis the multimodal interaction in AR [26]

## 3 Multimodal Fusion Levels

The section before we have discussed on multimodal in VR interfaces against the AR interfaces. This section will explain multimodal fusion levels in AR. Generally, data fusion methods are divided in three main categories: first fusion which happens at features levels;second fusion which concerns the intermediate decisions fusion and lastly is thehybrid fusion which is a mixture of the two modalities.During the multimodal fusion, the question mainly comes forward is *why* to integrate or combine these modalities input. *What* are they, the appropriate modalities that are going to integrate respectively? Next, once the desired modalities are selected, need to be addressed on *when* and *how* to combine them.

### 3.1 Decision Levels: *Why* and *What*

Multimodal interactive systems enable users to interact with computers through various input modalities like speech, gesture, and eye gaze. Meanwhile the output channels such as text, graphics, sound, avatars, and probably the synthesized speech. At the first stage to propose the multimodal interaction. The various input modalities like speech, gesture, and eye gaze are decided to be selected before be integrated in an interface. As far as what the modalities are concerned to be selected, the work in this phase is identifying the issues raised by unimodal interfaces and its' limitations. The reasons why multimodal considered as a greater option to improve the burdensome and limitation, remains at a very high level of abstraction more focusing on the identification of problems rather than proposing solutions.

In AR, multimodal considers as the solution to improve the interaction between the virtual and physical entities. It is an ideal interaction technique for AR applications since AR supports interactions in real and virtual worlds at the same time. The most critical concerns associated with multimodal at this decision level are on cases or combination of events that lead to clearly defining what the appropriate modalities are. Machine learning has been already applied to multimodal interfaces, mainly modality recognition like speech and gesture recognition. Multimodal interface is type of user interface which does not only beneficial for enhanced accessibility, but also its usability for greater convenience. For instance, the natural input mode recognition as well as flexibility when the adaptation to context of use, to tasks or to users' preferred interaction modalities. The goal would be to define the interfaces and its fusion that are reliable and usable. Multimodal fusion is commonly known as integration stages for multiple modalities, sometimes also referred to as the fusion engine. It soon will be detailed out in the next section.

### 3.2 Integration Levels: *When* and *How*

The fusion is the key technical challenge for multimodal interaction systems. In general, the meanings of input streams can vary according to context, task, user, and time. Modalities with very different characteristics for instance, speech and eye gaze, facial expression and haptics input, touch-based gesture, they may not have obvious points of similarity and easy ways to combine. Perhaps the most challenging aspect is the temporal dimension. Different modalities may have different temporal constraints and different signal and semantic endurance. Some modalities such as gestures provide information at sparse, discrete points in time while others generate continuous but less time-specific output like the affect. Some modal combinations are intended to be interpreted in parallel, which others may typically be offered sequentially.

When to integrate the modalities inputs is decided on how will computer learning the interaction techniquesaffect its fusion or how will the fusion may affect the interaction system. These questions should be properly addressed by practitioners in the field in order to characterize better the applicationsand problems that multimodal fusion able to improve the conventional unimodal interfaces.

In multimodal interactive systems, multimodal fusion is a crucial step in combining and interpreting the various input modalities. Once the desired modalities are selected, an important question to be addressed is how to combine them. To address this problem, it is helpful to know how the integrating modalities relate in AR environment. Some modalities, like speech and lip movements, are more closely tied than others, such as speech and hand gestures. It is also plausible to assume that integration of such different combinations of modalities should be explored at different levels of integration. Depending on the chosen level of integration, the actual fusion can then be performed using numerous methods, ranging from simple feature concatenation to complex interaction of interface agents.

Unlike unimodal interfaces, multimodal requires having multi-signal fusion architecture to merge two or more input commands in a natural and efficient way. We should have a history of each mode of signal. With the analysis of each signal, statistical characteristics will be obtained. Then, multi-channel signal fusion is available with the provided statistical characteristics. Additionally, environmental context and task context should be considered to provide better recognition result. The main difference between a unimodal interface and a multimodal interface is that the multimodal interface requires multimodal fusion architecture to merge two or more modality input in an efficient and effective way. As presented in Table 1, multimodal fusion systems can be classified in two groups: (1) feature level fusion and (2) semantic level fusion.

**Table 1.** Classification of multimodal fusion on *how* to integrate modalities

| Feature | Semantic |
|---|---|
| Fusion is finished before the input signals are sent to their respective recognizers | Fusion is finished after the signals are interpreted from their respective recognizers |
| Input signals are complex to model | Interpret the input signals independently |
| Difficult to train required high data training | Easy to train with existing unimodal data |

Feature level fusion is done before the input signals are sent to their respective recognizers. Feature level fusion is considered as a good strategy for integrating the closely coupled and synchronized input signals, for example, lip movement and speech input whose signals correspond to each other. Typical drawbacks of the feature level fusion are that it is complex to model, intensive to compute, and difficult to train. Mostly, feature level fusion requires a large amount of training data.

Semantic level fusion is done after the signals are interpreted from their respective recognizers. Semantic level fusion is appropriate for integrating two or more signals which provide complementary information, such as, speech and pen input. Individual recognizers are used to interpret the input signals independently. Those recognizers can be trained with existing unimodal training data. Therefore, input channels needed to have complementary information to each other and time-stamp played an important role to match two different modalities for integration. Semantic level fusion is that semantic representation of the recognized input was essential for multimodal fusion and that mutual disambiguation was necessary to improve error handling and resolution.

## 4 Conclusion

There are numerous potential benefits in integrating multiple modalities. The reasons range from the fact that natural human interaction itself. The interaction of humans with their environment including with other humans, is naturally multimodal. The human factors need to be addressed before moving to integrate the modalities. That is motivating people to study and explore multimodal interaction. When it comes to unimodal, however, we usually use only one interface device at a time like typing, clicking the mouse button, speaking, or pointing with a magnetic wand. The ease with which this unimodal interaction allows us to convey our intent to the computer is far from satisfactory. The practical reason can lead to consider the use of multimodal interaction. The task can be more practical and convenient with multimodal inputs.

Drawback of current advanced unimodal is that it lacks robustness and accuracy. Whether they use a stylus or a glove or are vision based, they are still constrained to the recognition of few predefined hand movements and are burdened by cables or strict requirements on background and camera placement. Gestures that complement speech, on the other hand, carry a complete communicational message only if they are interpreted together with speech and, possibly, gaze. The use of such multimodal messages can help reduce the complexity and increase the naturalness of the multimodal interface. In this studies we have explored on multimodal fusion in AR. Multimodal has been a topic of research in AR since decades. Though studies have been conducted to establish the feasibility of these novel modalities using appropriate sensing and interpretation techniques, their role is still being explored to compare the partial-immersive AR system against the fully-immersive VR systems. On the first section of this paper we have described about multimodal in general. Second section later has explained the multimodal in VR against the multimodal in AR environments. We have detailed out the previous works and researches have been done in multimodal that invites modalities gesture and speech as inputs. In Next section we identified the multimodal fusion in AR. We found fusion levels in dealing with multimodal in AR.

**Acknowledgments.**We would to express our appreciation to Universiti Teknologi Malaysia (UTM). We also thank the Human Interface Technology Laboratory New Zealand (HITLabNZ) at University of Canterbury. This work was funded by Ministry of Science, Technology and Innovation through eScieneFund Grant 01-01-06-SF1188

## References

1. Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., Blair, M.: Recent advances in augmented reality. IEEE Computer Graphics and Applications, 20–38 (2001)
2. Zhou, F., Duh, H.B.-L., Billinghurst, M.: Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. In: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality. IEEE Computer Society (2008)
3. Kaiser, E., Olwal, A., McGee, D., Benko, H., Corradini, A., Li, X., Cohen, P., Feiner, S.: Mutual disambiguation of 3D multimodal interaction in augmented and virtual reality. In: International Conference on Multimodal Interfaces, ICMI 2003, pp. 12–19 (August 2003)
4. Corradini, A., Cohen, P.: On the Relationships among Speech, Gestures, and Object Manipulation in Virtual Environments: Initial Evidence. In: Proceedings of the International CLASS Workshop on Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems, pp. 52–61 (2002)
5. Mitra, S., Acharya, T.: Gesture recognition: A survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 37(3), 311–324 (2007)
6. Lim, C.J., Pan, Y., Lee, J.: Human Factors and Design Issues in Multimodal (Speech/Gesture) Interface. JDCTA 2(1), 67–77 (2008)
7. Jewitt, C.: Technology, literacy and learning: A multimodal approach. Psychology Press (2006)
8. Haller, M., Billinghurst, M., Thomas, B.H. (eds.): Emerging technologies of augmented reality: interfaces and design. IGI Global (2007)
9. Irawati, S., Green, S., Billinghurst, M., Duenser, A., Ko, H.: An evaluation of an augmented reality multimodal interface using speech and paddle gestures. In: Pan, Z., Cheok, D.A.D., Haller, M., Lau, R., Saito, H., Liang, R. (eds.) ICAT 2006. LNCS, vol. 4282, pp. 272–283. Springer, Heidelberg (2006)
10. Sharma, R., Pavlovic, V.I., Huang, T.S.: Toward multimodal human-computer interface. Proceedings of the IEEE 86(5), 853–869 (1998)
11. Richard, A.: Bolt: Put-That-There: Voice and Gesture at the Graphics Interface. In: Proceedings of the International Conference on Computer Graphics and Interactive Techniques, vol. 14, pp. 262–270 (1980)
12. Cohen, P.R., Johnston, M., McGee, D.R., Oviatt, S.L., Pittman, J.A., Smith, I., Chen, L., Clow, J.: Quickset: Multimodal Interaction for Distributed Applications. In: Proceedings of the Fifth Annual International Multimodal Conference, pp. 31–40 (1997)
13. Nicholson, M., Vickers, P.: Pen-Based gestures: An approach to reducing screen clutter in mobile computing. In: Brewster, S., Dunlop, M.D. (eds.) Mobile HCI 2004. LNCS, vol. 3160, pp. 320–324. Springer, Heidelberg (2004)
14. Chu, C., Dani, T., Gadh, R.: Multimodal Interface for a virtualreality based computer aided design system. In: Proceedings of 1997 IEEE International Conference on Robotics and Automation, vol. 2, pp. 1329–1334 (1997)
15. Krum, D.M., Omotesto, O., Ribarsky, W., Starner, T., Hodges, L.F.: Speech and gesture control of a whole earth 3D visualization environment. In: Proceedings of Joint Eurographics-IEEE TCVG Symposium on Visualization, pp. 195–200 (2002)
16. LaViola, J.: MSVT: A virtual reality-based multimodal scientific visualization tool. In: Proceedings of the Third IASTED International Conference on Computer Graphics and Imaging, pp. 221–225 (2000)

17. Wang, J.: Integration of eye-gaze, voice and manual response in multimodal user. In: Proceedings of IEEE International Conference on Systems, Man and Cybernetic, vol. 5, pp. 3938–3942 (1995)
18. Fröhlich, C., Biermann, P., Latoschik, M.E., Wachsmuth, I.: Processing Iconic Gestures in a Multimodal Virtual Construction Environment. In: Sales Dias, M., Gibet, S., Wanderley, M.M., Bastos, R. (eds.) GW 2007. LNCS (LNAI), vol. 5085, pp. 187–192. Springer, Heidelberg (2009)
19. Rautaray, S.S., Agrawal, A.: Vision based hand gesture recognition for human computer interaction: a survey. Artificial Intelligence Review, 1–54 (2012)
20. Olwal, A., Benko, H., Feiner, S.: SenseShapes: Using Statistical Geometry for Object Selection in a Multimodal Augmented Reality System. In: Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2003), Tokyo, Japan, October 7-10, pp. 300–301 (2003)
21. Heidemann, G., Bax, I., Bekel, H.: Multimodal Interaction in an Augmented Reality Scenario. In: Proceedings of International Conference on Multimodal Interfaces, ICMI 2004, pp. 53–60 (2004)
22. Irawati, S., Green, S., Billinghurst, M., Duenser, A., Ko, H.: An evaluation of an augmented reality multimodal interface using speech and paddle gestures. In: Pan, Z., Cheok, D.A.D., Haller, M., Lau, R., Saito, H., Liang, R. (eds.) ICAT 2006. LNCS, vol. 4282, pp. 272–283. Springer, Heidelberg (2006)
23. Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K., Tachibana, K.: Virtual Object Manipulation on a Table-Top AR Environment. In: Proceedings of the International Symposium on Augmented Reality, pp. 111–119 (2000)
24. Lee, M.: Multimodal Speech-Gesture Interaction with 3D Objects in Augmented Reality Environments (2010)
25. Neumann, A., Schnier, C., Hermann, T.: &Pitsch, K. Interaction Analysis and Joint Attention Tracking In Augmented Reality. In: Proceedings of the 15th ACM International Conference on Multimodal Interaction, pp. 165–172 (2013)
26. Dierker, A., et al.: Mediated attention with multimodal augmented reality. In: Proceedings of the 2009 International Conference on Multimodal Interfaces. ACM (2009)
27. Neumann, A.: Design and implementation of multi-modal AR-based interaction for cooperative planning tasks. Bielefeld University, Bielefeld (2011)
28. Pitsch, K., Neumann, A., Schnier, C., Hermann, T.: Augmented reality as a tool for linguistic research: Intercepting and manipulating multimodal interaction. In: Multimodal Corpora: Beyond Audio and Video (IVA 2013 Workshop), pp. 23–29 (2013)

# Learning to Play Tetris from Examples

Somnuk Phon-Amnuaisuk

Media Informatics Special Interest Group,
School of Computing & Informatics,
Institut Teknologi Brunei, Brunei Darussalam
`somnuk.phonamnuaisuk@itb.edu.bn`

**Abstract.** We model a Tetris player using Artificial Neural Network (ANN). In contrast to most previous works which learned weights of predefined heuristic functions, our model learns the actual actions i.e., given a board state and a tetrominoe, where the piece should be placed on the board. The problem is formulated as a classification problem where the model learns to associate a board state with fruitful actions. We compare an ANN player with a random player and a Best First Search (BFS) player. The random player and the BFS player provide baselines for uninformed search and informed search, while the ANN player learns from the examples extracted from BFS runs and its performance lies between both baselines. We observe that generalisation appears to be very hard from the nature of the game itself. Fitting Tetrimino pieces together demands an exact match. This means we cannot expect similar solutions to closely related patterns. In this paper, we explore the problem from a soft computing perspective, present our experimental design and provide a critical discussion of the results of our experiment.

**Keywords:** Modeling Tetris players, Artificial neural networks.

## 1 Introduction

Tetris[1] is a well-known tile-jigsaw puzzle game created by Pajitnov in 1984 [1]. There are 7 kinds of tokens (aka tetrominoes) in Tetris, namely *I, J, L, O, T, S* and *Z*. The tetrominoes are named according to their similarity to the corresponding alphabets (see Figure 1). Tetris has a simple gameplay: random tetrominoe pieces fall from the top of the $20 \times 10$ board one piece at a time. These tetrominoes can be rotated at a 90-degree step (either clockwise or counter clockwise) while it is falling. Score is awarded for clearing complete horizontal lines (i.e., full width) of the Tetris board. Although Tetris has a simple gameplay, the plausible game states of the Tetris game is very large. The upper bound of the Tetris gamestates played on a $20 \times 10$ tiles board is $7 \times 2^{200}$ states. This is a complex problem and it is argued to be an NP complete problem [2].

Due to its computational complexity, despite a simple gameplay, the Tetris game has become a popular board game that captures the attention of many

---

[1] Read the article *Tetris* in *Wikipedia: The Free Encyclopedia for more information.*

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_25

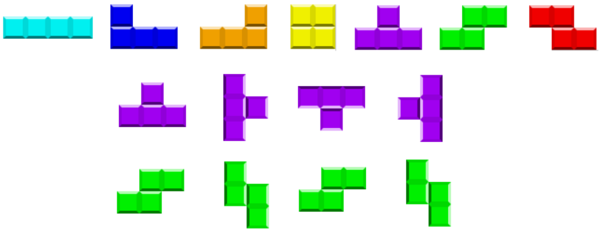


**Fig. 1.** Top row: All the seven tetrominoes *I, J, L, O, T, S* and *Z*; Second and Third rows: The four variations after 0, 90, 180, and 270 degree clockwise rotations of the *T* tetrominoe and the *S* tetrominoe.

AI game researchers. Emulating players' behaviours is one of the main research theme for AI game researchers. A rule-based expert system approach has been proven to be an effective approach to encode expert knowledge in a compact form. However, due to knowledge acquisition bottleneck in the traditional knowledge engineering approach, many AI game researchers have explored other soft computing approaches that perform knowledge acquisition process by learning from training examples.

In this work, we investigate the application of artificial neural network in learning a player's behaviours. We compare and contrast the results from the three player types: an ANN player, a random player (i.e., uninformed search) and a BFS player (i.e., an informed heuristic search). The rest of the materials in the paper are organized into the following sections: Section 2: Related works; Section 3: Problem formulation; Section 4: Experimental results & Discussion; and Section 5: Conclusion.

## 2 Related Works

Simulating chess playing skills is one of the early popular research themes explored by AI researchers (notably, Deep Blue and Hydra [3]). Intelligent chess moves emerge from an inference process based on various rules and heuristics. The simulation of intelligent behaviours using rules can be seen as a search through the search space guided by rules and heuristics. Although intelligent behaviours can be successfully emulated using a compact rule-based model, the approach requires a great deal of knowledge engineering effort and usually have a non-flexible and non-adaptive rules set.

Recent advances in soft computing address issues such as inexact reasoning, adaptive learning and knowledge engineering bottleneck inherited in the good old fashion AI approach. In the game AI scene, we also see the shift in encoding player character (PC) and non-player character (NPC) using complex rules toward modeling PC/NPC using soft-computing techniques [4]. Soft computing techniques such as genetic algorithms, reinforcement learning and neural networks have also been applied to Tetris by [5,6,7,8,9]. However, the majority of the work focuses on modelling high level heuristic functions (i.e., the actual action will later be determined from these heuristic functions by another subsystem). From our knowledge, only [10] has formulated the problem as learning the actual action using SVM.

It is intuitive to model a player's actions based on the information available on the Tetris board. The rules for players' actions are formulated based on the fact that tetrominoes should be placed in such a way that all the empty spaces in each row would be fully filled, or the placement does not create unreachable holes, or other undesired properties. The rules are expected to consider the shape of the tetrominoe at hand and the best way to place it on the board. This concept is common in all previous works. For example, the following properties: the topmost contour formed by the filled tetrominoes, the number of unreachable holes, the maximum and the minimum height of the filled tetrominoes, etc., have been employed to represent the board status (see summary in [11]). These desired board properties can be quantified as a value function $\mathcal{V}(s)$ [12]. The value of a board state $\mathcal{V}(s)$ is commonly expressed as a weighted linear sum of objective functions describing the desired board properties $f_i$ e.g., $\mathcal{V}(s) = \sum_{i=1}^{N} w_i f_i(s)$. The association between the board states and actions can then be determined using various approaches.

A rule-based approach commonly determines the next action using the generate and test tactic. All plausible next states are attempted and the best choice is taken [13]. The performance of the model depends on how far ahead the program look into the future. An evolutionary approach is commonly employed to search for the optimal weight $w_i$ of various heuristic functions $f_i$ [7]. A machine learning approach has also been explored by many researchers. The reinforcement learning technique learns the state and high-level-actions policy by playing many games and learn the association between the high-level-action and the accumulated board value $V(s)$ of the current action and the sequence of future actions [11].

In this work, we explore the use of ANNs to learn a player's action from examples. Patterns of board state and best action are collected from many games played by a BFS player. These patterns are then used to train the ANNs. Modelling a player's actions by learning from examples in this way can be formulated as a classification problem and this is not a common approach in previous works where high level actions such as minimising holes, etc. are modelled instead. Details of our approach are given in the next section.

## 3 Problem Formulation

Since a player makes her decision on how to place a tetrominoe based solely on information available on the board, it is intuitive to associate contour patterns formed by tetraminos to the actions. We approximate the number of possible contour shapes formed by the filled tetrominoes (see Figure 2) to be about $20^{10}$ states [2]. This poses quite a challenge for a machine learning technique aiming

---

[2] The maximum number of contour patterns that each tetrominoeneeds to learn.

to create a player model by learning from training data since (i) there are too many examples to be learned and (ii) there is a problem in preparing the training examples. This also poses a great challenge for an exhaustive search technique to locate an optimal action.

In this work, the neural networks learn a player's model from the training examples generated by calculating contour patterns of the filled tiles and the corresponding actions suggested by the BFS technique [14]. The BFS actions are guided by the heuristics adapted from previous handcrafted rules reported in the literature [8,12,13]. The following assumptions are made to the Tetris gameplay in this implementation:

1. A player only sees the current tetrominoe and the board. The information of the next tetrominoe will only be revealed after the action is made.
2. For each tetrominoe$\tau$, only one action (a combination of rotation and translation) is allowed.
3. No row eliminations is carried out during the game play. This means that a perfect game will require only 50 tetrominoes to fill all 200 tiles. Hence, in each game, a sequence of 50 tetrominoes are generated randomly by uniformly picking them from the set $\{I, J, L, O, T, S, Z\}$.

Given a sequence of tetrominoes $\mathcal{T} = \{\tau_1, \tau_2, ..., \tau_n\}$, the perfect gameplay is the gameplay with a sequence of actions $\mathcal{A} = \{a_1, a_2, ..., a_n\}$ that yields optimal board state i.e., the board with no holes (all tiles are filled).

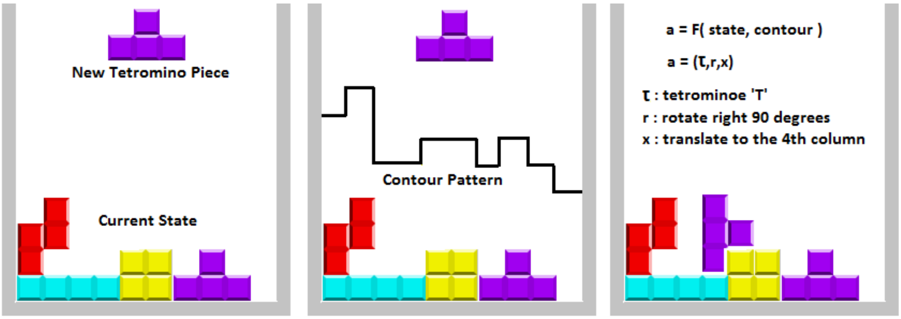


**Fig. 2.** Left: given a new tetrominoe piece; Middle: the contour of the current board is calculated; and Right: an action is determined.

### 3.1 Modelling a Tetris Player

Let us formally define a Tetris game using a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{F}, \mathcal{V})$, where $\mathcal{S}$ is a set of game states; $\mathcal{A}$ is a set of a player's actions; $\mathcal{F}$ is a transition function; and $\mathcal{V}$ is a value function.

**Definition 1.** *Game States $\mathcal{S}$: A state $s \in \mathcal{S}$ is a plausible arrangement of tetrominoes on the Tetris board area. In this implementation, the state $s \in \{0, 1\}^{20\times10}$ is represented as a binary matrix, where the entry 1 denotes a filled tile and the entry 0 denotes an empty tile.*

**Definition 2.** *Player's Actions $a \in \mathcal{A}$: During the gameplay, each new tetrominoe$\tau \in \{I, J, L, O, T, S, Z\}$ is placed on the board at the topmost row. At each time step, the tetrominoe piece falls down one tile due to its gravity. A player may rotate the tetrominoe (i.e., 90, 180, 270 degree) and translate the tetrominoe in the horizontal axis (i.e., move left/right). Here, $\mathcal{A}$ denotes a sequence of actions $\{a_1, a_2, ..., a_n\}$ a player has taken from the start of the game till the end of that game. An action a is expressed as a tuple $(\tau, r, x)$ where $\tau$ indicates the desired tetrominoe, $r \in \{0, 90, 180, 270\}$ indicates the desired rotation and $x \in \{1, 2, ..., 10\}$ indicates the left most position of the tetrominoe.*

**Definition 3.** *Transitional function $\mathcal{F}$: The transitional function $\mathcal{F}(s, a) : s_t \rightarrow s_{t+1}$ maps the current board state $s_t$ to the new board state $s_{t+1}$ as a result of applying a* **valid action** *a. A valid action is an action that conforms to the rules of Tetris.*

**Definition 4.** *Value Function $\mathcal{V}$: The value function $\mathcal{V}(s) : s \rightarrow \mathcal{R}$ maps the board state s to a real value. The value of a state is defined as a weighted sum of various predefined fitness criteria $f_i$.*

$$\mathcal{V}(s) = \sum_{i=1}^{n} w_i f_i$$

*In [8], various features (i.e., fitness criteria) have been proposed. Our fitness criteria share many similarities to those works. The two fitness criteria used here are based on the following heuristics:*

$h_1$ : *prefers to fill tetrominoes at the bottom-most of the board first*
$h_2$ : *prefers to rotate tetrominoes and place the tetrominoes such that the placement does not create unreachable holes*

### 3.2 Three Approaches to Play Tetris

We compare the results of three different player programs that take the following actions: random actions, BFS actions and ANN actions. These actions describe the rotation and the $x$ axis translation and each approach determines its actions as follow:

1. random player: Given a tetrominoe$\tau$, the program choose a valid action $a = (\tau, r, x)$ randomly e.g., rotation $r$ and translation $x$. This provides a baseline of an uninformed search approach.
2. BFS player: Given a tetrominoe$\tau$, the program tries all possible valid actions (i.e., the cartesian product of $\{0, 90, 180, 270\} \times \{1, 2, ..., 10\}$), then chooses the best action that maximises $\mathcal{V}(s)$ using the pseudocode below. This provides a baseline for an informed search approach.

```
function GET-ACTION(TOKEN) return $a$
  Input : $\tau$ is a TOKEN where $\tau \in \{I, J, L, O, T, S, Z\}$
  Output : a = $(\tau, r, x)$;
  Var : $s \in S$; $r \in \{0, 90, 180, 270\}$; $x \in \{1, 2, ..., 10\}$;
        $P$ is a sequence of recorded patterns and
        $T$ is a sequence of recorded actions
  Initialise a list [] $\leftarrow v$;
  for each pair $(r, x) \in \{0, 90, 180, 270\} \times \{1, 2, ..., 10\}$ {
    $s_{t+1} \leftarrow \mathcal{F}(s_t, a)$ where $a = (\tau, r, x)$
    // append the value of the next $s$ and the $a$ to the list v
    v.APPEND($\mathcal{V}(s_{t+1})$:a)
  }
  // Filter out the best action a from the list v
  $s_{t+1}, a \leftarrow FILTER(v)$;
  // Record contour patterns of the current board state $s_t$ and
  // its corresponding target actions $a$
  $P.APPEND(CONTOUR(s_t))$; $T.APPEND(a)$;
  return $a$
```

3. ANN player: Given a tetrominoe $\tau$, the program selects a valid action based on an ANN model. In this implementation, seven ANN models were created, one for each type of tetromino. The ANN models were trained with the contour patterns $P$ (i.e., contour formed by the filled tetrominoes) and the desired targets $T$ (i.e., best action for each input pattern).

## 4 Experimental Results and Discussion

**Experimental Design.** BFS is a useful technique when a system's responses to an action can be explicitly observed. By observing different behavioral responses to different actions applied to the system, the problem can be modeled as a state space search. In most real life problems, the state space is too large and it is impractical to search for the global optimal solution. BFS offers a reasonably good solution by heuristically pruning the unfruitful part of the state space. The quality of the solution often depends on the quality of the heuristics employed to guide search.

In this implementation, the patterns of the contour formed from the filled tetrominoes and the corresponding best action suggested by BFS were recorded. Then, the contour patterns for each tetrominoe type and the corresponding actions were extracted as training data. The training data were used to train different feedforward neural networks. The structure of the ANNs used here were the typical three layer feedforward structure with 10 input nodes, 100 hidden nodes and 2 output nodes. The contour patterns $P$ were fed to the ten input nodes and the target actions $T$ (i.e., translation on $x$-axis and rotation) were compared to the two output nodes where errors were back-propagated to adjust the weights of the ANN model. Important ANN parameters are listed below:

Network structure : 3 layers Feed Forward 10:100:2
Hidden layer transfer functions : Hyperbolic tangent sigmoid
Output layer transfer functions : Linear
Training error goal : 0.1
Learning rate : 0.5 %
Maximum training epochs : 50

To compare the results of the three different players, we decided to impose the following constraints in our implementation (i) the completely filled rows were not removed, hence each game only required a maximum of 50 tetrominoes (50 × 4 equals 200 tiles); (ii) actions were described using $x$-axis translation and rotation only, hence, the representation could not describe some trick moves that require two consecutive transitions or rotations; (iii) all three player types played the same tetrominoe sequences and the sequences were randomly generated by uniformly picking tetrominoes from the set $\{I, J, L, O, T, S, Z\}$; and finally (iv) the merit of each player was quantified based on the number of tiles left unfilled after the game was over i.e., a new tile could not be placed on the board.

Figure 3 shows that the average unfilled tiles after game over of the random player and the BFS player were stable when the number of games was increased.

| # Games | 100 | 500 | 900 | 1300 | 1700 |
|---|---|---|---|---|---|
| Random | 59.5 | 60.8 | 60.3 | 60.2 | 60.3 |
| std | 6.6 | 6.6 | 6.4 | 6.5 | 6.5 |
| ANN | 57.0 | 56.6 | 57.1 | 54.6 | 53.3 |
| std | 8.1 | 7.5 | 6.8 | 6.8 | 7.3 |
| BFS | 19.9 | 19.9 | 19.9 | 19.9 | 19.9 |
| std | 4.1 | 3.6 | 3.6 | 3.6 | 3.6 |

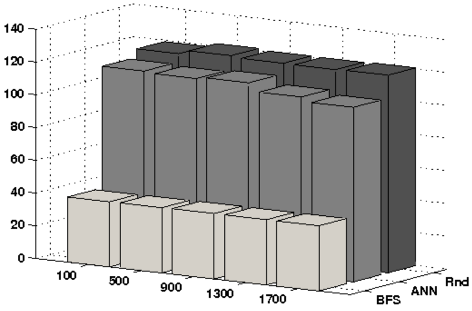


**Fig. 3.** The system played 100, 500, 900, 1300, and 1700 games. The more games played meant more training data was available for training the ANNs and improvement in their performances was observed. Entries on the table are in percentage of unfilled tiles and the height of the bar graph indicates the number of unfilled tiles.

This implies that no learning took place in those players. In contrast, a clear decrease in the number of unfilled tiles was observed in the ANN player.

Although we were convinced that some learning took place in the ANN player, the simulated results gave a less than satisfactory impression on the overall performance. Although the ANN players produced better solutions than those from the random search, huge gaps between the ANN players and the BFS players were observed. The result confirms that the Tetris game is not easy [2,13]. Counting possible contour patterns in a game with all seven types of tetrominoes will remind us about the complexity of this game and the limitation of an approach that learns new knowledge through examples. There are too many examples to learn from, i.e., $10^{20}$ contour patterns. This may be one of the reasons why majority of the previous works have focused on learning the weight of predefined heuristic functions, instead of learning the actual actions.

**Simplify the Problem.** To show that it is the capacity of the ANNs that limits the learning, we could (i) increase the number of games so that more examples are available for the training process; or (ii) reduce the number of patterns by reducing the board size or reducing the number of tetrominoes. We decided to reduce the complexity by playing the game with one kind of tetriminoe per game.

Figure 4 shows the number of average unfilled tiles of the three players. It is clear that, now, the ANN player had successfully learned how to play Tetris and the performance of the ANN player was comparable to the performance of the

| Tetrominoe | I | J | L | O | T | S | Z |
|---|---|---|---|---|---|---|---|
| Random | 62.4 | 61.6 | 61.0 | 49.4 | 60.0 | 60.5 | 60.7 |
| std | 6.9 | 6.2 | 6.3 | 7.8 | 6.3 | 6.3 | 6.6 |
| ANN | 0.0 | 38.5 | 23.0 | 0.0 | 26.5 | 29.0 | 9.0 |
| std | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| BFS | 3.5 | 30.0 | 23.0 | 0.0 | 29.5 | 29.0 | 17.0 |
| std | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

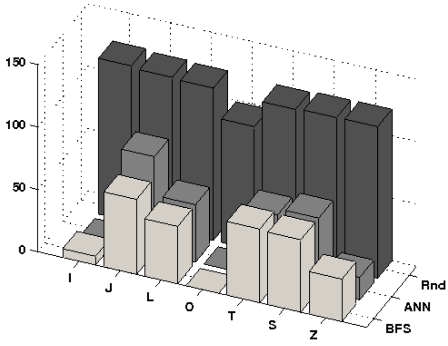


**Fig. 4.** The system played with a single kind of tetrominoe for 500 games (each player played a total of 3500 games for all seven tetrominoes)

BFS player. It is interesting to note that all the players found it easy to play with the *'O'* which was true since all the rotated positions did not create new patterns. The ANN players found it easy to fill up the *'O'*, *'I'* and *'Z'*, *'L'*, *'T'*, *'S'* and *'J'* repectively.

## 5 Conclusion

This work investigates the modelling of a Tetris player using ANN. This transforms the problem to a classification problem where the model learns to associate a board state with actions. It is conclusive from the experiment that the ANN player can learn to play the Tetris game. However, the performance of the player could be disappointing if there are not enough training patterns. By the nature of the game, the solutions of closely related contour patterns may not lie close to each other in the solution space. This implies that generalisation is particularly hard for this problem domain. The wrong move will produce unfamiliar contour patterns (i.e., the game moves to a new area in the search space where training data might not be sampled from that area). Can we ensure that the training data are uniformly sampled from the search space? This question cannot be answered now. To the best of our knowledge, none of the work has attempted to uniformly sample training data from the search space i.e., $10^{20}$ contour patterns. In future work, we will look into this issue and apply deep learning concept to the game.

**Acknowledgement.** We wish to thank anonymous reviewers for their comments, which help improve this paper. We would also like to thank GSR, Institut Teknologi Brunei for their partialfinancial support given to this research.

## References

1. "Alexey Pajitnov" in Wikipedia: The Free Encyclopedia (last retrieved April 14, 2014)
2. Demaine, E.D., Hohenberger, S., Liben-Nowell, D.: Tetris is hard, even to approximate. In: Warnow, T.J., Zhu, B. (eds.) COCOON 2003. LNCS, vol. 2697, pp. 351–363. Springer, Heidelberg (2003)
3. Donninger, C., Lorenz, U.: The chess monster Hydra. In: Becker, J., Platzner, M., Vernalde, S. (eds.) FPL 2004. LNCS, vol. 3203, pp. 927–932. Springer, Heidelberg (2004)
4. Phon-Amnuaisuk, S.: Learning chasing behaviours of non-player characters in games using SARSA. In: Di Chio, C., et al. (eds.) EvoApplications 2011, Part I. LNCS, vol. 6624, pp. 133–142. Springer, Heidelberg (2011)
5. Lundgaard, N., McKee, B.: Reinforcement learning and neural networks for Tetris. Technical Report, University of Oklahoma
6. Carr, D.: Adapting reinforcement learning to Tetris. BSc (Honours) Thesis, Rhodes University (2005)
7. Böhm, N., Kókai, G., Mandl, S.: An Evolutionary Approach to Tetris. In: Proceedings of the Sixth Metaheuristics International Conference (MIC 2005), Vienna, Austria, pp. 1–6 (2005)

8. Boumaza, A.: On the evolution of artificial Tetris players. In: Proceedings of the IEEE Symposium on Computational Intelligence and Games (CIG 2009), pp. 387–393 (2009)
9. Phon-Amnuaisuk, S.: GA-Tetris bot: Evolving a better Tetris gameplay using adaptive evaluation scheme. In: Siah, Y.K. (ed.) ICONIP 2014, Part III. LNCS, vol. 8836, pp. 579–586. Springer, Heidelberg (2014)
10. Zhang, D., Cai, Z., Nebel, B.: Playing Tetris using learning by immitation. In: Proceedings of GAMEON 2010, Leicester, UK, pp. 23–27 (2010)
11. Thiery, C., Scherrer, B.: Building controllers for Tetris. International Computer Game Association Journal 32, 3–11 (2009)
12. Tsitsiklis, J.N., Van Roy, B.: Feature-based methods for large scale dynamic programming. Machine Learning 22, 59–94 (1996)
13. Brzustowski, J.: Can you win at Tetris? Masters Thesis, University of British Columbia (1992)
14. Perl, J.: Heuristics: Intelligence Search Strategies for Computer Problem Solving. Addoson-Wesley (1984)

# Advances in Sports Informatics Research

Martin Sykora[1], Paul Wai Hing Chung[2], Jonathan P. Folland[3],
Benjamin J. Halkon[4], and Eran A. Edirisinghe[2]

[1] Centre for Information Management, School of Business and Economics,
Loughborough University, LE11 3TU, UK
[2] Computer Science Department, Loughborough University, LE11 3TU, UK
[3] School of Sport, Exercise and Health Sciences, Loughborough University, LE11 3TU, UK
[4] Wolfson School of Mechanical and Manufacturing Engineering,
Loughborough University, LE11 3TU, UK
`(m.d.Sykora,p.w.h.chung,j.p.folland,`
`b.j.halkon,E.A.Edirisinghe)@lboro.ac.uk`

**Abstract.** With advances in hardware and software, computer science based technologies within sports sciences are becoming more pervasive. This paper presents the growing field of applied computer science often referred to as *"sports informatics"*, and considers some of the advancements that have been made by its relatively small, but growing research community. The review includes: feedback systems, team play analysis, image/motion analysis, exertion interfaces and virtual reality, and data mining and artificial intelligence.

**Keywords:** Sports Informatics, Team Play Analysis, Feedback Systems, Image Processing, Data mining.

## 1 Introduction

A useful and thorough definition of sports informatics as a scientific discipline was proposed in [1]: *"Sport informatics is a set of multi- and interdisciplinary research programs including components of sport science and computer science. The material field is the application of tools, methods and paradigms from computer science on questions of sport science as well as the integration of sport scientific knowledge in computer science."* This definition highlights the two-way relationship between sport and computer science.

The broad scope of sports informatics research has applications in high performance sports, as well as the healthcare and entertainment sectors. Within elite sport the aim is typically to track, analyse and improve performance by providing feedback to athletes, coaches and/or sports scientists. With recent improvements in instrumentation (GPS, accelerometry, motion analysis) and data processing, contemporary work has focused on taking classical sports science research out of the laboratory and into more realistic field settings. For the wider population the primary aim of sports informatics, however, is to increase physical activity and exercise adherence in order to

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_26

maintain/improve physical fitness. This has helped foster physical activity and human movement related entertainment applications of sports informatics such as exergaming.

Applications of computers in sport have been reported as early as the mid-1960's [2] although these were somewhat rudimentary. According to [3], the term *sports informatics* was first coined at a congress in 1975 in Graz, Austria, organised by the International Organisation for Sports Information. Supplying Information and Communications Technology (ICT) in competitions is one of the most challenging practical examples of the applied use of ICT in sports. At the 2004 Olympics in Athens, more than 10,000 computers, 4,000 printers attached to 900 servers, 300 routers, and 2,000 switches were supported by 300-400 computer professionals [4]. [5] describes a virtual environment for enhancing athlete preparation ahead of the 2002 Salt Lake City Olympics, an example of sports informatics research aiming at improving athletic performance at major sports events. A major driver for sports informatics research was the dramatic improvement in computer hardware, which has become faster and smaller. The miniaturization of hardware has made it possible to attach a wide range of sensors onto athletes and equipment in order to acquire data without affecting the athletes during training. The following sections provide a review of the various relevant research streams identified during this study.

## 2 Research Streams

The research streams in sports informatics is categorized by [2] into: data acquisition, databases and expert systems; modelling and simulation; multimedia, presentation; virtual reality; information and communication technologies; biomechanics and sports technology. Due to advances in ICT and video processing algorithms, classical biomechanics data collection suites have become increasingly mobile and easily deployable in real world, field environments e.g. the entire pitch area within a soccer stadium.

Advanced modelling and analysis software have been applied to a wide variety of sports related problems. Many of these software are based on Artificial Intelligence (AI) techniques. Effective presentation and visualization of data can improve cognitive understanding of complex data outputs and be beneficial to coaches and athletes.

Virtual reality and immersive environment applications also have an important place in sports, given their demonstrated effectiveness in preparing athletes for competition and training.

Many applications in sports informatics are essentially a type of feedback system so the next section will review feedback systems specifically. Our study of literature identified the following research themes in sports informatics: team play analysis, image / motion analysis, exertion interfaces / virtual reality, and data mining / AI These themes are the focus of the rest of the paper.

## 3 Feedback Systems

Feedback enables athletes to modify their movements and produce optimum performance. [9] identifies and provides examples of eight types of feedback systems: 1-video information feedback, 2-three-dimensional virtual environments feedback, 3-intrinsic feedback under vibration conditions, 4-temporal feedback, 5-feedback about team performance, 6-feedback in aiming sports, 7-force platforms / transducers based feedback, 8-eye tracking technology based feedback. He also points out that combinations of these feedback types are quite common. [10] discussed the various performance lifecycle elements typical for instrumented feedback systems: 1-measurement and data collection, 2-data capture and storage, 3-data analysis and feedback in real time, 4-feedback post-event, 5-information and review, 6-data archiving. Achieving each within a complete system, represents a significant area of research, since the various individual tasks involved are non-trivial.

[11] presents an exemplary swimming performance monitoring and feedback system which caters for all elements of the performance lifecycle, capable of tracking and providing feedback on elements of interest, such as technique and physiological capability. A lot of other currently accessible systems track various performance parameters in isolation without proper integration to provide for the needs of the athlete, coach and all supporting personnel (biomechanist, nutritionist, strength-conditioning coach etc.). [12] provides good examples of three rapid feedback systems: rowing - pulling forces and reaction forces in the foot stretcher on a stroke-by-stroke basis; table tennis- monitor the impact positions of the ball on the table and the time interval between impacts of the ball during service; biathlon shooting – the orientation of the muzzle immediately before shooting is monitored. The latter system, for instance, uses a single video-camera positioned diagonally several meters in the direction of the shot. Image analysis, based on a simple colour histogram model, is used to pick up the muzzle position in relation to the shot-target. The classical use of laser reflective plates can be worn out and damaged quickly and are considerably more expensive and involved to set-up. As the shooter is aiming, a video is recorded and provided on a screen to the athlete for review.

[13,14] point out that a clear trend in sports informatics is the use of small ubiquitous sensors. The sensors used in sports can be categorised as those that are attached to the athlete's body (accelerometers, gyroscopes, GPS receivers, RFID tags, and a variety of physiological sensors), or mounted on parts of the sports equipment (i.e. equipment monitoring). They are used to measure physical quantities such as force, torque, pressure, acceleration / velocity / position (linear/angular) etc.

Feedback systems are not exclusive to high-performance sports. [15] presented a system called Mobile Motion Advisor (MMA), which is a mobile feedback system designed to support high school students by giving advice during physical exercises.

# 4 Team Play Analysis

The use of sports informatics relating to team sports is becoming popular in research [8, 16]. Previously, real-time player position and motion analysis in team sports was done using notation by hand or audio recording during a match, which was extremely time consuming particularly for the production of post-match analysis. [8] provides a review of observational analysis (for team sports in particular), and the technological advancements that have been progressively introduced into observational analysis. [17] presented one of the first computer-aided systems (goals/shots/specific plays were annotated with specific sections of video-action).

The most common contemporary approach in many sports, including soccer, rugby, ice-hockey or basketball, is to employ a system of multiple cameras (usually two to 12 strategically positioned cameras) around the playing field, and vision / image processing algorithms to track the players and the ball in a non-intrusive way [18]. For example, Prozone (http://www.prozonesports.com/) is a relatively well known commercial system. The problems of image processing associated with such systems include weather and lighting conditions and player occlusions. They are, therefore, typically semi-automated with a degree of manual verification [19, 20]. In football heuristics have been applied to automatically detect events such as simple ball-in / ball-out, ball contact, throw-ins, corner kicks, passes, possession, pass interceptions, shots on goal – derived solely on player and ball tracked data [21]. Using microwave-based player and ball tracking system can also provide coaches with player activity profiles, frequency distributions of tactical behaviour, and assess aspects of an opponent's tactical play [22]. This system employed AI techniques to classify the action into five layers: 1-position and motion; 2-action; 3-situation; 4-tactical and 5-tactical assessment – with each layer permitting an analysis at that level, and forming a basis for inferences at the next level.

Despite the drawbacks of visual camera based systems, advances in technology have enabled observational sports analysis to require less human intervention and to speed up the annotation process [8]. However, because of the need for manual verification with camera based systems a player tracking system based on sensors and microwaves tends to perform with far better accuracy [6]. Unfortunately, wearable sensors are not currently permitted during competition.

Analysis of team sports is also increasingly used to engage fans via television and web-based media. For example, the NHL'sonline Ice-Tracker system (http://www.nhl.com ) shows live matches on a graphic visualisation of the rink, with the capability to query the graphic for events of certain players, specific time-intervals, or types of events (body checks, penalty minutes, number of shots, types of shots, most frequent shooting position to goal angles, etc.). This arguably enriches the NHL website visitor's experience of a match.

In summary, comprehensive intelligent analysis has the capacity to enhance in-match and post-match coaching practice and decision making. However at present both the automation and sensitivity of event detection during competition (with visual camera systems), and the extraction of tactical patterns, require considerable research and development.

## 5 Image / Motion Analysis

One common methodology in biomechanics is the kinematic assessment of human movement, that is typically done with high-precision camera systems and custom hardware and software to track placed markers automatically. These systems tend to be expensive and mostly restrict data collection to indoor environments. Advanced image processing devices that do not require markers (e.g. Microsoft Kinect) use algorithms and infrared point clouds to gain depth perception and detect body segments have begun to appear. At present, however, the Kinect system underperforms considerably in some aspects of accuracy, compared to a conventional marker tracking system (VICON). On the other hand, the Kinect was surprisingly accurate in tracking the mass, size and inertia of body segments [23].

Gaining depth perception and subsequently performing pose estimation and accurate motion tracking, purely from video sequences (i.e. without the use of point clouds, as in the Microsoft Kinect) is an ongoing research challenge. The use of video footage to achieve the data collection is naturally a preferred choice due to wide availability of such data, both current and historical sporting events are available on high-quality video.

There is insufficient space for a complete review of research in video-based image processing for pose and motion estimation that is relevant to sports informatics. The special double issue edited by Sigal and Black [24] is recommended. The following issues and challenges are highlighted [24]:

- Tracking of complex activities from monocular imagery (or with fewer than four cameras);
- Dealing with general loose clothing;
- Tracking multiple interacting people;
- Automatically recovering from failure;
- Automatically adapting to different body shapes;
- Dealing with moving cameras / backgrounds and strong / variable lighting;
- Tracking through non-trivial environmental conditions.

In relation to the first item, with the increasing wide availability of high-quality monocular cameras within smart-phones, rather interesting solutions have been proposed. For instance, [25] proposed a marker-less body and motion tracking approach based on a dynamic colour model from a monocular (smart-phone recorded) video-sequence. The aim was to track motion for postoperative kinematic analysis, which would have applications in health and exercise. Acceptable segmentation and tracking results (related to correct form during execution of post-operative physical exercises) were obtained. Joint angles, however, were not very accurately tracked. The authors suggest augmenting the system with human-body models to improve joint angle accuracies. Including more (low cost) cameras tends to generally improve performance significantly. It is also expected that development of relevant algorithms will further improve performance relating to challenges two to seven above. In order to assist in the evaluation and validation of image processing algorithms for pose estimation and tracking of human motion, ground truth datasets are necessary. [25] suggest for instance the HumanEva (http://vision.cs.brown.edu/humaneva/publications.html) datasets to be used to benchmark future research in this area. The dataset contains calibrated videos in which subjects perform common actions such as gesturing, walking, jogging, etc., for which the corresponding ground-truth was obtained from an industry-standard motion capture system.

Although underwater and on water-based sports have not been discussed specifically, these sports present further challenges of their own in biomechanics data collection [26, 27, 28].

## 6 Exertion Inferfaces and Virtual Reality (VR)

The aim of interactive sports via web-based media is not to replace existing sports but to provide participants with a comparable activity when they are geographically separated. It enables participants to interact and compete in a manner comparable to co-located sports, and can provide a shared experience and social interaction. It can be difficult to find local fellow participants with similar physical capabilities in order to ensure a mutually enjoyable or productive experience. Web-based exercise and sports expands the range of potential exercise partners [29]. These applications have emerged in part from the fusion of sports and computer technology within computer games. However to date, not much work has been done on enhancing social sports experiences. Many computer games support network play but have been criticized for encouraging social isolation of the players. This began to change with the advent of exertion interfaces, such as the Nintendo Wii and Microsoft Kinect [30]. Exertional games, or exer-games, are computer games where players use grossmotor movements of the body to interact with the gameplay, rather than micro-manipulation such as using fingers to press buttons [30].

The development of exertion interfaces has stimulated further work and has been facilitated by a number of open source development libraries (e.g. Pointcloud library (http://www.pointclouds.org) and Open NI (http://openni.org/)), which open up the commercial technologies to researchers. [31] investigated the uses of Nintendo Wii, as exertion interfaces for the elderly, while Bekker and Eggen [32] looked at their potential benefits in use with children. It can be argued that interaction with fitness video games generally does not significantly improve a player's skill in real life. For example, playing bowling in Wii Sports does not improve bowling skills. Therefore some researchers, such as Campbell et al. [33], presented work where a daily fitness activity, for which there is a fitness application to track exercise performance, is augmented with an interactive and fun element, also often referred to as "gamification" of a real world activity [34]. It has been argued that "gamification" of real world sports potentially adds an extra element of social engagement to an activity, and can be useful in breaking-up sizeable exercising tasks into more achievable sub-goals. [33] present an example of Nike + iPod running experience augmented by a fantasy game that tries to improve on Nike + iPod. Nike + iPod system uses a small sensor placed in a runner's shoe to track runs in support of multiplayer challenges on the Nike+

website; however, the system is primarily focused on macro goals, and does not encourage frequent gratification associated with micro goals, but at the same time provides very limited social interaction within the online community. Kukini [33] is set in an imaginary world of elite class of swift runners of legend, known to undergo strenuous training as warriors, messengers, and foot racing athletes, with a significant social play element, integrated via communication layers throughout the game that support building relationships, forming teams, and interacting with other players both co-operatively and competitively. The game emphasizes micro goals by breaking up distances.

VR has been used effectively to help train astronauts, pilots, physicians, military personnel, and athletes [35]. In a 3D virtual environment, the coach may regulate important factors that influence perception, such as speed, orientation and directional changes, simply by operating a joystick or a keyboard. Training in this environment could enhance skill adaptation through progressive and repetitive practice in a systematic manner that is feasible in a highly controlled environment, but difficult to achieve in real environments. Skill may thereby result as a by-product of training in controlled simulated 3D environments. VR systems also have the potential to provide enhanced, immediate and direct feedback to participants [9]. Examples include a bobsleigh simulator for training the US bobsleigh team [36], a speedskating virtual environment for enhancing athlete visualisation and preparation for the 2002 Salt Lake City Olympics [5], and an archery system based on haptic feedback, motion tracking and custom graphical avatars [37]. There is some indication that various elements from VR and exertion interfaces are coming together in one solution [38]. The authors implement a martial arts game, based on an exertion interface, motivating training through a playful VR environment, and subsequent evaluation their system on 46 martial arts practitioners with encouraging results.

## 7 Artificial Intelligence (AI) and Data Mining

AI techniques are receiving attention in sports informatics literature. For example, sensor based feedback systems [39], team-play analysis systems [40], or many of the pose recognition applications [41] have benefited from AI classification algorithms. Baca et al. [6] point out that there is an increasing need for smart and intelligent systems and that research effort emphasis should shift from fundamental technologies to intelligent systems. The authors point out the need for applications of AI in coaching / training systems and especially the use of decision systems, which leverage machine learning and methods from AI. Such systems, it is argued, will help take the load off from coaching staff by semi-automatically discovering valuable insights and patterns in the datasets. [7] provides an introduction into AI / soft-computing modeling within sports informatics. A lot of classical sports science experiments and analysis of data could benefit from these techniques and the field of data mining [13]. Much historical performance data from different sports is readily available; however, the data mining and pattern discovery within these has been limited. For example an application of genetic algorithms was used to optimize alpine ski paths, given some constraints and fitness level parameters [7]. Erdogan et al. [42] provide a useful review of previous uses of Artificial Neural Networks in sports science related problems.

The increase in ubiquitous sensors has also facilitated the use of embedded intelligence in sports-related feedback systems. For example Neural Networks and Hidden Markov models have been used to recognise different movement patterns, based on force data [43]. There are numerous other potential applications in sports, healthcare, entertainment and leisure.

## 8 Conclusions

Feedback systems are not necessarily only applied in high-performance sports, although this has been often a common application area. Given current nationwide issues and health care costs associated with obesity, applications of such systems present a research area for future work. Non-intrusive or sensor based player tracking systems are highly sought after and receiving extensive research and development attention. The use of this information is also encouraging the development of intelligent decision support systems that may ultimately provide real-time guidance to coaches during match play. Image recognition is traditionally a significant research area in computer science. The application of image / video processing algorithms to sports, health and exercise applications, presents an interesting set of problems in constrained environment. Exertion interfaces, virtual reality, and AI / machine learning are all highly relevant to sports informatics and hold much promise. The integration of different technologies to cover the whole sports requirements also need to be considered. However, one must consider the criticism that technology and informatics use in professional sports can sometime be seen as unethical and inappropriate by some [44].

## References

1. Link, D., Lames, M.: Sport Informatics - Historical Roots, Interdisciplinarity and Future Developments. International Journal of Computer Science in Sport 8, 68–87 (2009)
2. Dabnichki, P., Baca, A. (eds.): Computers in Sports. WIT Press, UK (2008)
3. Recla, J., Timmer, R. (eds.): Kreative Sportinformatik (Creative Sport Informatics). Der Internationale Jubilaums-Kongres 1975 in Graz. Verlag Karl Hofmann, Austria (1976)
4. Sakac, B.: Information technology at the Olympic Games. In: Dabnichki, P. and Baca, A. (eds.) Computers in Sport. WIT Press, UK (2008)
5. Sorrentino, R.M., Levy, R., Katz, L., Peng, X.: Virtual Visualization: Preparation for the Olympic Games Long-Track Speed Skating. International Journal of Computer Science in Sport 4, 40 (2005)
6. Baca, A., Dabnichki, P., Heller, M., Kornfeind, P.: Ubiquitous computing in sports: A review and analysis. Journal of Sports Sciences 27, 1335–1346 (2009)
7. Perl, J.: Modeling. In: Dabnichki, P., Baca, A. (eds.) Computers in Sport. WIT Press, UK (2008)
8. Lames, M.: Coaching and computer science. In: Dabnichki, P., Baca, A. (eds.) Computers in Sport. WIT Press, UK (2008)

9. Liebermann, D.G., Katz, L., Hughes, M.D., Bartlett, R.M., McClements, J., Franks, I.M.: Advances in the application of information technology to sport performance. Journal of Sports Sciences 20, 755–769 (2002)
10. Atkins, G.: Challenges in the Analytics of High Performance Sport, Sports Informatics Symposium, Loughborough, UK (2012), `http://sportsinformatics.lboro.ac.uk/events/symposium/symposium-talks`
11. Mullane, S.L., Chakravorti, N., Conway, P.P., West, A.A.: Design and implementation of a user-centric swimming performance monitoring tool. Journal of Sports Engineering and Technology 225, 213–229 (2012)
12. Baca, A., Kornfeind, P.: Rapid feedback systems for elite sports training. IEEE Pervasive Computing 5, 70–76 (2006)
13. McCullagh, J., Beattie, M., Nugent, C.D.: Pervasive Technology to Facilitate Wellness. In: Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments, Samos, Greece (2010)
14. Chi, H.E.: Sensors and Ubiquitous Computing Technologies in Sports. In: Dabnichki, P., Baca, A. (eds.) Computers in Sports. WIT Press, UK (2008)
15. Preuschl, E., Baca, A., Novatchkov, H., Kornfeind, P., Bichler, S., Boecskoer, M.: Mobile Motion Advisor - a feedback system for physical exercise in schools. In: Proceedings of the 8th International Conference on The Engineering of Sport, Vienna, Austria (2010)
16. Peters, D.M., O'Donoghue, P. (eds.): Performance Analysis of Sport IX, Routledge, USA (2013)
17. Franks, I.M.: The structure of sport and the collection of relevant data. In: Baca A. (ed.) Computer Science in Sport. Obv and Hpt Publishers, Austria (2000)
18. Wilhelm, P., Thomas, P., Monier, E., Timmermann, R., Dellnitz, M., Werner, F., Ruckert, U.: An Integrated Monitoring and Analysis System for Performance Data of Indoor Sport Activities. In: Proceedings of the Tenth Australasian Conference on Mathematics and Computers in Sport, Darwin, Australia (2010)
19. Hedley, M., Mackintosh, C., Shuttleworth, R., Humphrey, D., Sathyan, T., Ho, P.: Wireless Tracking System for Sports Training Indoors and Outdoors. In: Litzenberger, S., Kafka, P., Sabo, C. (eds.) Proceedings of the 8th International Conference on the Engineering of Sport, Austria, Vienna (2010)
20. Frencken, W.G.P., Lemmink, K.A.P.M., Dellemann, N.J.: Soccer specific accuracy and validity of the local position measurement (LPM) system. Journal of Science and Medicine in Sport 13, 641–645 (2010)
21. Gudmundsson, J., Wolle, T.: Towards Automated Football Analysis: Algorithms and Data Structures. In: Proceedings of the Tenth Australasian Conference on Mathematics and Computers in Sport, Darwin, Canada (2010)
22. Beetz, M., Kirchlechner, B., Lames, M.: Computerized real-time analysis of football games. IEEE Pervasive Computing 4, 33–39 (2005)
23. Chopin, S., Wheat, J., Heller, B.: KINECT BIOMECHANICS: PART 1 and PART 2 (2011) (accessed), `http://engineeringsport.co.uk/2011/05/09/kinect-biomechanics-part-1/`, `http://engineeringsport.co.uk/2011/07/04/kinect-biomechanics-part-2/` (last accessed on May 1, 2012)
24. Sigal, L., Black, M.J.: Guest Editorial: State of the Art in Image-and Video-Based Human Pose and Motion Estimation. International Journal of Computer Vision 87(1), 1–3 (2010)

25. Lee, J., Karasev, P., Zhu, L., Tannenbaum, A.: Human Body Tracking and Joint Angle Estimation from Mobile-phone Video for Clinical Analysis. In: Proceedings of MVA 2011 IAPR Conference on Machine Vision Applications, Nara, Japan (2011)
26. Lauder, M.A.: Motion analysis in water sports. In: Dabnichki, P., Baca, A. (eds.) Computers in Sports. WIT Press, UK (2008)
27. Slawson, S.E., Conway, P.P., Justham, L.M., West, A.A.: The development of an inexpensive passive marker system for the analysis of starts and turns in swimming. In: Litzenberger, S., Kafka, P., Sabo, C. (eds.) Proceedings of the 8th International Conference on the Engineering of Sport, Austria, Vienna (2010)
28. Runciman, R.J.: Water-skiing biomechanics: a study of intermediate skiers. Journal of Sports Engineering and Technology 225(1), 231–239 (2011)
29. Mueller, F.F.: Long-distance sports. In: Dabnichki, P., Baca, A. (eds.) Computers in Sports. WIT Press, UK (2008)
30. Mueller, F., Gibbs, M.R., Vetere, F.: Towards understanding how to design for social play in exertion games. Personal and Ubiquitous Computing Journal 14(1), 417–424 (2010)
31. Gerling, K.M., Schild, J., Masuch, M.: Exergame design for elderly users: the case study of SilverBalance. In: Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology, Taipei, Taiwan (2010)
32. Bekker, T.M., Eggen, B.H.: Designing for children's physical play. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy (2008)
33. Campbell, T., Ngo, B., Fogarty, J.: Game design principles in everyday fitness applications. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW), San Diego, USA (2008)
34. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: Defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference, Tampere, Finland (2011)
35. Katz, L., Parker, J., Tyreman, H., Levy, R.: Virtual reality. In: Dabnichki, P., Baca, A. (eds.) Computers in Sports. WIT Press, UK (2008)
36. Kelly, A., Hubbard, M.: Design and construction of a bobsled driver training simulator. Sports Engineering 3(1), 13–24 (2000)
37. Göbel, S., Geiger, C., Heinze, C., Marinos, D.: Creating a virtual archery experience. In: Proceedings of the International Conference on Advanced Visual Interfaces, Rome, Italy (2010)
38. Haemaelaeinen, P., Ilmonen, T., Hoeysniemi, J., Lindholm, M.: Martial arts in artificial reality. In: Proceedings of the Conference on Human Factors in Computing Systems, Portland, USA (2005)
39. Holleczek, T., Ruegg, A., Harms, H., Troster, G.: Textile Pressure Sensors for Sports Applications. In: IEEE Conference on Sensors, Waikoloa, Hawaii (2010)
40. Beetz, M., Kirchlechner, B., Lames, M.: Computerized real-time analysis of football games. IEEE Pervasive Computing 4, 33–39 (2005)
41. Efros, A.A., Berg, A.C., Mori, G., Malik, J.: Recognizing action at a distance. In: Proceedings of IEEE International Conference on Computer Vision, Nice, France (2003)
42. Erdogan, A., Cetin, C., Goksu, H., Guner, R., Baydar, M.L.: Non-invasive detection ofthe anaerobic threshold by a neural network model of the heart rate–work rate relationship. Journal of Sports Engineering and Technology 223, 109–115 (2009)
43. Griesinger, F.; Kloos, U., Martinez, N., Tullius, G. (eds.): Informatics Inside: Pattern Recognition in Gait Analysis with the Ground Reaction Force. Hochschule Reutlingen, Germany (2012)
44. James, D., Sabo, A.: The Ethics of Using Engineering to Enhance Athletic Performance. In: Litzenberger, S., Kafka, P., Sabo, C. (eds.) Proceedings of the 8th International Conference on the Engineering of Sport, Austria, Vienna (2010)

# Experiments in the Application of Computer Vision for Ball and Event Identification in Indoor Sports

Kevin M. Oldham[1], Paul Wai Hing Chung[1],
Eran A. Edirisinghe[1], and Benjamin J. Halkon[2]

[1] Department of Computer Science,
Loughborough University, Loughborough, LE11 3TU, United Kingdom
[2] Wolfson School of Mechanical and Manufacturing
Engineering, Loughborough University, Loughborough, LE11 3TU, United Kingdom

**Abstract.** There is little evidence of thorough experimental research having been performed for the investigation of optimum computer vision algorithms suitable for table tennis umpire support systems and player performance analysis. The only practical way of evaluating and comparing such algorithms is to gather source video data from repeatable experiments in controlled conditions and pass multiple sequences through the same validated algorithms, techniques and processes. This paper presents a digest of software engineering issues, considerations and challenges observed as a direct result of performing over 250 experiments with computer vision and table tennis. Results demonstrate that variations in camera configuration have a direct impact on the success of a selected algorithm and that a comparison of algorithms against the source video data must be made for confirmation. Finally, determination of the location of the ball itself varies between algorithms and a parallel, non-obtrusive technology is required to provide baseline data.

**Keywords:** object segmentation, detection, ball, sports, computer vision, table tennis, algorithm, comparison.

## 1 Introduction

Identifying an optimum computer vision (CV) algorithm using monocular recording hardware for table tennis requires a root analysis. A small, yet fast ball with spin induced curved trajectories, high frequency collisions and occlusions, played across short distances and at a raised plane makes this a challenging CV prospect. Additionally, when restricted to a monocular system, consideration for scale of reference and ball location must be incorporated. Finally, current CV research in table tennis is limited when compared to other popular sports, such as football or baseball.

A review of the available CV based ball detection processes demonstrates an apparent success for several competing CV algorithms in identifying a table tennis ball from an image. Wong et al [1] recommends a binary threshold, followed by feature detection and classification using an artificial neural network. Zhang et al [2], alternatively, discusses a combination of binary thresholding, adjacent frame differencing,

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_27

Run Length Coding (RLD) and feature extraction whilst Teachabarikiti et al. [3] recommend a mean shift algorithm. The ability to compare these different algorithms and to reason which is the most successful (against a standard set of criteria) is regularly overlooked. However, implementation details are missing. A retrospective analysis of the available data is considered to be of little scientific value given the lack of camera hardware specification, lighting conditions, camera positioning, video format and player ability. The environments within which this sport is recorded are not rigorously defined; sports halls, open venues and arenas of varying characteristics are all used. Also there is little reason for a CV sports application to simply detect a ball; it should also generate motion and location data, for performance analysis and umpire decision support.

In deciding upon a method, the software engineer is presented with an extensive selection of algorithms. This choice is further compounded when existing research uses unspecified video recording hardware and the environment within which the game is being played is not controlled or consistent. This paper aims to assimilate key considerations for ball identification in table tennis (and similar sports) enabling future comparisons of algorithms to be made.

## 2 Apparatus and Experimentation

The ability to have direct comparisons between CV processing solutions requires a robust and systematic approach to data collection. Environmental factors, such as lighting, background and the ball surface itself all have an impact on the source data and need to be standardized. The recording device, the camera type, lens, exposure settings, location relative to the table and recording format all need to be documented for repeatability. The sports equipment must be of regulatory standard. For table tennis all the balls must be of 40mm diameter, with rackets and tables meeting the ITTF guidelines [4].

### 2.1 Positioning

Careful consideration must be given to the location of the camera, which must be fixed, on a rigid support and be oriented to the table such that it captures the ball in the maximum number of frames. This must be consistent for all video sequences. For many sports this location coincides with that of the umpire, where static. With table tennis this location is in-line with the net, at an incline of 10- $20^{o}$ to the plane of the table and 3.5m from its edge. Exact location must be detailed with the video.

### 2.2 Lighting

Fluorescent lamps should be avoided due to flicker apparent on the recording. This in turn can cause ball detection algorithm to detect multiple false positives within frames. Consistent levels of brightness are critical to the successful detection and for direct comparisons of the CV algorithms. Any additional lighting must be identified.

### 2.3 Camera Configuration

As the primary source of data, the selection of suitable video hardware, capable of capturing the scene in sufficient detail, is a priority. Previous research has used video sequences obtained from the internet, using YouTube or online table tennis training sites with no details of camera configuration used. Although sufficient to establish a principle of ball detection, it is not rigorous from a computer science perspective. Instead video source data with the following parameters must be carefully considered and documented in detail.

**Shutter Speed.** The speed at which the image is captured primarily defines the shape of the imaged ball. It may not be realistic (especially at the high speeds of professional players) to capture the ball as a circle, but a camera with the fastest shutter speed possible is optimal. Documented speeds of a table tennis ball vary, with a world record estimated to be 29.4m/s [5]. At this world record speed, a ball would have travelled 150mm during a shutter speed of 1/200th. Work during these experiments demonstrates the average player is much slower than this, but consideration must be given to shutter speeds of 1/100th second to reduce substantial blurring of the ball image. If a blurred ball region is the only option, then further processing based on the work by Caglioti et al [6] is necessary.

**Frame Rate.** The frame rate of the camera should produce data points capturing the ball during key events. However, with extremely fast cameras, the amount of data produced may introduce unwanted redundancy and overheads into the processing. A recommended approach is to ensure the ball does not move more than three times its own diameter between frames at the fastest match play.

**Lens.** The lens (combined with sensor) requires a field of view wide enough to capture the entire table and its immediate vicinity. A lens with a focal length of the 24mm will be sufficient on a full frame sensor to give the required angle of view of 90° at a distance of 3.5m (measured in line with the net). The greatest possible focal length is required to minimize spherical distortion [7]. Below 20mm and the image will become significantly distorted. Moving the camera away from the table decreases the ball resolution.

**Sensor.** Practical experiments in this research have found the resolution range of 2-8MP to be suitable. Sensor size when combined with the number of pixels, is more critical than number of pixels alone. The optimum is to implement the largest sensor possible. The current ubiquitous APS-C (22.3 x 14.9 mm) is suitable; successful results in this research have also been achieved with 17.3 x 13 mm and 12.8 x 10.2mm sensors. The size of sensors in most smartphones (4.54 x 3.42 mm) currently precludes their use for good source data. Standard webcam specifications are sensors of 4.60 x 3.97mm running at no more than 30fps. Cameras with small sensors have a narrower field of view and so need to be moved further away, reducing the pixel size of the ball. The output from the low resolution and large distances travelled between frames is of poor quality for ball identification and accurate event analysis.

**Focus.** The table dimensions are 1.525m wide and 2.74m in length. Focus variations across this field of view affects the diameter of the detected ball, decreases color saturation and reduces the contrast differential. It is has been found that manual focus should be used to ensure the camera does not attempt to re-focus automatically during play due to the wide depth of field. A Siemens star focus chart should be used to set focus at the center of the table. With the 24mm lens positioned as described, a focal ratio of f/1.4 is adequate to maintain the sufficient depth of focus of 1.6m [8] throughout the table area.

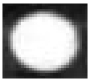
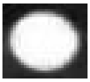
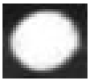
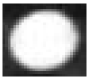
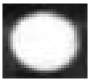
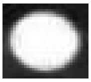
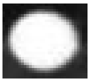
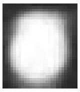
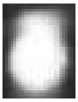
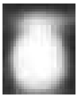
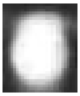
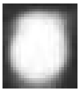
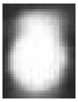
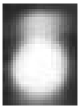
## 3 Candidate Filtering

Once a monocular video sequence is captured, identifying the ball from a set of candidates in a single frame is non-trivial. There are usually reflections, shadows, logos and other ball shaped regions confusing the scene. Results indicate that there is an optimal combination and prioritization of feature filtering to remove false positives. When comparing algorithms for ball detection and event identification, the candidate results must pass through this same filter to maximize the detection data set.

### 3.1 Reflections and Shadows

During controlled experiments designed to detect a table tennis ball using CV, observations have shown that, with particular table surfaces, as the ball approaches the proximity of the table (usually within 15cm) a reflection of the ball appears with similar feature characteristics to the ball. Output from an example experiment is presented in Table 1. The top row represents the ball; the bottom row is the reflected region in the same frame.

**Table 1.** The ball and its reflection

| | *a* | *b* | *c* | *d* | *e* | *f* | *g* |
|---|---|---|---|---|---|---|---|
| **Ball** |  |  |  |  |  |  |  |
| **Reflection** |  |  |  |  |  |  |  |

The shape, size and relative motion of the reflected image confuse a simple shape filter, finding more than one ball. As the ball moves towards a collision with the table (d-e), both the image of the ball and its reflection become similarly oval in shape. This problem is compounded when allowance is made for the shape of the ball becoming elongated due to exposure speed.

Evaluation of a tracking algorithm, the Kalman Filter [9], was made to overcome the loss of detection. However, this process becomes ineffective due to the abrupt change in ball direction at the point of collision; the Kalman Filter 'loses' the ball at the critical stage of impact, expecting the ball to continue through the table. By the time the Kalman Filter has ‘found’ the ball, the collision event has been missed. Detecting the precise point of contact with the table is essential for a successful CV application and must be resolved. The solution is to implement a rigorous feature filter, based on a prioritized list of *derived* shape and intensity features.

### 3.2 Shape

Given the finite exposure time of each frame, the circumference of the image of the ball during match play is rarely an ideal circle. Estimating the region *roundness* is therefore a more accurate measure. Building on the work of Wong et al [1], the experimentally proven list of shape characteristics is given in Table 2 (this feature set can apply to any sport with a ball as the primary object of interest).

**Table 2.** Ball feature properties

| **Feature** | **Definition** |
|---|---|
| Diameter (D) | Calculated using the bounding box method (average of height and width of the smallest box around the region), or the centroid calculation (the center of mass); |
| Intensity (I) | The color intensity of the region. The intensity may be used to filter reflections and shadows (see Section 3.1) |
| Perimeter (P) | The total distance around the boundary of the region. |
| Eccentricity (E) | The ratio of the distance between the foci of the ellipse and its major axis length. A value between 0 and 1. |
| Area (A) | The number of pixels in the segmented region. |
| Solidity (S) | The proportion of pixels in $A_c$ which are also in the region |
| Equivalent diameter ($D_e$) | Given the area (A) of the region, this is the diameter of a circle with the same area. |
| Convex Area ($A_c$) | The number of pixels bounded by the convex hull. |
| Perimeter to diameter (PD) | The ratio of P to D; should approach π. |
| Rounded top | $r^2 = (x-x_c)^2 + (y-y_c)^2$, see Wong et al [1] |

Additionally a number of derived parameters are used to provide further confidence of the region being the shape of the ball. An estimated ball diameter, ball area and an estimation of Pi are calculated using the following equations:

$$D_{est} = \frac{P}{\pi} \quad (1)$$

$$A_{est} = \pi \left(\frac{D}{2}\right)^2 \quad (2)$$

$$Pi_{est} = \frac{P}{A} \quad (3)$$

In particular (3) gives a strong indication of its circularity, reducing false positives from non-curved objects, such as a star or indeed the ball’s own reflection.

### 3.3 Color and Intensity

Table tennis balls are generally manufactured in either white or orange [10] and in matte only. However the precise color values recorded by the video hardware can vary considerably, depending on lighting, camera sensor, and internal color processing techniques and thus cannot be matched directly. During experiments with different ball variations, the color of the ball as a feature characteristic has proved surprisingly unreliable. Instead, the high intensity value in a greyscale representation is an improvement over the color.

### 3.4 Feature Priority

During complex filtering algorithms, processing resources are easily exhausted. However by isolating the experimental outputs, the feature filter characteristics have been shown to not deliver equal identification efficiency. A comparison of the detection improvement for each feature is presented in Fig. 1.

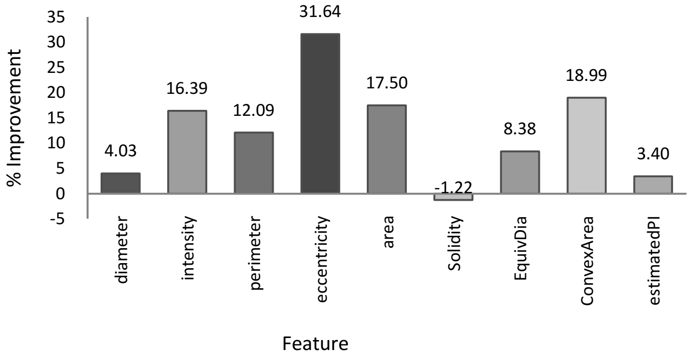


**Fig. 1.** A comparison of detection improvements of feature filters

It is apparent that not all features have an equal weighting in detecting the ball; an order of priority for having the greatest effect can be derived. By carefully eliminating the less effective feature filters, an optimized detector is established.

### 3.5 Feature Filtering Example

Applying an optimized feature filter can dramatically improve the quality of the data. An example comparison of ball detection with and without the optimized feature filter can be seen in Fig. 2. The chart clearly shows reflections without the use of the filter underneath the path of the ball; the point of contact is also missing. With the filter, the reflection is mitigated and the point of impact is recorded. The exclusion of false positives caused by many other sources, such as logos, lights or rackets has also been achieved by using this feature detection filter.

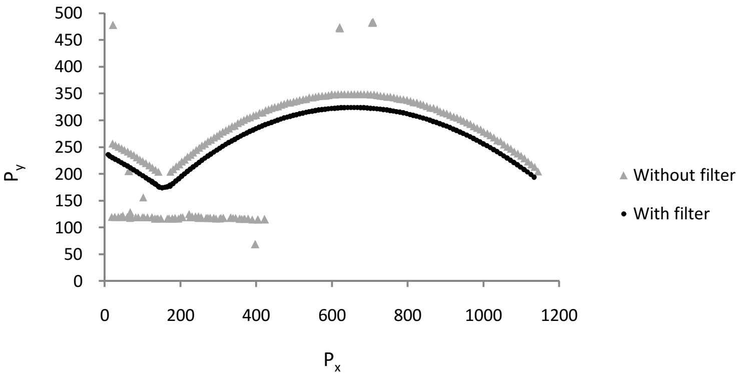


**Fig. 2.** A comparison of filtering and non-filtering on a table tennis bounce. Without filter ($P_y$) data has been shifted vertically for clarity.

# 4 Ball Location

When investigating different algorithms for performance analysis and umpire decision support, it is not only important to detect the ball, but also to determine its location. A finding of this research is that not all CV algorithms are able to equally determine the location of the center of the ball $P(x, y)$ within a given frame. As an example see **Fig. 3,** presenting a comparison of the Gaussian Mixture Model and the Optical Flow (Lucas Kanade) algorithms are made (the frame count has been normalized across the range for all sequences used). This is often due to the capability of an algorithm to find the boundary of an object when it is moving, when slightly out of focus, or when approaching another object such as the net or its own reflection.

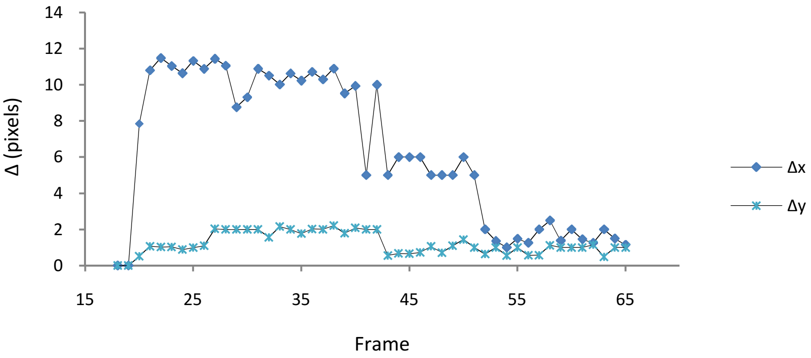


**Fig. 3.** Mean location Δ between Gaussian Mixture Model and Optical Flow (Lucas Kanade) for $P_x$ and $P_y$. Ball diameter = 18px

With each algorithm producing different location data, deciding which to use becomes an investigation into algorithm application and quality. In determining accuracy, additional consideration must be given to establishing the location of the ball in parallel without using CV; a measuring procedure must be developed to produce impartial, externally generated location data with which to validate the algorithm location output.

## 5 Critical Success Metrics

In measuring and comparing algorithms to select an optimum approach for ball detection in table tennis, a number of critical success metrics have been defined for algorithm comparison, as presented in Table 3.

**Table 3.** Critical success factors

| Metric | Group | Description |
|---|---|---|
| Initial identification | Technique Detection | First frame in which the ball is detected by the algorithm |
| Final identification | Technique Detection | Last frame in which the ball is detected by the algorithm |
| Identification count | Technique Detection | Number of frames with the ball positively detected |
| False identification count | Technique Detection | Frames showing the ball as visible, when it is not in the frame or is fully occluded |
| False identification(s) in frame | Technique Detection | Number of frames where one or more object is detected but none are the ball |
| Non-identification | Technique Detection | Number of frames where the algorithm is not identifying the ball and it is in the frame |
| Unidentified let frames | Match event | Number of frames at a let where the ball is not detected |
| Unidentified net frames | Match event | Number of frames at a net collision where the ball is not detected |
| Unidentified frames during bounce | Match event | Number of frames at a table bounce where the ball is not detected |
| Mean ball eccentricity | Ball metric | Average eccentricity of the ball for all detected frames |
| Mean ball diameter | Ball metric | Average diameter of the ball for all detected frames |
| Ball location | Ball metric | Cartesian co-ordinate accuracy |
| Processing speed | Processing | Average frame processing speed using the selected technique |

These metrics enable a direct comparison of identified algorithms, to determine which is the most accurate for the application. A given video, or selection of videos with a variety of scene complexities, must be processed and to generate the success metrics output. Using these metrics, the detailed performance of an algorithm, its strengths and weaknesses, can be defined, described and compared.

## 6 Conclusion

Using videos of unknown origin to develop ball extraction algorithms, whether for table tennis or any other sport, does not provide rigorous scientific data as a basis of CV algorithm selection and proof. There are many possible algorithms which could be used, and they themselves can be placed in many combinations with each other. A method of comparing the algorithms must to be used and evaluated against a predefined set of critical success factors.

The hardware for recording sequences must be carefully selected. Knowing the frame rate, for example, can give an indication of ball speed. It can also be used to interpolate missing data. The shape of the ball can vary due to camera settings, lenses and lighting. All of these must be optimized and standardized before any investigation into the most suitable CV algorithm can begin.

The algorithm must include a specific, consistent feature filter, which can allow for elongated ball representations and discard complications of reflections and shadows at key match play events, such as during a table bounce. Not all features, however, provide equal benefit to identifying the correct object of interest. Therefore, when hardware resources are restricted, features may be prioritized or omitted.

Tracking algorithms, such as The Kalman Filter, add only a marginal benefit to sports such as table tennis, where the ball is travelling fast and deviates due to collision many times per second. Using a local area of interest algorithm is a more successful solution in determining successful positive detections.

Finally it cannot be assumed that all CV algorithms detect the ball in the same location within an image. The derived location of the ball varies depending on the CV algorithm being used. An external, non-CV, measurement of the path of the ball must be created for each video sequence, so that a ground truth exists and the differences in location outputted by the investigated algorithm have a baseline from which to be assessed.

## References

1. Wong, P.: Identifying table tennis balls from real match scenes using image processing and artificial intelligence techniques. J. Simul. Syst. Sci. Technol. 7 (2009)
2. Zhang, Z.Z.Z., De Xu, D.X., Tan, M.T.M.: Visual Measurement and Prediction of Ball Trajectory for Table Tennis Robot. IEEE Trans. Instrum. Meas. 59(12), 3195–3205 (2010)
3. Teachabarikiti, K., Chalidabhongse, T.H., Thammano, A.: Players tracking and ball detection for an automatic tennis video annotation. In: 11th International Conference on Control, Automation, Robotics and Vision, ICARCV 2010, pp. 2491–2494 (2010)

4. ITTF, International Table Tennis Federation Handbook 2013-2014, 2013th–2014th edn. (2013)
5. Turberville, J.: How fast does a Table Tennis ball really go? (2004), `http://www.jayandwanda.com/tt/speed.html` (accessed: September 13, 2014)
6. Caglioti, V., Giusti, A.: Recovering ball motion from a single motion-blurred image. Comput. Vis. Image Underst. 113(5), 590–597 (2009)
7. Weng, J., Cohen, P., Herniou, M.: Camera calibration with distortion models and accuracy evaluation. IEEE Trans. Pattern Anal. Mach. Intell. 14, 965–980 (1992)
8. Mchugh, S.: Depth of Field Calculator. Cambridge in Colour (2014), `http://www.cambridgeincolour.com/tutorials/dof-calculator.htm` (accessed: July 14, 2014)
9. Kalman, R.E.: A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering 82, 35 (1960)
10. ITTF, The Laws of Table Tennis, in ITTF Handbook 2013 - 2014, ITTF, pp. 24–31 (2013)

# Semi-automatic, Landmark-Based Feedback Generation for Stand-Up Exercises

Pablo Fernández de Dios, Paul Wai Hing Chung, and Qinggang Meng

Department of Computer Science, Loughborough University, Loughborough, UK
{P.Fernandez-de-Dios,P.W.H.Chung,Q.Meng}@lboro.ac.uk

**Abstract.** This paper presents an approach to automatic human motion feedback generation for basic stand-up exercises. A semi-supervised classifier, based on the C4.5 decision tree algorithm and AdaBoost, is trained with ground-truth, manually labelled sequences of key body poses. Before calculating feedback features to be learnt, synchronisation of training and testing sequences –often incomplete and/or asymmetric– with a reference (the *tutor* or exemplary performance) is done. Finally, an algorithm for adaptation of a performance to that of the *tutor* is proposed, in order to further enrich the feedback delivered to the user. The proposed framework generates numerical feedback to the user and the main limitations of the proposed approach are discussed.

**Keywords:** exercises, assessment, supervised machine learning, feedback.

## 1 Introduction

In the field of gaming, technology has made big steps, to the point that human motion recognition has become almost trivial. Tracking the 3D coordinates of 23 joints of the user in real time, at an affordable cost, opened the door to a whole new world for the gaming industry. Moreover, smart mobile applications leverage with technologically advanced embedded sensors, can provide useful feedback regarding, among other things, fitness performance, timing and geospatial tracking. However, there remains the challenge of autonomously interpreting all this information and deliver human-understandable feedback.

This paper presents the results of an attempt to produce reliable feedback to the user, in the context of stand-up exercises for elderly people and athlete's rehabilitation. An approach for generic *key body poses matching* and automatic *feedback model learning and feedback generation* (based on numerical attributes) is presented. Also, a method for *adapting a motion data sequence to another*, in order to compliment visual feedback, is detailed.

The remaining of this paper is structured as follows. First, a summary of research works related to automatic and semi-automatic assessment of human exercise is presented. Then, the proposed approach for automatic feedback generation is explained, followed by a discussion of the experimental results using real human motion data captures. Finally, future work is discussed.

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems*,
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_28

## 2 Related Work

Research in automatic assessment of physical performances currently does not address the provision of specific feedback to the user, in the context of basic fitness primitives. Ontologies are used to represent knowledge of both context and expert evaluation information [2, 14, 19]. The feedback often consist of an encouraging comment based on information like bio-feedback, user's personal data or user's level of activity.

In the field of tele-medicine, remotely retrieved sensor data at the user's end are sent to the expert physician or therapist. [15] is an integrated system that incorporates a network architecture to facilitate the remote observation of patients performing functional assessment exercises in their homes. The data shown to the specialist consist of a stream of high level body posture and motion features like joint angles and joint range of motion (ROM). However, these approaches do not provide real-time feedback to the user, if any at all.

Motivation and level of engagement in exercise is considered in different fashions. Adaptation to the context is used to encourage the user and transmit a higher level of naturalness. For example, the level of difficulty in *exergames* can be adapted to the level of performance of the user [6, 18] or the avatar (representing one-self) can adapt its appearance [11, 13, 16]. Also, the nature of the exercises can change to adapt to the physical condition of the user [5].

Modification (or adaptation) of natural human motion is performed in order to achieve various aims such as manipulating robots [7, 8] or re-targeting [12] (i.e. re-scale a skeleton subject to certain constraints). A human-like, synthetic avatar is often used as a guide for performance [3, 11, 13].

## 3 A Method for Semi-automatic, Landmark-Based Assessment Model Learning and Feedback Production

The overall system has been conceived as a cyclic repetition of performances by the user, as depicted in Figure 1.

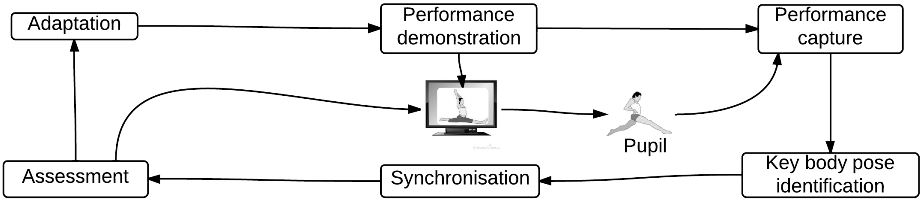


**Fig. 1.** Graphical representation of the life cycle of the assessment system

Ideally, the *pupil* will engage in physical activities on a daily or weekly basis (depending on their condition) for a certain period of time. Each performance will encompass the batch execution of several stages of analysis.

The process flow begins with the *pupil*, who is introduced with an animation demonstration of the exercise (*Performance demonstration*). Then, their performance is captured, recorded and pre-processed (*Performance capture*). The first motion data analysis sub-task consists of detecting the *un-assessed landmarks* (*Key body pose identification*). From this point onwards, the subsystems of the proposed approach are highlighted in Figure 2, performing *Synchronisation*, *Assessment* and *Adaptation* between *tutor*'s and *pupil*'s performances. These subsystems are described in details further on this section.

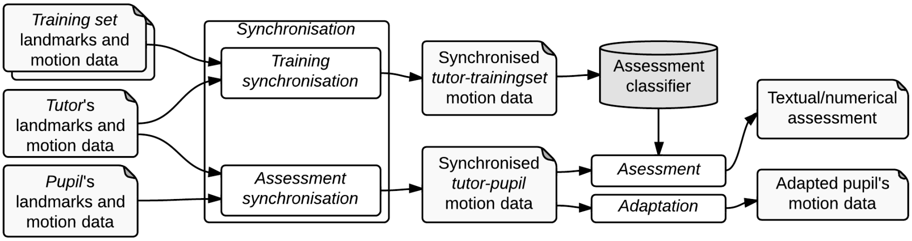


**Fig. 2.** System diagram depicting the synchronisation, assessment and adaptation subsystems

Before the assessment can be done, the *Assessment classifier* needs to be trained with matches between training samples ground-truth landmarks and those of the *tutor*. These are extracted from synchronised pairs of training sample and *tutor* (*Training synchronisation*)

Given the performance of both the *tutor* and the *pupil*, as well as the predicted landmarks, synchronisation between the formers is done first (*Assessment synchronisation*). This gives a matching between two sequences of body poses in order to be compared (*Synchronised tutor-pupil motion data*). Then, assessment of this matches is performed by the *assessment classifier*.

Finally, an adaptation of the *pupil*'s performance to that of the *tutor* is created and shown, along with the performance itself, in order to ease the overall understanding of what needs to be improved.

### 3.1 Landmark Sequences Synchronisation

The foundations for the current approach have been detailed in [9, 10], where a method for key body poses (*landmarks*) discovery from simple stand-up stretch exercises MoCap (*Motion Capture*) data from a Kinect camera has been materialised. A predicted *landmark sequence* $\overline{\chi}$ is produced by a semi-supervised (i.e. requiring prior training) classifier based on the C4.5 decision trees algorithm and the AdaBoost meta-classifier, being $\overline{\chi} = \{l_1 \, ... \, l_N\}$, with $l_i = \{\textsc{Frame}(i), \textsc{Type}(i)\}$ $\forall \; l_i \in \overline{\chi}$ and $\textsc{Frame}(l_i) > \textsc{Frame}(l_{i+1}) \; \forall \; l_i \in \overline{\chi}$. $\textsc{Frame}(i)$ is the time frame where the $i$-th landmark has been identified, whereas $\textsc{Type}(i)$ is the landmark type, i.e. range-of-motion limit within the exercise. This result simplifies enormously the task of retrieving specific key body poses on the input *pupil* performance to be matched against those of the reference motion (of the *tutor*).

Given two landmark sequences $\overline{\chi_1}$ and $\overline{\chi_2}$, a *synchronisation* between the two is defined as $S = \{s_1 \,...\, s_M\}$, with $s_i = \{t_i, p_j\} \forall s_i \in S$, where $t_i \in \overline{\chi_1}, p_j \in \overline{\chi_2}$. i.e. a one-to-one matching between the elements of $\overline{\chi_1}$ and those of $\overline{\chi_2}$. For an incomplete sequence $\overline{\chi_P}$, or *predicted landmarks* and a reference sequence $\overline{\chi_T}$, or *tutor landmarks*, two important constraints may apply.

First, *repetitiveness* may be different, i.e. the number of predicted landmarks and the number of ground-truth landmarks of the *tutor* may be different, or

$$|\{l_i \in \overline{\chi_T}/\textsc{Type}(l_i) = \tau_k\}| \neq |\{l_j \in \overline{\chi_P}/\textsc{Type}(l_j) = \tau_k\}| \;\forall\; \tau_k \in \textsc{Type}(\overline{\chi_T}). \quad (1)$$

Second, the *symmetry* of the performance may also be different and, for some subsequence on the left hand side of $\overline{\chi_T}$, the synchronised landmarks of $\overline{\chi_P}$ may be at the end of the sequence, and vice-versa. That is,

$$\exists\; l_i \in \overline{\chi_T} \;/\; max\; \textsc{Frame}(S(l_i...l_n)) < min\; \textsc{Frame}(S(l_1...l_{i-1})). \quad (2)$$

A sequential approach has been followed in order to synchronise two sequences of landmarks $\overline{\chi_1}$ and $\overline{\chi_2}$, as shown in Algorithm 1, taking into account the constraints identified in (1) and (2).

**Algorithm 1.** Landmark sequences synchronisation algorithm

```
function Synchronise($\overline{\chi_1}$, $\overline{\chi_2}$)
    $S \leftarrow \{\}$                                   ▷ Empty synchronisation sequence to begin with
    $lastAssigned \leftarrow \overline{\chi_1}(N_1)$      ▷ $\overline{\chi_1}(1)$ will be tried to be matched first
    $assigned \leftarrow$ Repeat(False, $N_1$)           ▷ None has been matched to begin with
    for all $l_j \in \overline{\chi_2}$ do                ▷ Iterate over the landmarks in $\overline{\chi_2}$
        $type_2 \leftarrow$ Type($l_j$)                   ▷ Check, in order, all unassigned landmarks of $\overline{\chi_1}$
        for all $l_i \in \overline{\chi_1}((lastAssigned \,...\, N_1] \cup [1 \,...\, lastAssigned))$ do
            $type_1 \leftarrow$ Type($l_i$)               ▷ Match if the Type is the same
            if $type_1 = type_2$ AND $assigned(i) =$ False then
                $S(l_i) \leftarrow l_j$
                $assigned(i) \leftarrow$ True
                $lastAssigned \leftarrow l_i$
                break
            end if
        end for
    end for
    return $S$
end function
```

The essential idea of the algorithm is to do a pair-wise assignment between the two sequences (nominally, the predicted landmarks and the ground-truth landmarks). The algorithm goes over the landmark sequence to synchronise with, registering a match if the landmark type is the same and it has not been matched before.

### 3.2 Automatic Performing Numerical Assessment for Synchronised Motion Series

In order to automatically infer an assessment level, we train a classifier (C4.5 with AdaBoost, as seen in [9]) with a number of observations like the following:

$$\begin{aligned} O(\overline{\chi}) = &\{\textsc{Frame}(l_i), \textsc{Type}(l_i), \\ &\quad \{A(j, \textsc{Frame}(l_i)), |\overline{M_P}(\textsc{Frame}(l_i), j) - \overline{M_T}(\textsc{Frame}(S(l_i)), j)| \;\forall\; j\} \\ &\quad \} \;\forall\; l_i \;\in\; \overline{\chi}\}, \end{aligned} \tag{3}$$

where each $\overline{\chi}$ is a sequence of predicted landmarks of the motion sample or *pupil* $M_P$, $\chi_T$ is the ground-truth landmarks of the *tutor*, $S = \textsc{Synchronise}(\overline{\chi}, \overline{\chi_T})$ and $A(j,t)$ is a numeric assessment assigned to the joint $j$ in accordance with the pose in the frame $t$.

Essentially, the angular moment drift between the *pupil* and the *tutor* is calculated and, together with each involved joint's assessment level ($A$), a *multi-label*, *multi-target* machine classifier is trained, with one label per assessment level (nominally, four) and one output per involved joint. Each pair of $\{A(j, \textsc{Frame}(l_i)), |\overline{M_P}(\textsc{Frame}(l_i), j) - \overline{M_T}(\textsc{Frame}(S(l_i)), j)|\}$ is processed separately for each $j$.

The result is a sequence $\overline{\psi} = \{l_1^{\psi} \ldots l_N^{\psi}\}$ of *assessed landmarks* with

$$l_i^{\psi} = \{\textsc{Frame}(i), \textsc{Type}(i), \{\mathrm{A}(\textsc{Frame}(i), j)\}\forall j\} \;\forall\; l_i^{\psi} \in \overline{\psi}, \tag{4}$$

i.e. an extension of a landmark $l_i$ including a touple $A(\textsc{Frame}(i), j)$ of assessment label for each involved joint on $t = \textsc{Frame}(i)$.

### 3.3 Adaptation of *pupil*'s Performance to the *tutor*'s to Provide Visual Feedback

In addition to the numerical feedback, a super-imposed animation of the skeleton of the *pupil* and that of a reference is shown, in order to illustrate to the *pupil* what needs to be improved. Direct use of the performance of the *tutor* present several difficulties: among others, the body orientation, the scale and the timing may be different between the *tutor* and the *pupil*.

An affine transformation algorithm, depicted in Algorithm 2, is applied over each subsequence of a synchronisation $S$ in order to adapt a *tutor* motion data series $M_T$ to that of the *pupil* ($M_P$). Given a motion data series $M$, a landmark sequence $\overline{\chi}$ and a landmark type $\tau$, let $\overline{P}$ be the *mean pose per landmark type*, calculated as follows:

$$\overline{P}(M, \overline{\chi}, \tau, j) = \sum^{L_i \in L(\overline{\chi})} M(\textsc{Frame}(L_i), j)/|L(\overline{\chi})|, \tag{5}$$

where $L(\overline{\chi}) = l_i \in \overline{\chi} \;/\; \textsc{Type}(l_i) = \tau \;\forall\; \tau \in \textsc{Type}(\overline{\chi})$. This is the average body pose of a performance resulting on all landmark time frames of type $\tau$.

**Algorithm 2.** Motion data series adaptation algorithm

```
function AffineOffsetTransformation($M_1$, $M_2$, $\overline{\chi_1}$, $\overline{\chi_2}$)
    $M_A \leftarrow M_1$
    for all $l_i \in \overline{\chi_1}(2 \ldots N_1)$ do            ▷ Iterate over the predicted landmarks
        $\tau \leftarrow$ TYPE($l_i$)
        $f_i \leftarrow$ FRAME($l_i$)
        $f_{i-1} \leftarrow$ FRAME($\overline{\chi_1}(i-1)$)
        for all $j$ do                                        ▷ Iterate over the involved joints
            $\delta_1 \leftarrow (M_A(f_{i-1}, j)$ - $M_1(f_{i-1}, j))$          ▷ Angular drift w.r.t. $M_A$
            $\delta_2 \leftarrow (\overline{P}(M_2,\overline{\chi_2},\tau,j)$-$M_1(f_i, j))$       ▷ Mean angular drift w.r.t. $M_2$
            $\varDelta \leftarrow (\delta_2 - \delta_1)./(f_i - f_{i-1})$          ▷ Average angular drift per frame
             ▷ Progressively correct the overall drift by adding $\varDelta$ cumulatively
            for all $s_k \in M_1(f_{i-1} \ldots f_i$, j) do
                $M_A(k,j) \leftarrow s_k + (\delta_1 + \varDelta \cdot (k - f_{i-1}))$
            end for
        end for
    end for
    return $M_A$
end function
```

Given the *tutor*'s and *pupil*'s motion data ($M_T$ and $M_P$, respectively) and their respective ground-truth and predicted landmarks ($\overline{\chi_T}$ and $\overline{\chi_P}$), the value of $\overline{P}(M_P, \overline{\chi_P}, \tau))$ is assigned to each landmark $l_i \in \overline{\chi_T}$ of Type $\tau$.

Then, the motion data in between landmarks is recalculated applying an affine transform on the offset of the angular information. Therefore, rather than showing the animation of $M_P$ and $M_T$ superimposed, the proposed approach uses $M_P$ and AffineOffsetTransformation($M_P$, $M_T, \overline{\chi_P}, \overline{\chi_T}$.

## 4 Results

A series of experiments, involving stretches of *ankles*, *arms*, *calves*, *inner thighs* and *shoulders*, have been conducted to evaluate the proposed approach. All samples are captured using a Microsoft Kinect device and the OpenNI [1] middleware framework is used to retrieve the body joint positions and Euler angles. A third party application, Brekel Kinect [4], manages the above in order to produce BVH formatted files, which are the input for the experiments.

Nine individuals from both sexes –eight of them aged 25 to 31 and one of them aged 69– executed a series of performances of each exercise class. They were asked to perform each exercise simulating different physical conditions and abilities (normally, slower, faster and worse).

### 4.1 Evaluation Measurements

Each ground-truth landmark for each performance was manually labelled, according to stretches limit criteria, producing a different ground-truth landmark set $G$ for each of them. A range of frames may represent a key body pose (extensively due to transitional phases), rather than just a specific time frame. A *compromise interval* $[\Lambda_0, \Lambda_1]$, only considered for evaluation measurement purposes, not to train the system, has been established for each ground-truth landmark, so that $G = \{\{\textsc{Frame}_j, \textsc{Type}_j, \Lambda_{j0}, \Lambda_{j1}\}\}$.

The compromise intervals of each ground-truth landmark is calculated automatically prior to any accuracy computation. For a given ground truth landmark $g_i \in G = \{g_1..g_n\}$ *difference threshold* $\Delta_\Lambda \in [0, 1]$ of a motion series $M = \{M_1..M_N\}$, the interval limits $\Lambda_0$ and $\Lambda_1$ are calculated as

$$\begin{aligned}
\Lambda_0 = min\ &\{\Lambda \in [\textsc{Frame}(g_{i-1}), \textsc{Frame}(g_i)]\ /\\
&\|M_t - M_{\textsc{Frame}(g_i)}\|_1 \leq \|M_{\textsc{Frame}(g_i)} - M_{\textsc{Frame}(g_{i-1})}\|_1 \cdot \Delta_\Lambda\\
&\forall t \in [\Lambda, \textsc{Frame}(g_i)]\}\\
\Lambda_1 = max\ &\{\Lambda \in [\textsc{Frame}(g_i), \textsc{Frame}(g_{i+1})]\ /\\
&\|M_t - M_{\textsc{Frame}(g_i)}\|_1 \leq \|M_{\textsc{Frame}(g_i)} - M_{\textsc{Frame}(g_{i+1})}\|_1 \cdot \Delta_\Lambda\\
&\forall t \in [\textsc{Frame}(g_i), \Lambda]\},
\end{aligned} \tag{6}$$

where $\|\cdot\|_1$ is the Taxicab norm, $\textsc{Frame}(g_0) = 1$ and $\textsc{Frame}(g_{n+1}) = N$.

*Precision* ($\pi$), *accuracy* ($\alpha$) and the *false positive* ($fp$) and *false negative* ($fn$) rates were chosen to measure the overall performance.

### 4.2 Precision Achieved on Landmark Prediction and Assessment

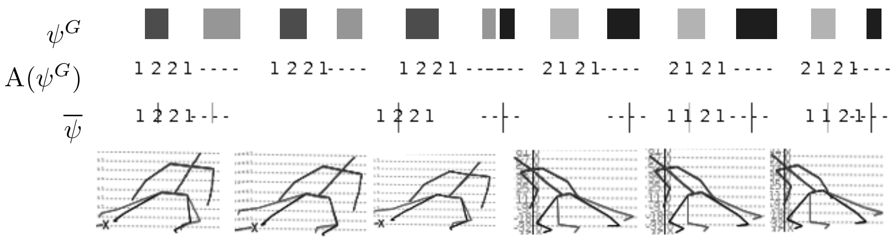
Table 1 shows the precision achieved with the proposed approach, including the compromise interval criteria for precision calculation with $\Delta_\Lambda = 0.10$ (i.e. maximum 10% motion drift between landmarks), using all 3 DOF ($x$, $y$ and $z$ axes) motion data of each involved joint and 4 assessment levels (i.e. $A(j, t) =$

**Table 1.** Precision achieved on both landmark prediction and assessment. $G$=avg. number of ground-truth landmarks per sample, $Ts$=number of training samples, $Tf$=avg. number of training frames and $Tsf$=avg. number of testing frames

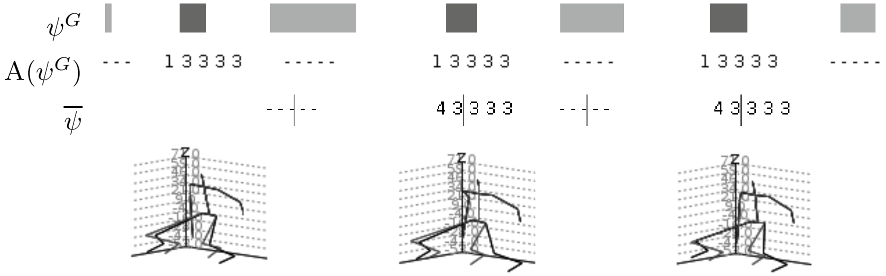
| Exercise | Statistics | | | | Prediction | | | | Assessment | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | G | Ts | Tf | Tsf | $\pi$ | $\alpha$ | $fp$ | $fn$ | $\pi$ | $\alpha$ | $fp$ | $fn$ |
| Ankles | 13.46 | 46 | 54 | 586 | 0.80 | 0.91 | 1.51 | 6.56 | 0.84 | 0.77 | 3.78 | 3.78 |
| Arms | 6.65 | 52 | 27 | 413 | 0.72 | 0.90 | 1.39 | 3.19 | 0.86 | 0.82 | 1.94 | 1.94 |
| Calves | 13.28 | 39 | 53 | 638 | 0.72 | 0.88 | 2.36 | 6.95 | 0.86 | 0.86 | 1.05 | 1.05 |
| Inner thighs | 6.28 | 50 | 25 | 340 | 0.84 | 0.93 | 0.64 | 1.78 | 0.81 | 0.74 | 4.17 | 4.17 |
| Shoulders | 9.21 | 48 | 37 | 391 | 0.66 | 0.72 | 1.40 | 6.20 | 0.88 | 0.85 | 1.36 | 1.36 |
| **Average** | **9.78** | **47** | **39** | **474** | **0.76** | **0.88** | **1.29** | **4.23** | **0.85** | **0.80** | **2.69** | **2.69** |

$[-,1,2,3,4]\ \forall j,t)$. Also, cross-validation of fixed-size folds (size: 7) and 3 repetitions were carried out for each exercise class.

Figure 3 shows a graphical representation of the results of two different exercises illustrating landmarks prediction and assessment, showing also the adapted poses of the *pupil* on a number of ground-truth landmarks after synchronisation with the *tutor*'s.



(a) Ankle stretches - Involved joints: left hip, left knee, right hip and right knee



(b) Inner thigh stretches - Involved joints: spine, left hip, left knee, right hip and right knee

**Fig. 3.** Sample results of landmarks prediction and assessment

The markers in $\psi^G$ depict the compromise intervals of the ground-truth landmarks. A($\psi^G$) is the assessment values of $\psi^G$ as shown in the figure. $\overline{\psi}$ shows the predicted landmarks and assessment values. Different grey tones represent different landmark TYPES, i.e., Range of Motion (*ROM*) extrema. The representation of the *pupil*'s pose is in black and its adaptation to that of the *tutor* is shown in grey. Each assessment level corresponds, in respective order, to an involved body joint. In this case, the smaller the value, the better the performance

## 5 Conclusions and Future Work

Empirical work showed that the data extracted with Kinect and OpenNI is not very accurate, finding the same kind of issues reported in [17]. Self-occlusion of body limbs produces only moderately satisfactory outcome on *landmark* prediction, producing both false positives and negatives (i.e., non-existent and missed landmarks, respectively). This becomes even more problematic when performance measurements, like speed, are to be calculated upon *landmark* sequencing information. However, the studied exercises involve repetitions so it is very likely that at least one landmark will be detected. Furthermore, the accuracy of the joint angles and positions calculated by the framework is accurate enough to prove the point of this research.

Assessment of a body pose in objective terms is a complex task. For example, if the judgement is done through visual comparison of a body joints representation then not every angle of perspective will be covered. Also, one may give more importance to factors like balance or posture than to the range of stretch itself. A further source of ambiguity can be found if all three degrees of freedom of each joint are taken into account. Likewise, the chosen motion sensing framework is not designed to discard (or correct) impossible poses.

Nevertheless, even though the motion capture resources were chosen upon affordability and ease-of-use criteria, it is our belief that using a more advanced motion capture dataset (e.g. Carnegie Mellon University Motion Capture Database at `http://mocap.cs.cmu.edu/`) should give better results. Future work will focus on further abstracting the way assessment is produced, like translating every touple $A(\textsc{Frame}(i), j)$ into human-understandable language. Further results are available in the research Web page at `http://homepages.lboro.ac.uk/~copf3/research.htm`.

## References

1. Ben-Horin, B.: Introducing OpenNI (2011), `http://www.openni.org/` (online; accessed 2012)
2. Bickmore, T.W., Schulman, D., Sidner, C.L.: A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. Journal of Biomedical Informatics 44(2), 183–197 (2011)
3. Blanchard, E., Chalfoun, P., Frasson, C.: Towards advanced learner modeling: discussions on quasi real-time adaptation with physiological data. In: Seventh IEEE International Conference on Advanced Learning Technologies (ICALT), pp. 809–813 (2007)
4. Brekelmans, J.: Brekel Kinect (2011), `http://www.brekel.com/?page_id=155` (online; accessed 2014)
5. Buttussi, F., Chittaro, L.: Mopet: A context-aware and user-adaptive wearable system for fitness training. Artificial Intelligence in Medicine 42(2), 153–163 (2008)
6. Buttussi, F., Chittaro, L., Ranon, R., Verona, A.: Adaptation of graphics and gameplay in fitness games by exploiting motion and physiological sensors. In: Butz, A., Fisher, B., Krüger, A., Olivier, P., Owada, S. (eds.) SG 2007. LNCS, vol. 4569, pp. 85–96. Springer, Heidelberg (2007)
7. Dasgupta, A., Nakamura, Y.: Making feasible walking motion of humanoid robots from human motion capture data. In: IEEE International Conference on Robotics and Automation, vol. 2, pp. 1044–1049 (1999)

8. Delson, N., West, H.: Robot programming by human demonstration: Adaptation and inconsistency in constrained motion. In: IEEE International Conference on Robotics and Automation, vol. 1, pp. 30–36 (1996)
9. Fernández de Dios, P., Chung, P.W., Meng, Q.: Landmark-based methods for temporal alignment of human motions. IEEE Computational Intelligence Magazine 9(2), 29–37 (2014)
10. Fernández de Dios, P., Meng, Q., Chung, P.W.: A machine learning method for identification of key body poses in cyclic physical exercises. In: IEEE International Conference on Systems, Man and Cybernetics (SMC), pp. 1605–1610 (October 2013)
11. Fox, J., Bailenson, J.N.: Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors. Media Psychology 12(1), 1–25 (2009)
12. Gleicher, M.: Retargetting motion to new characters. In: Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, pp. 33–42. ACM (1998)
13. Kim, S.K., Chun, J., Park, H., Kwon, D., Ro, Y.: Interfacing sensors and virtual world health avatar application. In: 5th IEEE International Conference on New Trends in Information Science and Service Science (NISS), vol. 1, pp. 21–25 (2011)
14. Klaassen, R., Lavrysen, T., Geleijnse, G., van Halteren, A., Schwietert, H., van der Hout, M., et al.: A personal context-aware multi-device coaching service that supports a healthy lifestyle. In: 25th BCS Conference on Human-Computer Interaction, pp. 443–448 (2011)
15. Kurillo, G., Han, J.J., Nicorici, A., Bajcsy, R.: Tele-mfast: Kinect-based telemedicine tool for remote motion and function assessment. In: Medicine Meets Virtual Reality 21: NextMed/MMVR21, vol. 196, p. 215 (2014)
16. Lisetti, C., Nasoz, F., LeRouge, C., Ozyer, O., Alvarez, K.: Developing multimodal intelligent affective interfaces for tele-home health care. International Journal of Human-Computer Studies 59(1), 245–255 (2003)
17. Obdrzalek, S., Kurillo, G., Ofli, F., Bajcsy, R., Seto, E., Jimison, H., Pavel, M.: Accuracy and robustness of kinect pose estimation in the context of coaching of elderly population. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1188–1193. IEEE (2012)
18. Pirovano, M., Mainetti, R., Baud-Bovy, G., Lanzi, P.L., Borghese, N.A.: Self-adaptive games for rehabilitation at home. In: IEEE Conference on Computational Intelligence and Games (CIG), pp. 179–186 (2012)
19. Wieringa, W., op den Akker, H., Jones, V.M., op den Akker, R., Hermens, H.J.: Ontology-based generation of dynamic feedback on physical activity. In: Peleg, M., Lavrač, N., Combi, C. (eds.) AIME 2011. LNCS, vol. 6747, pp. 55–59. Springer, Heidelberg (2011)

# A Survey on Energy-Aware Profiler for Mobile Devices

Muhammad Hanif Jofri, Mohd Farhan Md Fudzee, and Mohd Norasri Ismail

Multimedia and Mobile Group, Software & Multimedia Center,
Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn, Malaysia
`{farhan,norasri}@uthm.edu.my`

**Abstract.** Contemporary designs for mobile devices sometimes lack of energy-aware profiler thus require an optimized use of system power. Energy-aware profiler is one of the major parts in the energy management that plays an important role to minimize energy consumption in mobile devices. Many different methods and research efforts on energy-aware profiler have been discussed and proposed during the past few years. However, there is a little effort on summarizing and providing the generic requirement for energy-aware profilers in relation to mobile devices. In this paper we review several areas of classification in energy-aware profilers that have been recommended over the recent years. Moreover, we also discuss briefly on the categories, approaches and open challenges that related to energy profiler.

**Keywords:** Energy-aware, energy profiler classification, mobile devices.

## 1 Introduction

The mobile device technology nowadays has become more sophisticated. In order to keep up with this hi-tech gadgets, optimizing the phone energy is one of the primary intentions to enhance battery life of mobile devices. Energy-aware profiler is one of the important parts of energy management and it may be used to reduce energy being consumed by application and services on mobile devices. In general, an energy-aware profiler has a certain functionality to indicate and to collect data from the user, operating system, policy and application. It may also act as a set of Application Programming Interface (API) and adaptation policy that manage properties related to energy profiler and energy adaptation [1], [2]. Fig 1 is a generic energy-aware profiler application measures the energy consumption related to the beginning of the process to the end of all processes. Firstly the data from the user, the mobile operating system, services as well as hardware are captured by the energy-aware profiler. Then the energy-aware profiler will gather all the specific data such as energy level and hardware status, and transfers it to the energy management module where all data obtained will be analyzes and manage to determine the energy consumption. All the measured data will be segregated by the usage of energy consumption [3].

The analyzed data then will be used as output within mobile applications and services [4]. This design architecture usually consists of three main parts, energy source,

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_29

energy profiler and energy consumption [5], [6]. Then energy consumption on a mobile device can be measured based on the usages from profiling elements [7], [8]. Even though there are many researches on energy profilers, but there are less effort to summarize the requirements for energy-aware profiler.

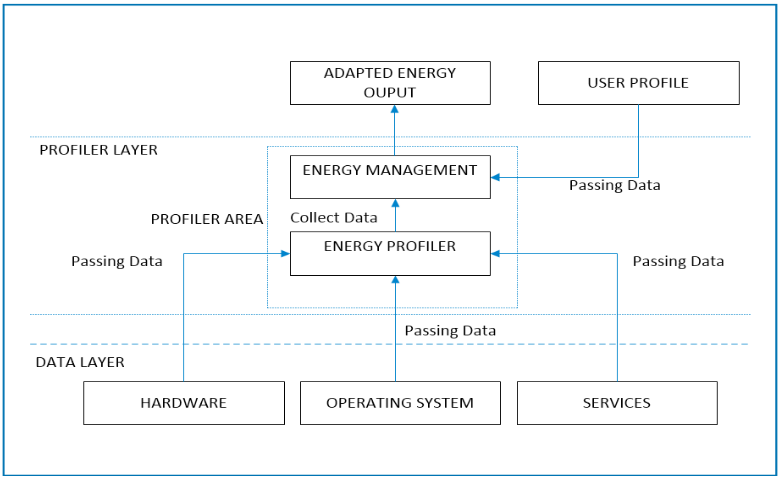


**Fig. 1.** Generic Energy-Aware Profiler Building Blocks in Mobile Device.

This paper aims to provide readers the categories and the approaches of classification in energy-aware that already exist. Furthermore, we discuss on understanding of a minimum building blocks in energy-aware profiler. The contribution of the paper can be divided into three: (1) a generic energy-aware profiler building blocks in mobile devices, (2) a classification of existing researchers into five non-overlapping categories, and (3) open challenges in relation areas of energy-aware profiler. The rest of this paper is organized as follows. In Section 2, we introduce the possible classification of energy profiler. We then present, in Section 3, interrelation between profiler categories. Finally, section 4 presents the open challenges in relation with the profiler to the research community to deal with.

## 2 Classification of Energy Profiler

Several efforts on these issues have been carried throughout the recent years to improve the energy consumption of mobile devices while concurrently providing all the necessary functionality. The elements can be classified into five categories; instrument, policies, perspective, procedures and areas. From these distinctive categories we can discretely define it to several approaches that rely on common energy profiler categories (refer Fig 2).

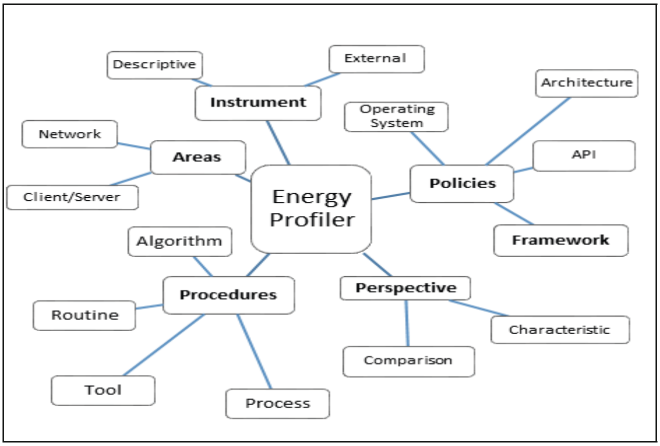


**Fig. 2.** Categories of Generic Energy-Aware Profiler

The first category is Instrument. It can be referred to actual experiment, testing and training on the device itself in order to get the specific of energy consumption result. There are two specific approaches in this category: Descriptive and External. These specific approaches are using certain devices or apparatus for testing and to measuring the energy consumption purpose [1], [2].

The second category is Policies. This category is referring to the reliable and useable set of outline structure of energy-aware profiler approaches. There are four approaches in this category: Operating System, API, Framework and Architecture. Every particular approach within this category provides a specific orientation related to energy-aware profiler [7], [9]. Moreover, it is vital method for energy-aware profiler on fundamental level. Also, a discussion of this category will be deliberated later on within this paper.

The third category is Perspective. In this category there are two approaches: Characteristic and Comparison of energy-aware profiler. These approaches are used to evaluate and distinguish a specific energy consumption based on result and finding [10]. Also by evaluating the comparison between two specific elements in energy-aware profiler, debugging process is available for estimating energy efficient tool [11]. The evaluating process of characteristic and comparison is important to measure how much the accuracy of input and output energy consumption on mobile devices.

The fourth category is the Procedure. Procedure defines the series of step and standard strategy that can be proceeded to reduce energy consumption on energy-aware profiler. The approaches in this category are Routine, Process, Tool and Algorithm. Manageable energy consumption data on mobile devices can increase the efficiency of battery life [4], [12].

The last category of a common energy-aware profiler is Areas. Area is referring to the locality and proportion of network running on energy-aware profiler. Within this category, there are two approaches: client/server and network. This category is used, in order to designing and structuring energy-aware profiler for mobile devices [13], [14], [15].

## 2.1 Instrument

In order to set an accurate reading or measurement of energy level, instrument is the key factor to determine good result. It can be refer to an experiment, testing and training on the device itself in order to get the specific result of energy consumption.

The first approach within this category is Descriptive. This approach requires tool not only from apparatus and devices but also includes advanced technology (e.g., digital multimeter) and software from computer to capture the energy consumption data [2], [6], [16]. By using the power modeling and used a native tool, energy-aware profiler can measure the performance of energy rating in order to minimize energy consumption [14], [17], [18], [19], [20]. Another approach in the instrument category is External. It refers to the actual testing on devices without including software such as apparatus, actual mobile device, measurement tool like analog multimeter to measure the power output from mobile device [21], [22], [23]. Table 1 shows the instrument category from existing researchers grouped by approaches.

**Table 1.** Instrument Categories

| Approaches | Researchers |
|---|---|
| **Descriptive** | Balasubramanian, et. al [1], Thiagarajan, et. al[2], Chung, et. Al[6], Hahm, et. al[13], Corral, et. al[14], Bunse, et. al[10], Xiao, et. al[19], Shin, et. al[21], Hu & Ruutu[25], Murmuria, et. al., [20] |
| **External** | Balasubramanian, et. al [1], Thiagarajan, et al[2], Ardito, et. al[4], Qian, et. al[9], Hahm, et. al[13], Chang, et. al[16], Bunse, et. al[10], Shin, et, al[21], Tsao, et. al[22], Sabharwal, et. al[11], , Cuervo, et. al[24], |

Instrument category of energy-aware profiler usually need context of work, mechanism of tool and energy comparison which allow executing defined scenarios on mobile devices and profiling the related power consumption through external measurement hardware [13], [11], [24]. In this category also, analyzed data from hardware implementation need to be done, as well by analyzing the implementation of mobile application. Moreover, it provides code-level estimate of energy that it will consume at runtime [9]. The instrument category combines a set of software profiler and energy measurement hardware that important to analyze the process-level and function [22], [25]. On hardware layer sometimes researchers use the Linux Kernel to manage energy consume by services of mobile devices. This is relatively determined by mobile application usage in order to saving more energy [26].

## 2.2 Policies

The second category within energy-aware profiler classes is the policies. This category governs the operational structure of energy-aware profiler approaches. Commonly, operating system approach has a special characteristic and it governs energy-aware operation of mobile devices. First approach is Operating System (OS).

On this level mobile operating system function is to manage the resources for energy saving purposed. This approach compares the mobile operating system in term of efficiency in profiler energy consumption [14], [27], [28].

Second approach is Application Programming Interface (API). Energy-aware function for API is to manage activity of energy usage pattern by every application running [24], [29]. To determine energy level, energy usage pattern on an application is catered by API services [11], [30].

The third approach is Framework. Framework design depends on how energy being consumed on devices and framework planning regarding energy, services and hardware interacted [17], [29], [31], [32], [33]. Framework design for energy-aware profiler is important to distinguish between service-quality and energy model [21], [34], [35].

Last approach for policies category is Architecture. It corresponds to the planning and the designing of energy-aware profiler model. Currently, the generic energy-aware profiler in a mobile device is not well-designed to provide energy profiler in the best way; therefore users cannot easily predict and switch the lifetime of battery policies. Table 2 shows the policies category from existing researchers grouped by approaches.

**Table 2.** Policies Categories

| Approaches | Researchers |
|---|---|
| **Operating System** | Ardito, et. al[5], Corral, et. al[14], Ding et. al[15], Cho, et. al[28], Kundu et. al[42] |
| **API** | Hao, et. al[37], Sabharwal et. al[11], Do et. al[30] |
| **Framework** | Chen, et. al[3], Shin, et. al[21], Wen et. al[17], Shin, et. al[21], Li et. al[29], Bareth, [33], Cho, et. al[28], Cuervo, et. al[24], Kundu, et. al[42] |
| **Architecture** | Chung et. al[6], Ding & Muntean, [15], Chang, et. al[16], Chang et. al[39], Tsao, et. al[22], (Carroll, et. al[36], Cuervo, et. al[24] |

## 2.3 Perspective

Basically, perspective category is the indication of features and aspect regarding energy-aware profiler on mobile devices. First approach is Characteristic. In this approach, it will segregate and divided the process functionality based on the energy characteristic. Energy layer component is the main function of energy profiler in mobile devices [5], [14], [34], [38]. Estimating how the profiler works required some modification to the operating system and application [39], [40].

The second approach is the Comparison. In this approach, the main idea is to define the contrast analysis on energy-aware profiler. The analysis is an important process which data gather from power tier and energy consumption characteristics quantify the relative energy ratio during usage between a general purpose mobile device and a set of dedicated devices [13], [16], [19], [25]. This also shows how these characteristics and comparison can be translated into viewpoint about energy consumption and energy profiler regarding battery life under a number of usage patterns. Table 3 shows the perspective category from existing researchers grouped by approaches.

**Table 3.** Perspective Categories.

| Approaches | Researchers |
|---|---|
| **Characteristic** | Balasubramanian, et. al[1], Thiagarajan, et. al[2], Bernal, et al[5], Ardito, et. al[8], Ismail, et. al[12], Corral, et. al[14], Wen, et. al[17], Bunse, et. al[10], Yu Xiao, et. al[19], ], Jelschen, et. al[34], Metri, et. al[35], Carroll, et. al[36], Bedregal, et. al[38] |
| **Comparison** | Balasubramanian, et al.[1], Thiagarajan, et. al[2], Bernal, et. al.[5], Ardito, et. al[8], Hahm, et. al[13], Corral, et. al[14], Chang, et. al[16], Xiao, et. al[19], Ruutu, [25], Metri, et. al.[35], Kundu, et. al.[42] |

## 2.4 Procedure

There are various procedures have been developed in relation to energy-aware profiler. The first approach is the Routine. Routine is a technique of profiler ability to arrange possible outcome for energy consumed by mobile devices. The energy measurement system is significant for assisted user to perform an automation technique to reduce energy consumption in mobile devices [2]. Table 4 shows the procedure category from existing researchers grouped by approaches.

**Table 4.** Procedure Categories

| Approaches | Researchers |
|---|---|
| **Routine** | Thiagarajan, et. al[2], Chen, et. al [3], Ardito, [4], Chung, et. al[6], Qian, et. al[9], Corral, et. al[14], Chang, et. al[16], Ding, et. al[26], Tsao, et. al[22], Ge, et. al[23], |
| **Application** | Chung et al[6], Hao et al[31], Pathak et al[18], Schubert, et. al[7], Hao, et. al[37], |
| **Process** | Bernal, et. al[5], Schubert, et. al[7], Hahm, et. al[13], Corral, et. al[14], Chang, et. al[16], Ding, et. al[26], Murmuria, et. al[20], Li, et. al[29], Bareth, [33], Bedregal, et. al[38], Cuervo, et. al[24] |
| **Tool** | Ardito, [4], Chung, et. al[6], Hao, et. al[31], Pathak, et. al[18], Bunse, et. al[18], Ge et al[23], Kwon, et. al[27], Hao, et. a.[37], Kjrgaard, et. al.[32], Bareth, [33], Cho, et. al[28], Metri, et. al[35], Do, et. al[30], Cuervo, et. al[24] |
| **Algorithm** | Qian, et. al[9], Bunse, et. al[10], Murmuria, et. al[20], Kwon, et. al[27], Jelschen, et. al[34], Li, et. al[29], Kjrgaard, et. al[32], Bareth, [33], Cuervo, et. al[24], |

Second approach is Application. Some of applications are build-in with smart power consumption to reduce energy usage. By using a workload generator energy consumption can be reduced considerably [7]. Some of the applications were instrumented to record the energy state transitions of various resources such as the processor and network interface [11].

The next approach is Process. This approach is crucial to ensure the flow of energy consumption analysis. It is a state to verify the flow of energy consumption in different scenarios [11] , [20], [26].

Next approach is Tool. A pTop and PowerPack is one of the tool that estimated the usage of energy of application and how much services effected energy on mobile devices [23], [37].

The last approach is the Algorithm. A sorting of algorithms that were used in application to reduce energy consumption [28]. To reduce energy usage in a mobile device the decent algorithm can be produced by optimizing code functionality of a program [18], [27].

## 2.5 Areas

Area is the last category of energy-aware profiler building blocks. It reflects the network-related of an energy-aware profiler on mobile devices. There are two approaches; Client/Server approach and Network approach. The first approach is Client/Server. Server optimization will enhanced energy-oriented system profiling and content adaptation of network ability within devices usage [9], [15].

Second approach is the Network approach. This method uses the mobile network characteristic such as 3G, GSM or Wi-Fi to profiling energy consumption in mobile devices. Wi-Fi protocol is used to overcome energy consumption based on developed energy profiler [19]. Energy-aware profiler go through every these mobile network platform to get good result on energy usage. Table 5 shows the areas category from existing researchers grouped by approaches.

**Table 5.** Areas Categories

| Approaches | Researchers |
|---|---|
| **Client/Server** | Balasubramanian, et. al[1], Thiagarajan, et. al[2], Qian, et. al[9], Ding and Muntean [15], Wen, et. al[17] |
| **Network** | Thiagarajan, et. al[2], Xiao, et. al[19], Ding, et. al[26], Shin, et. al[21], Tsao, et. al[22], Hu and Ruutu [25], Cuervo, et. al[24] |

## 3 Interrelations between Categories

The current solutions on energy-aware profiler usually focus on the delivery application without considering applications that run the environment and device features that pose different energy constraints on the whole content and delivery process. Some of existing profiler interacts with two or three categories in order to achieve good result of energy-aware profiler and measurement tool [2], [3], [6], [10]. Some of reserchers specification more prior in some categories but up until now there are no research on combining all five categories. To the best of our knowledge, there is no minimal specific interaction between categories. However, several existing researches have conducted some experiment and testing where they use more than one category for designing, developing and structuring their energy-aware profiler.

## 4 Open Challenges in Energy-Aware Profiler

There are several open challenges for researchers regarding energy aware profiler. Firstly most of the challenges in energy-aware profiler can be related to content adaptation. Content adaptation is the action of transforming content to adapt to various device interfaces. This process needs high energy usage. In order to manage this matter, energy-aware profiler manages will capture raw data from hardware, operating system, services and the user. Existing content adaptation systems deploy various techniques which have been developed for specific purposes and goals [40], [41], [42]. Second challenge is within network performances. In network performances, energy-aware profiler main challenge is to cater the speed, jitter and robust of the network coverage. This process will enable user to hold the battery power longer in their mobile devices. The third challenge is environments. Environment is referring to the mobile operating system area. There are many environments and platforms exist nowadays. Different architecture of mobile operating system may consume more energy usage in mobile devices. However, some of these challenges are on working progress and several more still need modification and enhancement.

## 5 Conclusion

This paper develops on and contributes to the work of profiling in relation to energy-aware and the future work potentially be carried out by other researchers. First we classified the researchers on energy-aware profiler into five non-overlapping building blocks or categories. Then we present the five categories that include instrument, policies, perspective, procedure, and areas. Second we discussed a minimum building block of energy-aware profiler. Finally, open challenge in relation of energy-aware profiler is discussed in order to saving energy in mobile devices.

**Acknowledgment.** The authors would like to acknowledge Universiti Tun Hussein Onn Malaysia and Ministry of Education, Malaysia for the Fundamental Research Grant Scheme 1239 provided in financing the research activities.

## References

1. Niranjan, B., Aruna, B., Arun, V.: Energy Consumption in Mobile Phones: A Measurement Study and Implications for Network Applications. In: IMC 2009, Chicago, Illinois, USA, November 4-6 (2009)
2. Thiagarajan, N., Aggarwal, G., Nicoara, A., Boneh, D., Singh, J.P.: Who Killed My Battery: Analyzing Mobile Browser Energy Consumption. In: WWW 2012 – Session: Mobile Web Performance, Lyon, France (2012)
3. Chen, H., Luo, B., Shi, W.: Anole: A Case for Energy-Aware Mobile Application Design. In: 2012 41st International Conference on Parallel Processing Workshops, Pittsburgh, USA (2012).
4. Ardito, L.: Energy Aware Self-Adaptation in Mobile Systems. In: 35th International Conference on Software Engineering (ICSE 2013), San Francisco, CA, USA (2013)
5. Bernal, J.F., Mejia, A.L., Morisio, M., Falcarin, P.: Towards an Efficient Context-Aware System: Problems and Suggestions to Reduce Energy Consumption in Mobile Devices. In: ICMB/GMR 2010, 2010 Ninth International Conference on Mobile Business / 2010 Ninth Global Mobility Roundtable, Athen, Greece (2010)
6. Chung, Y.-F., Lin, C.-Y., King, C.-T.: ANEPROF: Energy Profiling for Android Java Virtual Machine and Applications. In: IEEE 17th International Conference on Parallel and Distributed Systems (ICPADS 2011), Tainan, Taiwan (2011)
7. Schubert, S., Kostic, D., Zwaenepoel, W., Shin, K.: Profiling Software for Energy Consumption. In: 2012 IEEE International Conference on Green Computing and Communications (GreenCom), Besançon, France (2012)
8. Ardito, L., Procaccianti, G., Torchiano, M., Migliore, G.: Profiling Power Consumption on Mobile Devices. In: ENERGY 2013: The Third International Conference on Smart Grids, Green Communications and IT Energy-aware Technologies, Lisbon, Portugal (2013)
9. Qian, F., Wang, Z., Gerber, A.: Profiling Resource Usage for Mobile Applications: A Cross-layer Approach. In: MobiSys 2011, Bethesda, Maryland, USA (2011)
10. Bunse, C., Hopfner, H., Mansour, E., Roychoudhury, S.: Exploring the Energy Consumption of Data Sorting Algorithms in Embedded and Mobile Environments. In: 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, Taipei, Taiwan (2009); Pathak, A., Hu, Y.-C., Zhang, M.: Where is the energy spent inside my app? Fine Grained Energy Accounting on Smartphones with Eprof. In: EuroSys 2012 Proceedings of the 7th ACM European Conference on Computer Systems, Bern, Switzerland (2012)
11. Sabharwal, M., Metri, G., Huang, C., Agrawal, A.: Towards Fine Grain Power Profiling Tools for SoC based Mobile Devices. In: 2013 4th Annual International Conference on Energy Aware Computing Systems and Applications (ICEAC), Istanbul, Turkey (2013)
12. Norasri, M.-I., Rosziati, I., Farhan, M.-F.: A Survey on Content Adaptation Systems towards Energy Consumption Awareness. Advances in Multimedia 2013, Article ID 871516, 8 pages (2013)
13. Hahm, O., Adler, S., Schmittberger, N., Gunes, M.: Energy Profiling for Wireless Sensor Networks. In: GI/ITG KuVS Fachgespräch Sensornetze, Berlin, Germany (2011)

14. Corral, L., Georgiev, A.-B., Sillitti, A., Succi, G.: A Method for Characterizing Energy Consumption in Android Smartphones. In: 2013 2nd International Workshop on Green and Sustainable Software (GREENS), San Francisco, CA, USA (2013)
15. Ding, R., Muntean, G.-M.: Device Characteristics-based Differentiated Energy-efficient Adaptive Solution for Video Delivery over Heterogeneous Wireless Network. In: 2013 IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China (2013)
16. Chang, H.-C., Agrawal, A., Cameron, K.: Energy-Aware Computing for Android Platforms. In: 2011 International Conference on Energy Aware Computing (ICEAC), Istanbul, Turkey (2011)
17. Wen, Y., Zhang, W., Luo, H.: Energy-Optimal Mobile Application Execution: Taming Resource-Poor Mobile Devices with Cloud Clones. In: The 31st Annual IEEE International Conference on Computer Communications: Mini-Conference, Orlando, FL, USA (2012)
18. Bunse, C., Hopfner, H., Mansour, E., Roychoudhury, S.: Exploring the Energy Consumption of Data Sorting Algorithms in Embedded and Mobile Environments. In: 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, Taipei, Taiwan (2009)
19. Xiao, Y., Cui, Y., Savolainen, P., Siekkinen, M., Wang, A., Yang, L., Jaaski, A.-Y., Tarkoma, S.: Modeling Energy Consumption of Data Transmission over Wi-Fi. IEEE Transactions on Mobile Computing PP(99), 1–14 (2013)
20. Murmuria, R., Medsger, J., Stavrou, A., Voas, J.: Mobile Application and Device Power Usage Measurements. In: 2012 IEEE Sixth International Conference on Software Security and Reliability, Gaithersburg, MD, USA (2012)
21. Shin, D., Kim, K., Chang, N., Lee, W., Wang, Y., Xie, Q., Pedram, M.: Online Estimation of the Remaining Energy Capacity in Mobile Systems Considering System-Wide Power Consumption and Battery Characteristics. In: 2013 18th Asia and South Pacific Design Automation Conference (ASP-DAC), Yokohama, Japan (2013)
22. Tsao, S.-L., Kao, C.-C., Suat, I., Kuo, Y., Chang, Y.-H., Yu, C.-K.: PowerMemo: A Power Profiling Tool for Mobile Devices in an Emulated Wireless Environment. In: 2012 International Symposium on System on Chip (SoC), Tampere, Sweden (2012)
23. Ge, R., Feng, X., Song, S., Chang, H.-C., Li, D., Cameron, K.: PowerPack: Energy Profiling and Analysis of High-Performance Systems and Applications. IEEE Transactions on Parallel and Distributed Systems 21(5), 1–14 (2010)
24. Cuervo, E., Balasubramanian, A., Cho, D.-K., Wolmanx, A., Saroiux, S., Chandrax, R., Bahlx, P.: MAUI: Making Smartphones Last Longer with Code Offload. In: Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services, MobiSys 2010, San Francisco, California, USA (2010)
25. Hu, Z., Ruutu, J.: Comparison of Energy Consumption Between A Mobile Device and A Collection of Dedicated Devices. In: 2011 IEEE International Symposium on Sustainable Systems and Technology (ISSST), Chicago, IL, USA (2011)
26. Ding, F., Xia, F., Zhang, W., Zhao, X., Ma, C.: Monitoring Energy Consumption of Smartphones. In: 2011 IEEE International Conferences on Internet of Things, and Cyber, Physical and Social Computing, Dalian, China (2011)
27. Kwon, Y.-W., Tilevich, E.: Reducing the Energy Consumption of Mobile Applications Behind the Scenes. In: 2013 IEEE International Conference on Software Maintenance, Eindhoven, Netherlands (2013)
28. Cho, J., Woo, Y., Kim, S., Seo, E.: A Battery Lifetime Guarantee Scheme for Selective Applications in Smart Mobile Devices. IEEE Transactions on Consumer Electronics 60(1), 155–163 (2014)

29. Li, D., Hao, S., Halfond, W.-G.: Calculating Source Line Level Energy Information for Android Applications. In: International Symposium on Software Testing and Analysis, ISSTA 2013, Lugano, Switzerland (2013)
30. Do, T., Rawshdeh, S., Shi, W.: pTop: A Process-level Power Profiling Tool. In: Proceedings of the 2nd Workshop on Power Aware Computing and Systems, HotPower 2009 (2009)
31. Hao, S., Li, D., Halfond, W.G.J., Govindan, R.: Estimating Mobile Application Energy Consumption using Program Analysis. In: 35th International Conference on Software Engineering (ICSE 2013), San Francisco, USA (2013)
32. Kjergaard, M.B., Blunck, B.H.: Unsupervised Power Profiling for Mobile Devices. In: 8th International ICST Conference, MobiQuitous 2011, Copenhagen, Denmark (2011)
33. Bareth, U.: Simulating Power Consumption of Location Tracking Algorithms to Improve Energy-efficiency of Smartphones. In: 2012 IEEE 36th International Conference on Computer Software and Applications, Washington, DC, USA (2012)
34. Jelschen, J., Gottschalk, M., Josefiok, M., Pitu, C., Winter, A.: Towards Applying Reengineering Services to Energy-Efficient Applications. In: 2012 16th European Conference on Software Maintenance and Reengineering, Szeged, Hungary (2012)
35. Metri, G., Agrawal, A., Peri, R., Brockmeyer, M., Shi, W.: A Simplistic Way for Power Profiling of Mobile Devices. In: 2012 International Conference on Energy Aware Computing, Guzelyurt, Cyprus (2012)
36. Carroll, A., Heiser, G.: An Analysis of Power Consumption in a Smartphone. In: Proceedings of the 2010 USENIX Conference on USENIX Annual Technical Conference, USENIXATC 2010, Berkeley, CA, USA (2010)
37. Hao, S., Li, D., Halfond, W.: G., J., Govindan, R.: Estimating Android Applications' CPU Energy Usage via Bytecode Profiling. In: GREENS 2012, Zurich, Switzerland (2012)
38. Bedregal, J.C.V., Gutierrez, E.G.C., Arisaca M, R.E.: Optimizing Energy Consumption per Application in Mobile Devices. In: 2013 International Conference on Information Society (i-Society), Toronto, ON (2013)
39. Chang, K.-C., Kao, C.-Y., Su, B.-G.: A Low-Overhead Design of Context-Sensitive Profiler for Android Applications. In: 2013 IEEE 17th International Symposium on Consumer Electronics (ISCE), Zhubei City, Taiwan (2013)
40. Md Fudzee, M.F., Abawajy, J.: A Protocol for Discovering Content Adaptation Services. In: Xiang, Y., Cuzzocrea, A., Hobbs, M., Zhou, W. (eds.) ICA3PP 2011, Part II. LNCS, vol. 7017, pp. 235–244. Springer, Heidelberg (2011)
41. Md Fudzee, M.F., Abawajy, J.: Multi-criteria Content Adaptation Service Selection Broker. In: 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing (CCGrid), Melbourne, Australia (2010)
42. Mohamed, J., Md Fudzee, M.F., Abawajy, J., Kasim, S., Ismail, M.-N.: An SLA Evaluator for Multi-media Content Adaptation Services. In: 2014 International Conference on Information Science and Applications (ICISA), Seoul, South Korea (2014)

# Operational Rules for Implementing Sincere Software Agents in Corrective and Preventive Actions Environment

Nur Huda Jaafar[1], Mohd Sharifuddin Ahmad[2], and Azhana Ahmad[2]

[1] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA,
KM 12, Jalan Muar, 85009 Segamat, Johor Malaysia
nurhu378@johor.uitm.edu.my
[2] College of Information Technology, Universiti Tenaga Nasional, Putrajaya Campus,
Jalan IKRAM-UNITEN, 43000 Kajang, Selangor
{sharif,azhana}@uniten.edu.my

**Abstract.** Organizational management that manifests favourable behavioural characters among organizational members brings many benefits to the organizations and their team members. Such successful management ideas, offer software agent technology researchers a model to adopt these behaviours in software agent-based systems. In this paper, we propose operational rules for implementing sincere and sacrifice characters of agents in agent-based systems for corrective and preventive actions in quality management environments. To ensure that a sacrifice agent still produces the best services for its own increasingly progressive task, our proposed agent-based system compares the urgency and importance of its current task and an incoming task. The rules for implementing sincere and sacrifice agent help us to formulate sincere behavior in the agent-based systems.

**Keywords:** Software agent, sacrifice, sincerity, corrective and preventive actions, quality management.

## 1 Introduction

In quality management, the enforcement of corrective and preventive actions is very important to ensure that the quality management process of an organization still maintains its high quality products or services. Corrective actions are taken to fix existing problems, while preventive actions are taken to avoid the occurrence of potential problems [1]. To achieve an organization's objective, teamwork spirit is a priority in enforcing corrective and preventive actions. Strong moral values instilled in organizations such as sincerity in doing jobs may help organizations to maintain consistent quality management.

Previous research findings show that strong moral values in an organization can bring positive impacts to the organization [2][3]. We propose that sincerity is one of the strong moral values to be instilled among members as an important organizational

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_30

culture. According to [3], sacrificing culture can also tremendously help in developing cohesiveness among the different parties working in an organization including operatives, managers, and owners, which ultimately help in achieving the goals of the organization.

Software agent technology is one of the technologies that have been proven to assist humans in performing tasks. According to [4], an autonomous software agent is a component that interacts with its environment and with other agents on the user's behalf. Most software agent's characteristics are adapted from human's behavioral characteristics.

In our research, we attempt to introduce human's sincere attitude in a multi-agent environment. For this purpose, a questionnaire has been distributed to 153 respondents to identify the factors that manifest sincere attitude in human's organizations. Factor analysis has been conducted on the survey results to identify the factors that contribute to sincere attitude in human's organization. One such factor is the sacrifice attitude while completing a task. Based on this finding, we adopt the characteristic of sacrifice with sincerity in a multi-agent environment and implement it in a corrective and preventive actions environment.

In this paper, we present a novel concept of an agent which is sincerely willing to sacrifice its tasks and resources to assist other agents in performing tasks in corrective and preventive actions environment. In this concept, our proposed agent-based system considers the sacrifice agent's current tasks and incoming tasks to avoid the agent from failing to timely complete the tasks. While enforcing corrective and preventive actions, it may involve teamwork within a department and between departments. However, in this paper, we only focus on the instillation of sacrifice attitude of teamwork agents in a department.

## 2 Related Works

A conflict among staff in any organization is one of the biggest problems that management must resolve. Failure to do so could results in bad reputation to the organization. Conflicts that arise during corrective and preventive actions should be handled carefully and immediately to avoid bad practices dominating the quality management and the organization in general. One of the causes of conflicts is the existence of insincere attitude among staff. Insincere staff can be a burden to any organization because this bad attitude manifests negative effects to the organization, especially in getting cooperation among the staff. Indirectly, sabotage among the staff may happen due to personal interests. Such disturbances promote negative impacts to organizations [5].

Sincere behavior among organizational members brings positive impacts to the organization especially when corrective and preventive actions are enforced. An organization needs its team member's support at cooperation, while performing their own tasks. If members are willing to sacrifice with extra responsibility, it would ensure that the common goals of the organization can be achieved. Sincerity culture can boost up quality management and brings many benefits to staff, clients and organizations. As mentioned by [6], if we know what we wish to sustain and if it is a righteous cause, then through sincerity, we can sustain it only by replacing them with a more positive and constructive culture of nonviolence. This shows the importance of sincerity in our life especially when managing processes that involve many parties.

Negotiation in agent-based system involves many parties [7]. According to [8][9], negotiation in agent-based system can be (i) one-to-one (ii) many-to-one or (iii) many-to-many negotiation. In agent-based system, sincere behavior can help agents to negotiate successfully. Sincere behavior can be applied in cooperative negotiation, which occurs when agents have a global goal envisioned for the system [10]. As mentioned by [11], cooperation among team members is important in assisting organizations to solve problems and achieving its goals.

When applying sincere behavior in agent-based systems, the sacrifice behavior should also be considered. Sacrifice can occur sincerely by an individual or via instruction from the leader but void of sincerity. Ahmad et al. [12] has implemented the sacrifice behavior in their OP-RND framework, which focuses on how agents need to sacrifice their resources to ensure that tasks can be completed as scheduled and to avoid penalty. In our research, we instill the sacrifice behavior in agents to help their team members to alleviate their responsibilities while carrying out their own tasks.

## 3 The Scenarios of Potential Sacrifice Behavior in Corrective and Preventive Actions

The literature defines sincerity as doing things voluntarily for the sake of God without hoping for any worldly rewards [13][14]. While sacrifice is to give up something that is important or valuable to us in order to get or do something that seems more important for another person [15]. According to [16], sacrifice refers to what we give or offer up, sometimes willingly, sometimes under pressure. Here, we define a sacrifice with sincerity software agent as an agent that is willing to help other agents without hoping for any reward, although at the same time it has the task to carry out its own duty in the organization.

There are two scenarios in which our agent-based system instills potential sacrifice behavior. We only focus on the agents that exist in a localized department isolated from others. To instill potential sacrifice behavior, the system is to be based on:

1. Work Progress Log File: This log file stores the history of an agent work progress. It is important to the system environment for detecting the level of work-in-progress for each agent. The work progress log file can only be accessed by the owner and the Head of Department (HOD) agent.

2. Experience base: Experience base is used to keep all the agents' experience. From this experience base, an agent with a problem (problem agent) and the HOD agent can view all agents' experience in order to identify potential agents to help the problem agent.

In our proposal, there are four types of agent involved:

1. *HOD agent,* $A_{hod}$: $A_{hod}$ plays the role as a leader in a department or unit. It monitors the work progress of all agents under its responsibility. The $A_{hod}$ agent also determines the time when a problem agent, $A_p$, is offered a help from a team mate agent, $A_{tm}$. It can also offer the experienced $A_{tm}$ to assist $A_p$ in completing a difficult task.

2. *Problem agent,* $A_p$: $A_p$ refers to the agent that faces problems in completing its tasks.

3. *Team mate agent,* $A_{tm}$ : A team mate agent is a team member under the same department or unit. It can offer to solve a problem faced by a $A_p$. A favor from $A_{tm}$ can help the organization to achieve its goals.

4. *Sacrifice agent,* $A_s$: Here, the agent can be considered as a sacrifice agent, $A_s$, if it has its own tasks and at the same time is willing to help a problem agent.

We deploy a signal from the work progress log file in our proposed system to alert the $A_p$ that it can get help from a $A_{tm}$. The $A_{hod}$ can also use this signal to identify the agent under its supervision that needs help from a $A_{tm}$. The system displays the signal based on the following rules:

1. The percentage of task completion; $CT$is less than or at 50%.
2. The remaining duration from the previous milestone to the next milestone; $M_t$ is less than or at 40%.

For example, if an agent only completes 50% of a task and the remaining time to reach the last milestone is only 35%, then:
$CT <= 50\ AND\ M_t <= 40$ → Display signal.

### 3.1 Scenario 1

In this scenario, an agent faces a problem in completing a task (henceforth, it is known as $A_p$). The system signals $A_p$ to alert that the level of task completion has reached a level at which the problem agent must get help from a $A_{tm}$. If the $A_p$ cannot complete the task on time, it accesses the experience base to identify a $A_{tm}$ that can help. It selects a $A_{tm}$ that has experiences in doing a similar task. Then, $A_p$ requests help from the selected $A_{tm}$. If the selected $A_{tm}$ agrees to help without any conditions, while it has its own task to complete, then this $A_{tm}$ is considered as a potential $A_s$. Figure 1 shows the Scenario 1 in agent-based system of corrective and preventive actions.

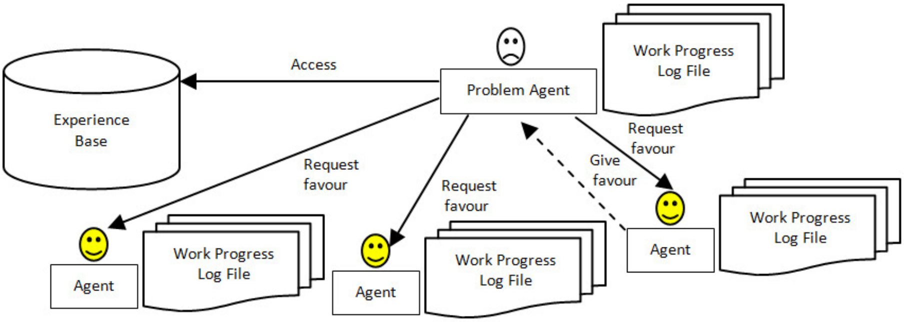


**Fig. 1.** Problem Agent Requests a Favour to the Team Mate Agent

### 3.2 Scenario 2

Figure 2 shows the Scenario 2 in agent-based system of corrective and preventive actions. The $A_{hod}$ realizes that one of the agents under its supervision is facing a problem in completing a task from the agent's work progress log file. The log file shows that the work progress reaches the level that the $A_p$ must get help from a $A_{tm}$. The $A_{hod}$ can take responsibility to appoint any suitable $A_{tm}$ that can help $A_p$. By accessing the experience base, the $A_{hod}$ may identify a suitable $A_{tm}$ by matching their experiences with the task from $A_p$. Then, the $A_{hod}$ can open the offer to a suitable $A_{tm}$. The $A_{tm}$ is considered as a potential $A_s$ if it takes the offer without any conditions, while it has its own tasks to complete.

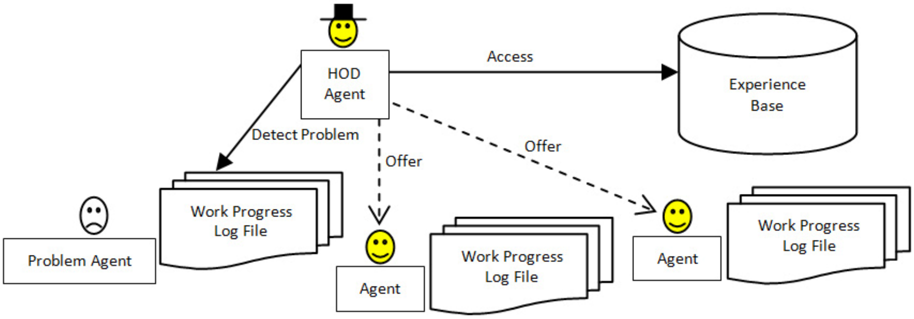


**Fig. 2.** HOD Agent Help Problem Agent to Get Help from Team Mate Agent

## 4 Comparing The Level of Urgency and Importance of Tasks

Although we can identify a potential $A_s$ from the corrective and preventive actions environment, the agent-based system imposes conditions before the potential $A_s$

takes the task from $A_p$. These conditions help to avoid the potential $A_s$ from failure to complete its on-going task.

In quality management, the use of a ranking system to evaluate issues is a good strategy to maintain quality management through the CPA. In our research, the ranking issues are discussed by humans and finalized by a human Head of Department (HOD). From the rank, the agent-based system evaluates the level of urgency and importance of the task. According to [17], urgency and importance are top items that need to be considered in setting the task according to priority. Here, we adapt the time management matrix suggested by [17] to design the urgency and importance matrix as shown in Figure 3.

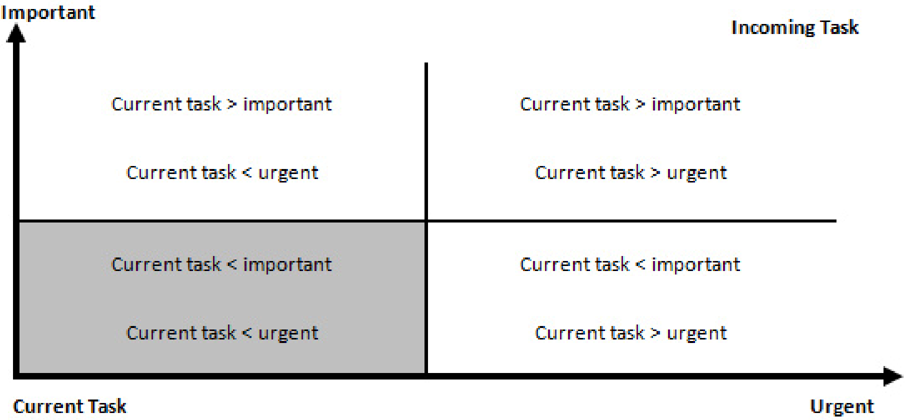


**Fig. 3.** Urgency and Importance Matrix

In our research, the level of urgency and importance for both current task; $T_c$ and incoming task; $T_i$ are compared. $T_c$ refers to the task that is still carried out by the $A_{tm}$, while $T_i$ is the task the $A_{tm}$ gets from $A_p$.

We assume that the $A_{tm}$ can only take $T_i$ if it fulfills the following conditions:

1. Level of importance, I, $T_c$ less important than $T_i$.
   $I_{T_c} < I_{T_i}$
2. Level of urgency, U, $T_c$ is not urgent compared to $T_i$.
   $U_{T_c} < U_{T_i}$

The first and second conditions are based on urgency and importance matrix as shaded in Figure 3.

3. The $T_c$ can be completed on schedule, although $T_i$ is still in progress.

The agent-based system prevents a $A_{tm}$ from taking $T_i$ if it does not fulfill the above conditions. The agent is only considered as a sincere agent if it takes the task without putting any conditions.

## 5 Conclusion and Future Work

In software agent environment, we can find many human-like behaviors adapted such as team cooperation, diligence and assistance. Many research show that such existing behaviors in the software agent environment help the agent-based systems to provide effective and efficient services. In this paper, we demonstrate the operational rules in implementing the concept of sincere agent which is willing to sacrifice its resources in agent-based system for corrective and preventive actions. The rules are summarized as follows: (1) Acquire problem signal from a log file, (2) Evaluate the urgency and importance of current tasks and incoming tasks and (3) Infer sincerity of agent when agent accepts incoming task without any condition.

While we need sincere agents in the system for helping the organization in achieving the goals, our agent-based system checks the capability of agents to complete its tasks and also the incoming task within the time frame. This is to maintain the high quality of services. In our future work, we shall formulate sincere behavior in software agents. We shall deploy a merit point system as a reward to sincere agent which is willing to help its team although having its own task to complete.

## References

1. Motschman, T.L., Moore, S.B.: Corrective and preventive action. Transfusion Science 21(2), 163–178 (1999), `http://www.ncbi.nlm.nih.gov/pubmed/10747525` (retrieved)
2. Upadhyay, A.Y.A., Upadhyay, A.K., Palo, S.: Strategy Implementation Using Balanced Scorecard: Achieving Success through Personal Values of Leaders and Employees. Management and Labour Studies 38(4), 447–469 (2014), doi:10.1177/0258042X13516596
3. Hoque, N., Khan, M.A., Mowla, M.M.: Organisational culture: features and framework from Islamic perspective. Humanomics 29(3), 202–219 (2013), doi:10.1108/H-06-2013-0040
4. Lesser, V.R.: Cooperative Multiagent Systems: A Personal View of the State of the Art 11(1), 133–142 (1999)
5. Lee, J.(J.), Ok, C.M.: Understanding hotel employees' service sabotage: Emotional labor perspective based on conservation of resources theory. International Journal of Hospitality Management 36, 176–187 (2014), doi:10.1016/j.ijhm.2013.08.014
6. Gandhi, A.: The Sincerity of Purpose: Sustainability and World Peace. In: Madhavan, G., Oakley, B., Green, D., Koon, D., Low, P. (eds.) Practicing Sustainability, pp. 45–49. Springer, New York (2013), doi:10.1007/978-1-4614-4349-0
7. Sánchez-anguix, V., Julian, V., Botti, V., García-fornes, A.: Analyzing Intra-Team Strategies for Agent-Based Negotiation Teams. In: Proceeding of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011), pp. 929–936 (2011)

8. Jaafar, N.H., Ahmad, M.S.: An Agent-based Negotiation Model Selector for Managing Negotiation Processes in Corrective and Preventive Actions: A Conceptual Framework. International Journal of Management & Information Technology 5(2), 478–482 (2013)
9. Wooldridge, M.: An Introduction to Multiagent Systems, 2nd edn. John Wiley & Sons Ltd., United Kingdom (2009)
10. Green, S., Hurst, L., Nangle, B., Cunninghan, P., Somers, F., Evans, R.: Software Agents: A review (1997)
11. Robinson, W.N., Volkov, V.: Supporting the Negotiation Life Cycle. Communication of the ACM 41(5), 95–102 (1998)
12. Ahmad, A., Yusof, M.Z.M., Ahmad, M.S., Ahmed, M., Mustapha, A.: Resolving conflicts between personal and normative goals in normative agent systems. In: 2011 7th International Conference on Information Technology in Asia, pp. 1–6. IEEE (2011), doi:10.1109/CITA.2011.5999538
13. Jamaluddin, A.-S.M., Ad-Dimasqi, A.-Q.: Mutiara Ihya' Ulumuddin: Hak Milik Muslim - Imam Al-Ghazali, 5th edn., pp. 1177–1202. Shah Alam: Illusion Network (2013); Sharul, M., Sharul, B., Harun, A. (eds.)
14. Al-Mishri, M.: Ensiklopedia Akhlak Muhammad (SAW), 2nd edn. Pena Pundi Aksara, Jakarta (2011); Mufti, A.M., Fahrie, A., Nurchasanah, Nurmalasari (eds.)
15. Turnbull, J., Lea, D., Parkinson, D., Philips, P., Francis, B., Webb, S., Ashby, M. (eds.): Oxford Advanced Learner's Dictionary, 8th edn. Oxford University Press, United Kingdom (2010)
16. Esquith, S.L.: Sacrifice, Dignity and Suicide. Political Theory 41(2), 336–342 (2014), doi:10.1177/0090591712470629
17. Covey, S.R., Merrill, A.R., Merrill, R.R.: First Things First. Simon & Schuster UK Ltd., London (1994)

# A Survey on Distributed Service Discovery Mechanisms with the Focus on Topology Awareness

Mohd Saleem Nazmudeen[1] and Seyed Mohamed Buhari[2]

[1] Institut Teknologi Brunei, Brunei Darussalam
mohamed.saleem@itb.edu.bn
[2] King Abdul Aziz University, Kingdom of Saudi Arabia
mesbukary@kau.edu.sa

**Abstract.** Distributed Service (or resource) Discovery (DSD) is becoming an important research area in Service Oriented Computing (SOC) because many software applications are now developed with services from different vendors. The query routing mechanism of the current DSD applications functions purely on the overlay without incorporating the topological and routing knowledge of the underlying physical topology of the network. Consequently, Internet Service Providers (ISPs) are tested to their limits due to underlay-ignorant query forwarding that are employed by the overlay applications such as Peer-to-Peer and DSD. This paper surveys the existing query routing approaches in various domains of Distributed Service Discovery and summarizes their level of awareness with respect to the underlying network topology. We have identified various characteristics required for the query routing algorithms to be intelligent and the comparisons are performed based on those characteristics.

**Keywords:** Service discovery, topology awareness, intelligent query routing.

## 1 Introduction

The query routing mechanism of the current DSD applications functions purely on the overlay without utilizing the topological and routing knowledge of the underlying physical network topology. Consequently, Internet service providers (ISPs) are tested to their limits due to underlay-ignorant query forwarding that is employed by the overlay applications such as Peer-to-Peer (P2P) and DSD. This paper surveys the existing query routing mechanism in DSD and their implementation approaches with respect to underlying topological awareness. In order to critically analyze various existing literatures in the area of DSD, a thorough comparison of those works with respect to the parameters that are relevant to this survey need to be performed. The main characteristics that a query routing algorithm which supports underlay based query routing are, Service Discovery (SD) architecture that can perform informed searches [1], intelligent query processing SD system [1], underlay location awareness [2, 3], reduction of inter-ISP traffic [4] and avoiding redundant traffic in the underlay [5, 6].

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_31

Taking the above mentioned criteria into consideration; the following parameters were identified and used in the following sections for comparison of existing literatures,

1. System Architecture
2. Query routing approach
3. Underlay and location awareness
4. Range queries
5. Intelligent query processing
6. ISP consideration
7. Layer in which query routing is performed
8. Traffic redundancy in the underlay

The rest of the paper is organized as follows, section 2 discusses the service discovery in the P2P domain, sections 3 and 4 discuss that of the grid and web service domains respectively. Section 5 discusses the existing query routing algorithms that provide location awareness to some extent and section 6 concludes with the existing research opportunities in location aware query routing process.

## 2 Service Discovery in P2P Systems

Gnuetella [7, 8, 9, 10-13, 3, 4, 14, 15, 16, 17, 18, 19-23] is one of the early P2P file sharing applications which is based on an unstructured approach. Each peer in the Gnutella system maintains a connection to at least one other peer. Neighbours are formed as they come to know about their existence. The query for file discovery is flooded into the network. A peer sends the query to all its neighbours and the neighbours in turn forward it to their neighbours and the process is triggered. Any query hit is propagated back to the query initiating peer. The flooding approach is fine with a small number of nodes and users. However, when the number of users using the system proliferates the performance deteriorates. As a consequence many variations of the protocol have been introduced following the Gnutella protocol. All of those routing protocols function purely on the overlay.

Kalogeraki et al. [24] has provided an implementation of the extended Gnutella protocol with two improvements. The problem area they improved is actually the searching and routing mechanism among the nodes of Gnutella. Those improvements are modified Breadth First Search (BFS) and intelligent search mechanism based on past behaviour. Unlike the original Gnutella protocol, the modified Breadth First Search chooses only a subset of the neighbours. Intelligent searching is performed by monitoring the past queries so that only those neighbours which are likely to reply to the queries are used for forwarding. The authors' objective was to minimize the search cost and to improve the search efficiency of the system. However, it still remains an unstructured overlay without any underlying topological consideration.

Freenet [25] is another unstructured P2P system which employs a better query routing process compared to Gnutella. It makes use of key-based query processing. The neighbour for query forwarding is selected if the key has been answered by the neighbour earlier. Here the links between the neighbours are dynamically altered by making new links and dropping old links based on the responsiveness of the peers. The overlay layer does not possess any knowledge about the underlying topology.

Chord [10] is the first structured system which employs the Dynamic Hash Table (DHT) as a way of implementing a distributed look up service in the Internet. It employs a key mapping mechanism in which a given keyword of a particular resource is mapped into a node with the help of a hash function. If there is no exact node identifier for the given key, then the node next in the sequence will be used for storing the data. This node is called the successor node. The complexity of the resource look up in Chord is O(log N). The main drawback of the Chord system is its tight administrative overhead and difficulty in providing support for range and partial queries.

The Content Addressable Network [22] is another popular structured P2P system which makes use of tree structured nodes and a virtual Cartesian coordinate space with a d-dimension. A node learns the set of IP addresses of the nodes that falls as neighbours in the coordinate space and not in the physical topology. In this way there is no information about the physical location information of the underlying network. Also, the routing complexity in CAN is O(dN1/2) where d is the dimension of the Cartesian coordinate space. Therefore, when d is small, the complexity is worse than Chord which is a DHT based implementation.

An improvement to CAN was proposed by [26], which tried to modify the way the routing was performed in the coordinate space of CAN. It was argued that in CAN, routing was inefficient especially for small values of dimension for the coordinate space because a node in the coordinate space maintained only the set of IP addresses of the neighbours which was constant irrespective of the number of peers. Therefore, some 'Long Distance Pointers (LDP)' were added which were IP addresses of some distant peers. This could eventually reduce the routing complexity. The underlying topological consideration and the traffic redundancy control were not in their scope of work.

FastTrack [27] is a proprietary protocol that uses encryption for message exchanges. Until today the operation of FastTrack still remains a mystery. It has been found through packet sniffing that it operates in a hybrid structure with super nodes. How the topology is constructed and whether there is any underlay awareness in the overlay layer is not known.

HyperCuP (Hypercube P2P) [28] is an attempt to reduce the flooding mechanism usually employed in the unstructured P2P systems. It also introduces a partial structure to the unstructured systems by building a graph for each formation of neighbour. The main objective of the HyperCuP is to form a TCP overlay which is used for the formation of graph structure on the nodes, so that the network diameter for finding a resource in the worst case can be reduced as compared to that of unstructured systems. Here again, it is a pure overlay construction and the underlay awareness is out of their scope.

Kademlia [12] is a structured P2P system which is based on DHT similar to that of Chord [11]. However, Kademlia differs from Chord in the way the configuration messages are learned and in the topology construction. Kademlia uses an XOR metric based tree topology for the routing of query messages. In pure DHT based systems the path which has a higher latency has the same weight as the one with a lower latency so that the routing process cannot differentiate between them. In Kademlia, these configuration messages are automatically spread to the nodes so that nodes can route queries with enough information.

Pastry [29] is a hybrid architecture which can perform either tree based routing like Kademlia or ring based routing like Chord. It resorts to ring based routing if the tree routing fails. For this purpose each node maintains a set of leaf, routing and neighbour tables. The neighbour table is used to perform routing based on proximity. However, redundancy and ISP issues are not taken into account.

The characteristics of the reviewed P2P systems are summarised in Table 1.

**Table 1.** Service discovery in P2P applications

| | *Architecture* | *Query routing mechanism* | *Underlay location awareness* | *Partial, Range, and Multi attribute queries* | *Intelligent Query processing* |
|---|---|---|---|---|---|
| Gnutella [13] | Unstructured | Flooding | No | Yes | No |
| Kalogeraki et al. [24] | Unstructured | Controlled flooding | No | Yes | Yes(Intelligent search mechanism based on past behaviour) |
| Freenet [25] | Unstructured | Controlled flooding | No | No | No |
| Chord [10] | Structured | DHT | No | No | No |
| CAN [22] | Structured | Selective neighbour selection based on Cartesian coordinates | No | No | No |
| Sahin et al. [26] | Structured | Selective neighbour selection based on Cartesian coordinates Adds 'Long Distance Pointers (LDP)' to the set of neighbours' list | No | No | No |
| FastTrack (Proprietary) [27] | Hybrid | N/A | N/A | - | N/A |
| HyperCuP [28] | Hybrid | Controlled flooding | No | No | No |
| Kademlia [12] | Structured | DHT | No | No | Partial |

## 3 Service Discovery in Web Services

Fatih et al. [30] contributed in providing a distributed service discovery so that trust and Quality of Service (QoS) are taken into account during service discovery. They incorporated this feature by providing a ranking system for similar services based on trust and QoS. Therefore, if there is an option in choosing more than one service in the discovery process, the one with the highest ranking will be chosen. A sample scenario provided by the authors is that of a user who wishes to purchase a book online with the most secured payment service and the fastest possible delivery. These choices can be automated in the SD process of their system. The system here is based on the Chord based structured P2P system.

Liu et al. [31] provided a two layer P2P model for distributed service discovery based on semantics. Their approach in organizing the peer structure is based on ontology. They argued that ontology based SD is more time consuming because of the logical reasoning, and therefore needs to be distributed. The structure here is similar to the super peer architectures. The super peer is the 'Ontology Agent' under which a particular 'Ontology Community' is organized. The 'Ontology Agent' forwards the query to other 'Ontology Agent' if the query requests services from different 'Ontology Communities'. It also maps the ontology from one form to another form as per the requirement. Here, they employ a hybrid approach similar to cluster formation and functions by forming an overlay.

Lin Zhang [32] provided another two layer approach which has the cluster based super peer structure. However, the main difference from this work and [32] is that the centralized discovery is performed within the cluster and the P2P model of discovery is performed only across the clusters. Here, the clusters are formed based on semantic similarity. Ling Zing's approach uses the overlay hybrid model and the topology consideration is out of his scope. The distributed structure of Lin Zhang's implementation is shown in Figure 2.6. The black nodes are the super nodes and the white nodes are the normal nodes grouped according to their semantic similarity. The nodes that need to be searched for service discovery are reduced due to the number of super nodes in the system.

Sioutas et al. [33] provided a remarkable contribution in distributed web service discovery. The main advantage of the DHT based system is that the lookup takes O(log N) complexity where N is the number of nodes. Sioutas et al. have gone a step further and have come up with a structured system which can provide the complexity of O(log log N). Their system is called 'NIPPERS (Network of InterPolated PeERS)'. This system is in fact very suitable for highly static systems. However, the tight administrative overhead due to the high structured nature makes it not suitable for situations where dynamic partnership among organizations (frequent exit from and entry into the P2P system) is of higher priority.

Scoutas et al. [34] has given an approach where semantic web service discovery can be applied in both the centralized and P2P environments. In order to improve the response time they adopt an approach called progressive query matching where better matches are returned first while the query processing proceeds further. They make use of the structured P2P organization which employs spatial coordinates.

Chord4S [11] is a system which attempts to improve the original Chord [10] so that the services are distributed evenly across nodes rather than concentrating on a particular node. The services have the tendency to get hashed to the same node if they belong to the same category as they are usually mapped to closer hash values. Chord4S has avoided this by adding some provider bits to the service bits for the purpose of hashing so that they end up with different hash values and eventually different nodes. Figure 2.7 shows an example of how the provider bits are added to functional bits. In this example the bits corresponding to service provider IP address '10.0.0.1' is added with the functional bits of the service provided.

**Table 2.** Service discovery in the domain of web services

| | *P2P Approach* | *Problems Addressed* | *Underlay awareness* | *ISP Considerations* |
|---|---|---|---|---|
| Fatih et al. [30] | Chord-based, structured | Addresses QoS and trust attributes for service discovery | No | No |
| Liu et al. [31] | Super peer cluster (Hybrid) | Performance improvement in ontology processing | No | No |
| Lin Zing [32] | Super peer cluster (Hybrid) Centralized within the cluster and distributed across | Semantic classification | No | No |
| Sioutas et al. [33] | Tree structure with interpolated peers | Routing efficiency | No | No |
| Scoutas et al. [34] | Spatial coordinates, structured | Progressive query matching for improved response time | No | No |
| Chord4S [11] | Chord | Load balancing | No | No |
| JaxSON [35] | Hybrid, Clusters based on semantic affinity | Semantic based service discovery | No | No |
| Zhou et al. [36] | Tree-based, structured | Service discovery in cloud | No | No |
| Schmidt et al. [37] | Structured | Range queries | No | No |

JaxSON [35] is a semantic overlay which is hybrid in nature with cluster formation based on 'Semantic Affinity (SA)'. According to this structure the cluster which can answer a particular query can be identified by the system based on SA. The SA function is used to identify the group in which a new node can join. The SA would be compared with the existing groups and the entering peer. The group which has the highest value of SA with the entering peer would add the peer into its group. All the processes that are taking place here are completely on the overlay of the system and have nothing to do with the underlay. This approach is similar to Lin Zang [32], however JaxSON uses SA for identifying the cluster to which a node needs to be joined. In JaxSON, it is possible that a node can join more than one group with slightly different SA. During the query processing SA will determine to which group a node belongs.Zhou et al. [36] provided an interesting P2P model for service discovery in the cloud environment. They provide a Kd-tree approach for implementing range queries with a semantic and locality enabled service discovery process. The multidimensional query is routed based on the Kd-tree of the peers. However, it still functions in the overlay layer. A sample Kd-tree is shown in Figure 2.9 in which the point (200,50), which represents 200GB and 50Mb/s, splits the data space into two subspaces and accordingly other nodes and points are formed. If a node has branches the rectangle is split into two subspaces otherwise the point is just marked without splitting into further subspaces. This approach helps in localizing the region for searching range queries.

Schmidt et al. [37] provides another distributed web service discovery model which uses Chord as the P2P structure and Hilbert's space filling curve to map the multi-dimensional queries. It is a structured overlay system. Table 2 summarises the contributions of the above works.

## 4 P2P and SD Approaches with Location Awareness

PIPPON [15] attempts to bring the underlay awareness to the overlay by clustering the peers based on the proximity. The proximity of the peers is measured with two parameters namely, Longest prefix IP matching (LPM) and Round Trip Time (RTT).

Apart from cluster formation, it forms a dynamic "Key Tree" for routing the query messages on the overlay. This tree is employed to route the queries based on the IP prefixes as shown in Figure 2.10. However, PIPPON does not take into account the similarity of queries in the service discovery which is a key factor in improving efficiency. Also, it does not solve the problem of redundant traffic in the underlay.

TOPLUS [38] has adopted a slightly different approach to PIPPON in terms of underlay awareness. The cluster formation in TOPLUS involves three tier hierarchies such as groups, super-groups and hyper-groups so that the queries can be easily forwarded with a minimal number of hops. In contrast to the longest prefix match of the IP addresses, TOPLUS makes use of the XOR metric for identifying the proximity. In each tier the routing of the query messages is performed with the XOR metric. TOPLUS claims that the routing performance it provides is very close to IP routing. TOPLUS does not cover the aspects of inter-ISP, redundant traffic and interlayer communication overhead.

In general there are two approaches with respect to finding the locality information of the underlay:

- Dynamic (on the fly calculation with RTT and LPM), and
- Static (Prior knowledge of the Internet such as AS numbers).

Plethora [16] follows the second approach which is static whereas TOPLUS and PIPPON follow the first approach. Plethora makes use of cSpace, which uses a local broker, and gSpace which uses a global broker. Hence, it ends up with two layer architecture. If a request is generated, it checks the cSpace for answers; if cSpace could answer it, then the query is not forwarded to gSpace. It only forwards it to the gSpace if the query is not answered by the cSpace.

P4p [39] addresses the locality awareness problem by two interfaces iTrackers and appTrackers. The purpose of iTrackers is to provide the locality parameters of a particular node to the querying node. iTrackers are designed to reside in each of the ISPs. appTrackers are designed to reside globally so that they have a complete picture of the P2P applications. appTrackers would decide whether a particular peer "a" could establish a neighbour relation with another peer "b" based on the information from iTrackers. A sample scenario of the tracker interfaces are shown in Figure 2.11.

**Table 3.** SD approaches with location awareness

| | *Locality awareness approach* | *Targeted performance criteria* | *Inter-ISP traffic reduction* | *Security* | *Query routing layer* |
|---|---|---|---|---|---|
| PIPPON [15] | Pure overlay based on RTT and LPM | Improved locality awareness | No | No ISP related security issues as it is overlay-based | Overlay |
| TOPLUS [38] | Pure overlay based construction with IP prefix-based location awareness | Improved locality awareness | No | No ISP related security issues as it is overlay-based | Overlay |
| Plethora [16] | Overlay-based, uses static ISP information | Improved locality awareness | No | No ISP related security issues as it is overlay-based | Overlay |
| P4P [39] | ISP Assisted approach | Traffic control with P2P application and ISP cooperation | Yes | Security concerns relating to exposing ISP information to the peers in the overlay are not addressed | Overlay |
| Seedorf et al. [2] | ISP Assisted approach | Traffic control with P2P application and ISP cooperation | Yes | Security concerns relating to exposing ISP information to the peers in the overlay are not addressed | Overlay |
| SLUP [3] | RTT based proximity identification without ISP involvement | Traffic control with locality based clustering | No | No ISP related security issues as it is overlay-based | Overlay |
| Bindal et al. [40] | Purely ISP - based. Specific for BitTorrent applications | Traffic control for BitTorrent | Yes | No specific security issues as the ISP information is not exposed to peers in the overlay. | Overlay |
| P4P Pastry [41] | ISP Assisted approach | Traffic control with P2P application and ISP cooperation | Yes | Security concerns relating to exposing ISP information to the peers in the overlay are not addressed | Overlay |
| Haja M. Saleem et al. [42] | AON routing | complete locality awareness | Yes | Security concerns with respect to AON route corruption | Underlay |

IETF formed an Application Layer Traffic Optimization (ALTO) group in 2009 and has been tasked with identifying the means of optimizing the traffic in the underlay that are generated by the ever-increasing P2P applications. Seedorf et al. [2] elaborate on the efforts of the ALTO group which intends to provide the ALTO service to the P2P applications. With reference to the ALTO approach, any P2P application which needs to know the topological and proximity information could obtain that information from the ALTO server. Information like location of the target peer, operational cost and policies from the ISPs could be provided to the querying peers. A sample scenario is shown in Figure 2.12. In this approach, letting the peers be aware of these network related parameters could be misused by any malicious user. As an alternative the proposed approach in this research for location awareness could be used for the same goal as that of IETF in optimizing traffic in the underlay.

SLUP [3] is another approach which forms clusters based on semantics and RTT. It forms a three-tier clustering: normal peer, level-2 super peer and level-1 super peer. Redundancy in the underlay and location awareness are not considered in this work.

Bindal et al. [40] provided a strategy for improving the traffic locality in the BitTorrent network. They highlighted the fact that ISPs often end up throttling the P2P traffic as BitTorrent like unstructured P2P systems do not take ISPs cost into consideration. Their work focuses on biased neighbour selection in which, a peer chooses the majority of its neighbour from the same ISP so as to reduce the cross-ISP traffic. The authors point out here that if there are N users within the ISP wanting to download a particular file, then in the current BitTorrent system the file is downloaded N times, which heavily increases inter-ISP traffic. In their work they have demonstrated that biased neighbour selection could improve this situation. Here, their application is specific to the BitTorrent system which uses the rarest first piece algorithm at the client side.

P4P Pastry [41] is an approach to bring the locality features to the PASTRY [29] structured system. Its features are similar to that of P4p [38] except that it is applied in the context of PASTRY [29]. It has the same limitations as that of P4p [38].

Haja M. Saleem et al. [42] has proposed a radical approach in terms of location awareness as the routes to the queries are dynamically learned and stored for further query routing. Here they have proposed to perform the query routing process in the underlying network layer with the help of Application Oriented Networking (AON). Table 3 summarises the contributions of the above works and highlights the need for underlay awareness.

## 5 Conclusion and Future Work

This paper provided a comprehensive review of SD approaches in the field of P2P systems, web services and current topological awareness techniques. It can be concluded from the literature survey that there is a need for a mechanism which aids the process of DSD with intelligent message routing so that location awareness can be incorporated in the overlay of the distributed systems. There are still research needed to be done in the following areas such as, a study on how ISPs could be encouraged to provide location awareness for query routing, the performance of the DSD systems while clustering the peers purely based on service classes needs to be investigated further and the effect of clustering the registries across ISPs also demands a careful study.

## References

1. Meshkova, E., et al.: A survey on resource discovery mechanisms, peer-to-peer and service discovery frameworks. Computer Networks 52, 2097–2128 (2008)
2. Seedorf, J., et al.: Traffic localization for P2P-applications: The ALTO approach. In: IEEE Ninth International Conference on Peer-to-Peer Computing, P2P 2009, pp. 171–177 (2009)
3. Xin, S., et al.: SLUP: A Semantic-Based and Location-Aware Unstructured P2P Network. In: 10th IEEE International Conference on High Performance Computing and Communications, HPCC 2008, pp. 288–295 (2008)
4. Karagiannis, T., et al.: Should internet service providers fear peer-assisted content distribution? In: Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement, p. 6. USENIX Association, Berkeley (2005)
5. Abboud, O., et al.: Underlay awareness in P2P systems: Techniques and challenges. In: IEEE International Symposium on Parallel & Distributed Processing, IPDPS 2009, pp. 1–8 (2009)
6. Jie, D., et al.: The disparity between P2P overlays and ISP underlays: issues, existing solutions, and challenges. IEEE Network 24, 36–41 (2010)
7. Cai, M., et al.: MAAN: A Multi-Attribute Addressable Network for Grid Information Services. Journal of Grid Computing 2, 3–14 (2004)
8. Talia, D., Trunfio, P., Zeng, J.: Peer-to-Peer Models for Resource Discovery in Large-Scale Grids: A Scalable Architecture. In: Daydé, M., Palma, J.M.L.M., Coutinho, Á.L.G.A., Pacitti, E., Lopes, J.C. (eds.) VECPAR 2006. LNCS, vol. 4395, pp. 66–78. Springer, Heidelberg (2007)
9. CISCO, Cisco AON: A Network Embedded Intelligent Message Routing System, `http://www.cisco.com/en/US/prod/collateral/modules/ps6438/prod_bulletin0900aecd802c201b.html`
10. Ion, S., et al.: Chord: A scalable peer-to-peer lookup service for internet applications. SIGCOMM Comput. Commun. 31, 149–160 (2001)
11. Qiang, H., et al.: Chord4S: A P2P-based Decentralised Service Discovery Approach. In: IEEE International Conference on Services Computing, pp. 221–228 (2008)
12. Maymounkov, P., Mazières, D.: Kademlia: A peer-to-peer information system based on the XOR metric. In: Druschel, P., Kaashoek, M.F., Rowstron, A. (eds.) IPTPS 2002. LNCS, vol. 2429, pp. 53–65. Springer, Heidelberg (2002)
13. Gnutella-Protocol, `http://rfc-gnutella.sourceforge.net/src/rfc-0_6-draft.html` (accessed March 7, 2012)
14. Hung-Chang, H.: A Near-Optimal Algorithm Attacking the Topology Mismatch Problem in Unstructured Peer-to-Peer Networks. IEEE Transactions on Parallel and Distributed Systems 21, 983–997 (2010)
15. Hoang, D.B., Le, H., Simmonds, A.: PIPPON: A Physical Infrastructure-aware Peer-to-Peer Overlay Network. Presented at TENCON 2005 2005 IEEE Region 10 (2005)
16. Ferreira, R.A., Grama, A., Jia, L.: Plethora: An Efficient Wide-Area Storage System. In: Bougé, L., Prasanna, V.K. (eds.) HiPC 2004. LNCS, vol. 3296, pp. 252–261. Springer, Heidelberg (2004)

17. Wei-peng Chen, Y.G., et al.: J-Sim 1.3, `http://sites.google.com/site/jsimofficial/`
18. Saroiu, S., Gummadi, K.P., Gribble, S.D.: Measuring and analyzing the characteristics of Napster and Gnutella hosts. Multimedia Systems 9, 170–184 (2003)
19. Guttman, E., et al. (eds.): Service Location Protocol, Version 2. RFC 2165 (1997), `http://www.ietf.org/rfc/rfc2608.txt`
20. Arnold, K., et al. (eds.): Jini Specification, 1st edn. Addison Wesley Longman Publishing Co., Inc, Boston (1999)
21. Leibowitz, N., et al.: Deconstructing the Kazaa network. In: Proceedings of the Third IEEE Workshop on Internet Applications, WIAPP 2003, pp. 112–120 (2003)
22. Ratnasamy, S., et al.: A scalable content-addressable network. SIGCOMM Comput. Commun. Rev. 31, 161–172 (2001)
23. Ripeanu, M.: Peer-to-peer architecture case study: Gnutella network. In: Proceedings of the First International Conference on Peer-to-Peer Computing, pp. 99–100 (2001)
24. Vana, K., et al.: A local search mechanism for peer-to-peer networks. In: Proceedings of the Eleventh International Conference on Information and Knowledge Management, pp. 300–307. ACM, McLean (2002)
25. Freenet,, Available: freenetproject.org/
26. Sahin, O.D., et al.: Techniques for efficient routing and load balancing in content-addressable networks. In: Fifth IEEE International Conference on Peer-to-Peer Computing, pp. 67–74 (2005)
27. Liang, J., et al.: The FastTrack overlay: A measurement study. Computer Networks 50, 842–858 (2006)
28. Schlosser, M.T., Sintek, M., Decker, S., Nejdl, W.: HyperCuP - Hypercubes, Ontologies, and Efficient Search on Peer-to-Peer Networks. In: Moro, G., Koubarakis, M. (eds.) AP2PC 2002. LNCS (LNAI), vol. 2530, pp. 112–124. Springer, Heidelberg (2003)
29. Rowstron, A., Druschel, P.: Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems. In: Guerraoui, R. (ed.) Middleware 2001. LNCS, vol. 2218, pp. 329–350. Springer, Heidelberg (2001)
30. Fatih, E., et al.: A peer-to-peer framework for Web service discovery with ranking. In: Proceedings of the IEEE International Conference on Web Services, pp. 192–199 (2004)
31. Zhizhong, L., et al.: A two-layered P2P model for semantic service discovery. In: 2010 4th International Conference on New Trends in Information Science and Service Science (NISS), pp. 41–46 (2010)
32. Lin, Z.: A Scalable Model for Service Discovery in P2P Environment. In: 2011 International Conference on Intelligence Science and Information Engineering (ISIE), pp. 495–498 (2011)
33. Sioutas, S., et al.: Dynamic Web Service discovery architecture based on a novel peer based overlay network. Journal of Systems and Software 82, 809–824 (2009)
34. Skoutas, D.N., Sacharidis, D., Kantere, V., Sellis, T.K.: Efficient Semantic Web Service Discovery in Centralized and P2P Environments. In: Sheth, A.P., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T., Thirunarayan, K. (eds.) ISWC 2008. LNCS, vol. 5318, pp. 583–598. Springer, Heidelberg (2008)
35. Bisignano, M., Di Modica, G., Tomarchio, O.: JaxSON: A Semantic P2P Overlay Network for Web Service Discovery. Presented at 2009 World Conference on Services - I (2009)
36. Zhou, J., Abdullah, N.A., Shi, Z.: A Hybrid P2P Approach to Service Discovery in the Cloud. IJITCS 3 (2011)
37. Schmidt, C., Parashar, M.: A Peer-to-Peer Approach to Web Service Discovery. World Wide Web 7, 211–229 (2004)

38. Garcés-Erice, L., Ross, K.W., Biersack, E.W., Felber, P., Urvoy-Keller, G.: Topology-Centric Look-Up Service. In: Stiller, B., Carle, G., Karsten, M., Reichl, P. (eds.) NGC 2003 and ICQT 2003. LNCS, vol. 2816, pp. 58–69. Springer, Heidelberg (2003)
39. Xie, H., et al.: P4p: provider portal for applications. Presented at the Proceedings of the ACM SIGCOMM 2008 Conference on Data Communication, Seattle, WA, USA (2008)
40. Bindal, R., et al.: Improving Traffic Locality in BitTorrent via Biased Neighbor Selection. In: 26th IEEE International Conference on Distributed Computing Systems, ICDCS 2006, pp. 66–66 (2006)
41. Zhengwei, G., et al.: P4P Pastry: A novel P4P-based Pastry routing algorithm in peer to peer network. In: 2010 The 2nd IEEE International Conference on Information Management and Engineering (ICIME), pp. 209–213 (2010)
42. Saleem, H.M., Hassan, M.F., Asirvadam, V.S.: Proxy-Based Selective Forwarding in Distributed Service Discovery Using Application Oriented Networking. Advances in Information Sciences and Service Sciences 4(8) (2012)

# Ubiquitous Shift with Information Centric Network Caching Using Fog Computing

Ibrahim Abdullahi[1,2], Suki Arif[1], and Suhaidi Hassan[1]

[1] InterNetWorks Research Laboratory, School of Computing,
Universiti Utara Malaysia, 06010 UUM Sintok, Malaysia
`ibrahim@internetworks.my`, {`suki1207`,`suhaidi`}`@uum.edu.my`
[2] Department of Mathematics/Computer Science,
Ibrahim Badamasi Babangida University, Lapai, Niger State, Nigeria
`ibrojay01@ibbu.edu.ng`

**Abstract.** With the immense growth of information gathering, sharing and processing in the world through sophisticated smart gadgets, it is therefore paramount that the Cloud would be better managed through other paradigms such as Fog computing and Information Centric Network (ICN). These would drive the need for the shift to the Internet of Everything (IoE). Jacobson et al., Cisco and Gantz et al., in different studies forecasted the engrossment of IP addresses and as such these along other vital reports have necessitated the need of replacing IP addresses with names through the Content Centric Networking. Ciscos' report submitted the forecast of smart gadgets out numbering the world population soon. This paper therefore presents a conceptual framework of introducing Information Centric Networking (ICN) as API to Ubiquitous computing. Taking ICN cache at the edge nodes through Fog computing technique (as off-network cache) by referring object with names in lieu of IP addresses. This practice will make access to information residing on the Cloud in the Internet of Everything closer to the user. Fog in this paper is used as edge processing node in the content store to depict the caching that makes the ICN achievable off-path. Fog concept proposed the computing, processing and storage at the edges (devices) with the added advantage of heterogeneity for Internet of Things (IoT).

**Keywords:** BInformation centric network, ubiquitous, fog computing, cloud computing, off-path caching, Internet of everything.

## 1 Introduction

The need for ubiquitous computing makes information dissemination easier by connecting all Internet enabled gadgets closely useful to the user. Yet, making the ubiquitous computing challenging by assigning IP addresses to all connected devices [3,10,12]. The consumption of IP addresses by this paradigm could as well come to play prior to the earlier years forecasted by Ciscos' reports. The reports states that extra ordinary rise in IP usage has hit 1.5 exabytes per month at the end of 2013 from 820 petabytes per month in 2012 [5]. Gantz et al. in

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems*,
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_32

[6] additionally submitted that huge amount of information shared by IP hit up increasingly from 2008 to 2010 and through the years. As such, cloud computing introducing fog [4] as the Information Centric Networking (ICN) cache side (off-network) would make information dissemination faster with a lower latency, less excessive bandwidth consumption and reduce the streaming times.

However, challenges posed at the ubiquitous will become clearer to identify when the user connects barely the entirety of his household items (such as electronic refrigerators for monitoring, electric doors, light bulbs, toilet handlers, sprinklers, game consoles, smart TV, etc.) to the Internet or cloud to receive updates possibly through a wearable device. More so, for countries like China and India with large population sizes, if the Man-to-all equipment is practiced and overall items are assigned IP addresses, then the exhaustion of IP addresses will be more expended and delays will therefore increase. These facts support Jacobson et al. idea of naming objects as content centric on the network [8]. Referring object with names in place of IP will therefore drive the ubiquitous to the new ICN paradigm. These along other advantages will reduce the consumption of IP addresses and benefit from the user-in-the middle phases of the Content Centric Networking (CCN) [8].

For the purpose of this paper, an idea on how to benefit immensely from the connectivity and disappearance of physical gadgets inter connectivity through ubiquitous computing [18] is presented. A strong motivation of vending into the Internet of names ahead of the IP based could be traced to the recent release of survey and forecast by Cisco which states that in 2013, mobile traffic alone hits nearly 18 times the size of the entire global Internet in the year 2000. With the submission of this magnitude, information processing practice is drastically shifting into mobile devices. Consequently, in the report by Cisco in [5], it was understood that global mobile devices connected to the Internet between a short span of 2012 to 2014 has increased from 6.5 billion in 2012 to 7 billion with lots of the connectivity attributed to the smart gadgets.

Internet of Things (IoT) has become the most comfortable idea to users and barely every user is fully addicted to making updates and information outgrowth online. If the world population is about 7 to 8 billion, before 2018 user enabled gadgets will outnumber the world population to about 10 billion including Machine-2-Machine (M2M) communication on the Internet [5]. An issue of this strength requires borrowing idea of ICN to manage the gross increase of information sharing. Moon et al., in their paper stated that ubiquitous computing must support services to query, processing and writing as shown in the example on a warfare [11]. The connectivity of all gadgets will become possible through Global Positioning System (GPS) with the goal of making mobile fog, vehicular interactions, and location and situation awareness with the Platform as a Service (PaaS) on the cloud [7]. With the aforementioned advantages of lowering the latency and improving QoS, ICN is seen as an add-on to ubiquitous as fog node.

The paper is further structured as follows: Section II presents Internet of Things (IoT) and its features, Section III is focused on discussing the possibilities and nature of the Information Centric Networking. Describing how the cloud could be linked through the name object in ICN. Section IV of the paper briefly looked at the ubiquitous environment thereby adding some features of the cloud (private and public) to enable users make communication flow with ease. Section V presents a conceptual proposal of the ICN-Fog to the Internet of Everything (IoE) while Section VI concludes the paper.

## 2 Internet of Things (IoT)

In IoT, clouds are the central point of support whereby fog computing techniques makes scalability and access to heterogeneous information possible. A typical ubiquitous system connects all gadgets and equipments of a user to the Internet as shown in figure 1. The user is connected centrally to the network which in-turn has all devices and facilities managed over the network. Network sensors, actuators and controllers perform instinctive tasks making the communication and information dissemination a lot easier. The User owns a cars, Smart Connected Vehicles (SCV), mobile devices, wearable devices such as spy-glasses, flight booking information via a smart flight monitor, weather forecaster, traffic controls through the help of the Cloud etc. On the Cloud, there are various inter-cloud communications such as the Smart Grid (SG), smart homes and smart environment which can be referred to as the private and the public clouds [12]. The essence of the User connected to other clouds as shown on the figure 1, include lack of dependency on a single cloud and source among other challenges. The major advantage aimed is to curtail the problems of data loss through disasters and other scheduling problems. All of these challenges will require having an intermediary between clouds for easy access to infrastructures, platform and software as services. Therefore, with the user fully connected as seen on the figure, the IoT becomes easier on the ubiquitous via clouds.

However, in the IoT, scheduling and power management are still open to research due to the wide connectivity of heterogeneous devices, geo-location distribution, location awareness to achieve low latency. This work foresee fog concept as a probable solution to the IoT problems alongside ICN off-network caching. Presenting the necessary parts that makes up the IoT include better Quality of Service (QoS) expected, mobility, sensors, location awareness tools and actuators, communication protocols etc. and their interconnections as shown on the figure 2 below.

IoT provides all connectivity via its connection protocols, sensors and location awareness tools. These enable IoT to support configuration, mobility in and out of the network and QoS. The cloud on the other hand with its facilities of platform as a service (example seen in Google AppEngine), software as a service (e.g. Gmail) and infrastructure as a service, additionally enable users achieve processing, visualization, storage and service discovery on the cloud easier. Applications running on smart gadgets, e-health, smart environs etc. provide the user-end the interactivity with the cloud thereby making it pervasive or ubiquitous. ICN from this point provides the names of services and information using its concept of pending interest, forwarding information and content storage mechanisms.

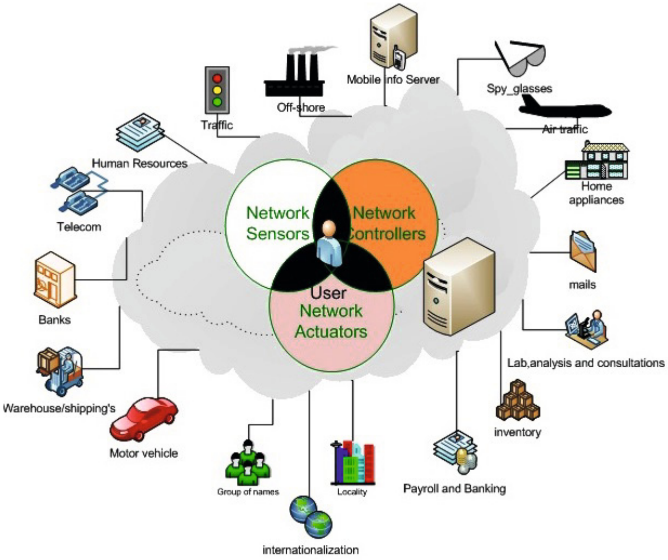


**Fig. 1.** Ubiquitous environment

## 3 Information Centric Networking

Since the predictive encroachment of IP addresses going less with users being more concerned about objects not addresses, ICN has since had an eye to solving some problems of information dissemination through name network approach [1,8]. It is therefore worthy of trying a shift in ubiquitous computing to in-cooperate this new paradigm to manage, process and retrieve information from the cloud by caching in-and off-path controlled or undemonstratively [13,14,16,17]. The motivating idea tends to have the desire of saving time to achieve low latency by retrieving demanding information on the ubiquitous heavy connectivity through ICN advantages of closest node cache. Content centric networking (CCN) has since been engaged in research by looking for means to experiment both off and on-network caching [9].

Fog computing [4,7], however with its flexibility of processing at the leaf (either through mobile devices, smart devices and Machine-2-Machine) node of the cloud, makes it inter-operable to achieve the off-network processing and caching on the network by content centric approach. The questions becomes clear as how can ICN off-network be feasible using fog computing? Take a scenario of a warfare as discussed in [11] for example, when the infantry makes a discovery or the airborne army locate a glob of enemies, all it need to do is to process the bearing location and publish (results) to the network. This in a little time will be subscribed by the first fighter Jet airborne. Reducing the vast majorities of each Jet resubscribing and wasting the bandwidth. Since Traffic Collision Avoidance Systems (TCAS) communicate (as case in aircraft), then it makes it a lot easier for fog to be practice on ICN interaction on cloud. When information is lodged

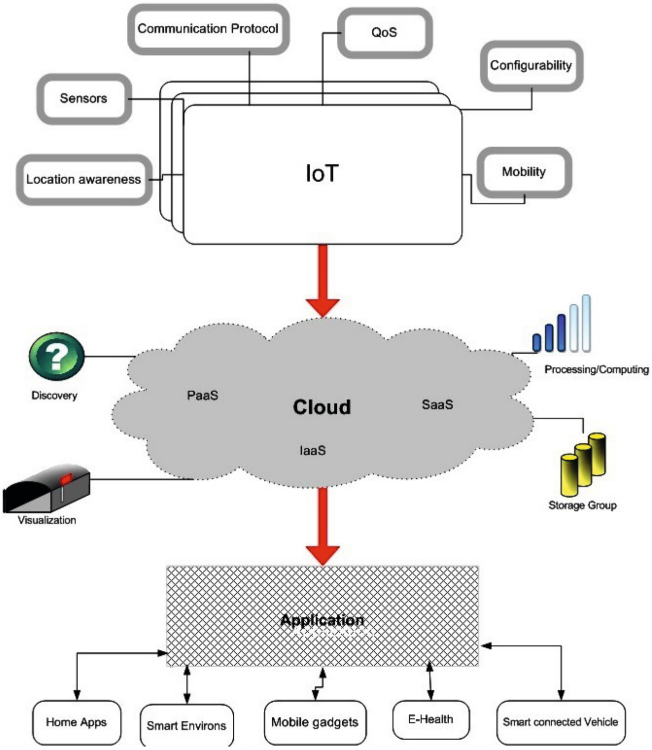


**Fig. 2.** IoT Framework

on the cloud, a user subscribes to the Internet using the name of the object requested as described in [1]. Hierarchical, flat and distributed hash functions are used in the naming. However, the cloud based on its IaaS, SaaS and PaaS would set the name to a matched publisher of such interest. Once a user gets a feedback, all nodes and router along ICN path cache the resulting information since they are cache-enabled. The ICN advantage of caching for subsequent request is thereby granted easily. Fog in ICN therefore will reduce the delay of checking on the special table in ICN known as Pending Interest Table (PIT). Upon receiving the information from the closest device through fog concept, the Forwarding Information Base (FIB) in ICN push the results into the fog unit which is saved on an autonomous system Content Store (CS) on the cloud.

This description however is open to further research to suit the objectives of not requesting by address but instead by using object names as submitted by Jacobson et al. in [8]. From the figure below, the battalion (army) discovered the enemies through smart binoculars and other gadgets with Internet connectivity facilities. Infantry then publish to the network about the requested enemy using ICN naming. The information acquired is further processed at the device node and pushed to the network. The information is accordingly cached on the first point in the network making it possible to place the results from the information

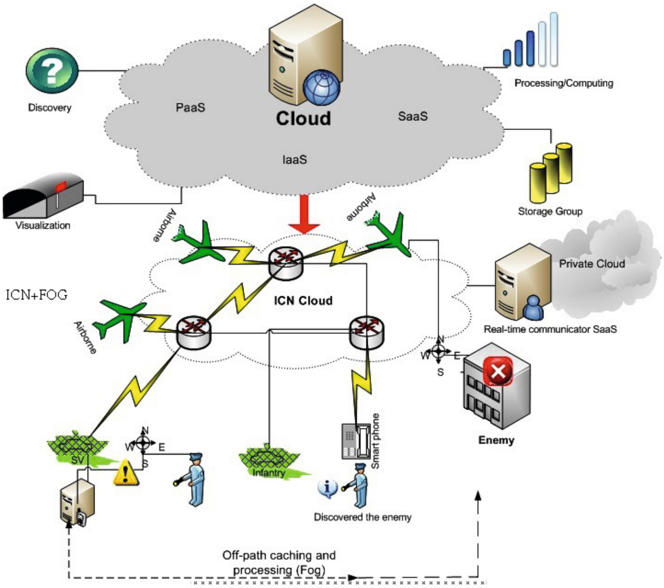


**Fig. 3.** ICN+Fog

into the content store and the forwarding base. For lower latency, reduction in bandwidth, and IP consumption, a publish query is in-turn posted by the next router point until the information gets to the air fighters as described in [15].

## 4 Ubiquitous

Ubiquitous as earlier discussed in section 1, enables all devices to be connected to a user. It was observed through various reports and forecast that all devices are needed to make life a lot easier for users through information management, dissemination, read/writes and updates in and out of the network. Since cloud computing, ubiquitous and pervasive computing is appearing to be grossly in-cooperated into our daily equipment, this would make big data handling a lot easier.

However, the suggestion and launching of the fog computing [4] makes the tedious linkage of the cloud and devices far more comfortable with processing and computations offered at the device level. This proposed in-cooperation of ICN concepts aims to curtail some challenges posed at the fog to cloud correspondences. Therefore, an optimal routine will be best suggested when cloud-fog ICN are well managed.

## 5 Conceptual Framework

To present a framework of adding ICN off-path caching to the IoT, there will need to make a linkage of the cloud with the fog idea. Figure 3 depicts the

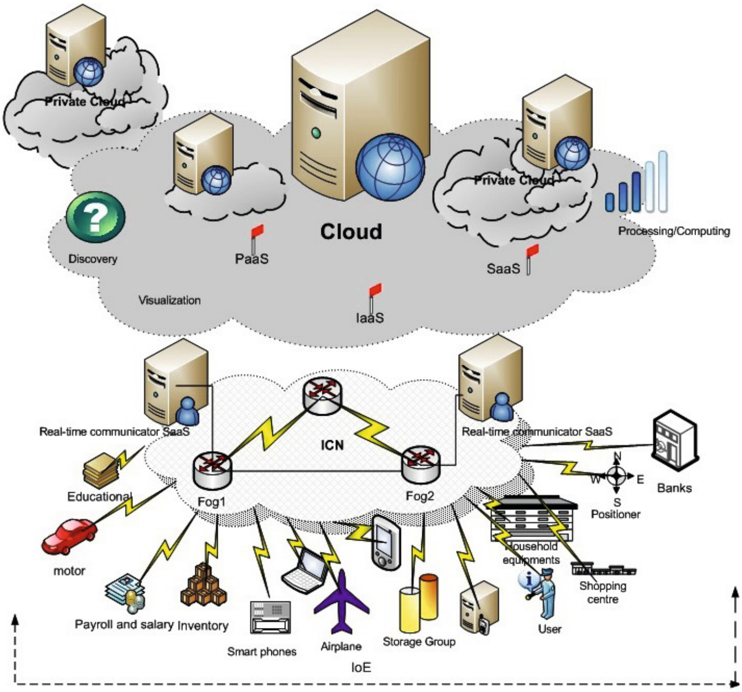


**Fig. 4.** Ubiquitous shift with ICN

branches of the cloud environment where the PaaS, IaaS and SaaS are used to make discovery, processing, visualization and storage. ICN+Fog section handles the challenging part of the off-path caching. The advantage of adding the fog leaf makes ICN the suitable paradigm for the user connecting all equipment and devices. The challenges in assigning IP addresses to all nodes and stations are thus minimized by content centric-naming. Additionally on the naming issue, propose use of the ICN concepts by either choosing the flat naming, hierarchical or the distributed hash naming schemes [1,2,19,20] would be more beneficial. Where as to optimize the case of caching, predictive distribution via subscription-publish will make private clouds connects with public cloud through the means of autonomous stations on the general cloud. Fog will then be used as an off-path of the network caching operation as all smart gadgets and devices will produce leaves of the nodes for the fetch, decode, read and write operations. Conceptual overview of the proposed caching off-path ICN-Fog to drive Internet of Everything (IoE) describes a user that set out all Internet enabled devices and appliances unto the Internet with the intention of getting information disseminated, coordinated and updated as shown on figure 4. Since scheduling is also seen as one of the research open issue of optimizing the ubiquitous, it is therefore suggested that the inclusion of the multilevel feedback queue scheduling on the ICN to send and receive information (objects) on the cloud will reduce the scheduling problem.

# 6 Conclusion

In conclusion, ubiquitous computing will make a shift of not mainly concerned on the address of where to get services, but mainly about who produces what service? This paper therefore concludes to propose a shift of ICN along ubiquitous computing by emphasizing on the off-path caching technique in the new ICN paradigm. Fog was seen as the must suiting way to implement the caching. Without adequate caching in place, there will be no way the ICN dream would be actualized. A migration of IP addresses and using name shall thus contribute to the optimal performance of ubiquitous computing to solving the problems of big data management, naming, scheduling, mobility and security. The major advantages of adding ICN to the ubiquitous computing benefits users to process at node level and publish back to the network with ease. The paper has therefore presented the concept of adding ICN off-path caching to fog computing to improve the QoS, lower bandwidth consumption among other ICN naming advantages.

**Acknowledgment.** This research is one of the Malaysian Government funded projects under Ministry of Higher Education (MOHE) Long-Term Research Grant (LRGS) with reference number LRGS/TD/2011/UKM/ICT/02.

# References

1. Ahlgren, B., Dannewitz, C., Imbrenda, C., Kutscher, D., Ohlman, B.: A survey of information-centric networking. IEEE Communications Magazine 50(7), 26–36 (2012)
2. Ahlgren, B., D'Ambrosio, M., Marchisio, M., Marsh, I., Dannewitz, C., Ohlman, B., Pentikousis, K., Strandberg, O., Rembarz, R., Vercellone, V.: Design considerations for a network of information. In: Proceedings of the 2008 ACM CoNEXT Conference, CoNEXT 2008, pp. 66:1–66:6. ACM, New York (2008), `http://doi.acm.org/10.1145/1544012.1544078`
3. Banerjee, A., Mukherjee, A., Paul, H.S., Dey, S.: Offloading work to mobile devices: An availability-aware data partitioning approach. In: Proceedings of the First International Workshop on Middleware for Cloud-enabled Sensing, MCS 2013, pp. 4:1–4:6. ACM, New York (2013), `http://doi.acm.org/10.1145/2541603.2541605`
4. Bonomi, F., Milito, R., Zhu, J., Addepalli, S.: Fog computing and its role in the internet of things. In: Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing, MCC 2012, pp. 13–16. ACM, New York (2012), `http://doi.acm.org/10.1145/2342509.2342513`
5. Cisco: Cisco visual networking index:global mobile data traffic forecast update, 2013-2018. Tech. rep., Cisco (2014)
6. Gantz, J.F., Reinsel, D., Chute, C., Schlichting, W., McArthur, J., Minton, S., Xheneti, I., Toncheva, A., Manfrediz, A.: The expanding digital universe: A forecast of worldwide information growth through 2010. Technical reoprt, An IDC White Paper- Sponsored by EMC (March 2007)

7. Hong, K., Lillethun, D., Ramachandran, U., Ottenwälder, B., Koldehofe, B.: Mobile fog: A programming model for large-scale applications on the internet of things. In: Proceedings of the Second ACM SIGCOMM Workshop on Mobile Cloud Computing, MCC 2013, pp. 15–20. ACM, New York (2013), `http://doi.acm.org/10.1145/2491266.2491270`
8. Jacobson, V., Smetters, D.K., Thornton, J.D., Plass, M., Briggs, N., Braynard, R.: Networking named content. In: Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies, CoNEXT 2009, pp. 1–12. ACM, New York (2009)
9. Li, Y., Xie, H., Wen, Y., Zhang, Z.L.: Coordinating in-network caching in content-centric networks: Model and analysis. Tech. rep., Center for Innovation, Huawei Central Research Institute (2012)
10. Madsen, H., Albeanu, G., Burtschy, B., Popentiu-Vladicescu, F.: Reliability in the utility computing era: Towards reliable fog computing. In: 2013 20th International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 43–46 (July 2013)
11. Moon, Y.W., Jung, H.S., Jeong, C.S.: Context-awareness in battlefield using ubiquitous computing: Network centric warfare. In: 2010 IEEE 10th International Conference on Computer and Information Technology (CIT), pp. 2873–2877 (June 2010)
12. Petcu, D.: Multi-cloud: Expectations and current approaches. In: Proceedings of the 2013 International Workshop on Multi-cloud Applications and Federated Clouds, MultiCloud 2013, pp. 1–6. ACM, New York (2013), `http://doi.acm.org/10.1145/2462326.2462328`
13. Psaras, I., Chai, W., Pavlou, G.: In-network cache management and resource allocation for information-centric networks. IEEE Transactions on Parallel and Distributed Systems PP(99), 1–1 (2014)
14. Psaras, I., Chai, W.K., Pavlou, G.: Probabilistic in-network caching for information-centric networks. In: Proceedings of the Second Edition of the ICN Workshop on Information-centric Networking, ICN 2012, pp. 55–60. ACM, New York (2012), `http://doi.acm.org/10.1145/2342488.2342501`
15. Vasilakos, X., Siris, V.A., Polyzos, G.C., Pomonis, M.: Proactive selective neighbor caching for enhancing mobility support in information-centric networks. In: Proceedings of the Second Edition of the ICN Workshop on Information-centric Networking, ICN 2012, pp. 61–66. ACM, New York (2012), `http://doi.acm.org/10.1145/2342488.2342502`
16. Wang, S., Bi, J., Wu, J.: Collaborative caching based on hash-routing for information-centric networking. In: ACM SIGCOMM 2013 Conference on SIGCOMM, Hong Kong, China, SIGCOMM 2013, pp. 535–536. ACM (2013)
17. Wang, S., Bi, J., Wu, J., Li, Z., Zhang, W., Yang, X.: Could in-network caching benefit information-centric networking? In: Proceedings of the 7th Asian Internet Engineering Conference, AINTEC 2011, pp. 112–115. ACM, New York (2011), `http://doi.acm.org/10.1145/2089016.2089034`
18. Weiser, M.: The computer for the 21st century. IEEE Pervasive Computing 1(1), 19–25 (2002)
19. Zhang, X., Chang, K., Xiong, H., Wen, Y., Shi, G., Wang, G.: Towards name-based trust and security for content-centric network. In: 2011 19th IEEE International Conference on Network Protocols (ICNP), pp. 1–6 (October 2011)
20. Zhu, Y., Nakao, A.: A deployable and scalable information-centric network architecture. In: 2013 IEEE International Conference on Communications (ICC), pp. 3753–3758 (June 2013)

# Cloud Location Verification and Adversary Identification

Asadullah Al Galib[1] and Rashedur Mohammad Rahman[2]

[1] Department of Computer Science, University of Calgary, 2500 University Drive N.W., Calgary, Alberta T2N 1N4, Canada
aagalib@ucalgary.ca
[2] Department of Electrical and Computer Engineering, North South University, Plot-15, Block-B, Bashundhara, Dhaka 1229, Bangladesh
rashedur@northsouth.edu

**Abstract.** When it comes to cloud computing, people get concerned not only of privacy and integrity of their data but its location as well. This is particularly important due to the existence of variations of laws and regulations governing data storage and access mechanisms in different geographical regions. We will consider a distributed environment such as the cloud and try to verify whether the client-requested servers are holding the file they actually claim to hold. We make some plausible assumptions on the adversarial model and propose a scheme called CLVI for finding out geographical location of servers. Our scheme uses a challenge-response protocol instead of traditional pinging mechanism to accomplish its task. CLVI is able to detect location forgery and identify the adversary involved in the forgery. We validated our technique of detecting location forgery and approximating the adversary's location using measurements from PlanetLab.

## 1 Introduction

Finding out the geographical location of Internet hosts has been an important issue to ponder upon over the recent years, due to the numerous services, restrictions, fraud prevention systems that depend on it. Sites like Hulu would geo-locate the users before allowing them to access a video. Online gambling centers would need to know the location of their users. This is necessary to verify whether the region where the user resides enforces law against gambling. This mechanism of finding out location (i.e. geo-locating) can be used for enforcing SLA (Service Level Agreement) on location of data in a distributed system such as the cloud. There are many organizations that could benefit from cloud computing; however there are certain laws which confine their choice of storage regions. This confinement of choices generally arises from the imposed regional data privacy and data access law.

Cloud computing gives users and providers greater freedom to place their storage in specified location. While storing files, users can specify the region where they would like their files to be stored. The users can later verify if their

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems*,
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_33

files have been stored properly in the specified locations. Therefore, the task of ensuring the integrity of the file as well as identifying the location of the file becomes important to the user. Consider a scenario, where the user's file has not been stored properly in a requested region. If the user is able to verify such phenomenon he can take legal actions against the store-provider since it violates the SLA.

We make some plausible assumptions on the adversarial model which allowed us to construct a scheme that can reliably verify storage location. [1] provides a survey on the different geo-locating algorithm and then discusses and evaluates the ways adversaries can attack them. Typically, adversaries committing location forgery rely on shortening or lengthening the delay time. For instance, consider a scenario where we ping a location which is within a distance of 1000 KM from us and record the round trip time (RTT). This recorded RTT must be shorter than the RTT which can be obtained by pinging a location 100,000 KM away. Therefore, if a terminal 100 KM away takes abnormally large amount of time to respond while pinged, the user on the other end will definitely think that the terminal is located more than 1000KM away. This is how the adversary can manipulate the delay time to falsify its location when pinged or challenged. We take into consideration these kinds of adversarial behaviors in a cloud environment and propose a technique which can detect location forgery and subsequently identify the adversary lengthening the RTT. Our technique used for finding out the location of a specific server and detecting a location forgery is based on trilateration as discussed in [2]. Trilateration is the process of estimating a position using a sufficient number of distances to some fixed points. In our scheme, we send challenges from some reference/fixed points to the point of our interest or target server. We call this reference points *landmarks* throughout this paper. Additionally, we propose a new technique which is able to approximate the location of the adversary which elongated the delay time and committed location forgery. In a nutshell, our goal is to guarantee a scheme which would be able to verify a storage location and detect location forgery in a cloud environment. If a user can prove formally that the storage location is not the valid/requested one; he will be capable of taking legal actions against the service provider, since it will violate the SLA. To facilitate this process our scheme even extends to identify (approximately) the corrupted adversary's location.

The rest of the paper is organized as follows: Section 2 provides discussion on our scheme. Section 3 & 4 discuss about forgery detection and adversary's location identification receptively. Section 5 provides empirical results. Related works are explored in Section 6. And finally, we conclude in Section 7.

## 2 The CLVI Scheme

### 2.1 Environment Setup and Threat Model

Figure 1(a) shows a distributed file storage system, such as the cloud, where the user sends a request to the store-provider for saving a file at a particular region. So the store provider receives a request containing the file, $z$, and the region, $R_1$, where it needs to be stored. The store-provider then runs an auction among the servers. Auction is a process whereby the store-provider is able to select a group of eligible servers at a user requested region. The process of auctioning is outside the scope of this paper. In Figure 1(a), the store-provider selects server $s_1$, $s_2$ and $s_3$ in the region $R_1$ to store the user's file $z$. We assume that the user has requested the store-provider to store his files in any three servers in the region $R_1$. Later, the user can get a list of servers which currently hold his files just by querying the store-provider. The user also has the ability to employ certain trusted reference points or landmarks which will be able to send challenges to the target server and verify the integrity of the file and record the challenge and response time to calculate the RTT. The user can then convert this delay time into geo-graphical distance and verify the location of the target servers. Thus the aim of the user here is to verify both the integrity and location of the file storage and identify adversary in case of location forgery. We now consider some non-adversarial and adversarial models and explore in each case, whether the user is successful in achieving his goals.

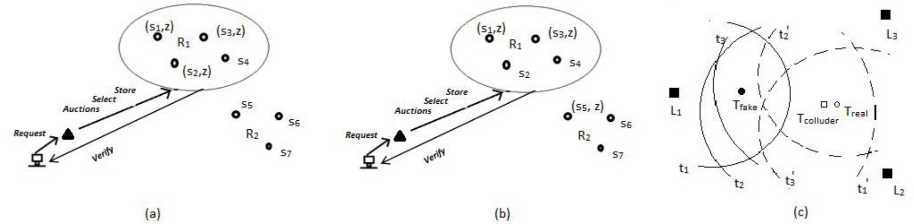


**Fig. 1.** (a) Non-Adversarial Model-I, (b) Non-Adversarial Model-II, (c) Adversarial Model-I

*Non-Adversarial Model-I:* Figure 1 (a) is an example of a non-adversarial environment. The store-provider is honest and stores the user's file $z$ in the region $R_1$ as requested. However, the store provider does not verify the location of the file once it has been stored. Therefore, we consider our first non-adversarial model to be an environment where the store-provider is unwilling to verify the location of the file. All the entities of this environment, the user, store-provider, and servers are honest. The only way a file can get stored in a server in a non-user requested region might be due to some protocol error or mismanagement of physical infrastructure. For example, the file might get stored in server $s_5$ instead of $s_3$ accidentally. The user will first query the store-provider to provide him with a list of servers which are currently holding the file. The store-provider then sends a list of servers containing $s_1$, $s_2$ and $s_3$ to the user. The user employs the landmarks to initiate a challenge-response protocol with each of the three servers in the list. However, server $s_3$ will not be able to response properly since it does not have the file and the user will detect a failure in response.

*Non-Adversarial Model-II:* We consider another non-adversarial model. Here the store-provider is not malicious but wants to make some profit by storing the files in low cost servers. Figure 1(a) shows that the user's file $z$ is getting stored in $s_1$, $s_2$ and $s_3$ which are in user's request region $R$. However, consider server $s_5$ offers a cheaper storage system, therefore, the store-provider stores a copy of file $z$ in server $s_5$ instead of storing it in server $s_2$ in region $R_1$, Figure 1 (b). Notice that $s_5$ is located in another region. This time when the user requests the store-provider to send a list of server, the list will consist of server $s_1$, $s_2$, $s_5$. The challenge and response protocol will be initiated for each of the servers. And for the server $s_5$, the integrity test will pass, but when the delay time will be translated to geo-graphical distance the user will come to know that $s_5$ falls out of his requested region.

*Adversarial Model-I*: Now we consider an environment where the store-providers and the servers are malicious and they want to deceive the user into thinking that his file is stored in his requested region, but in fact it is stored somewhere else. Consider that the server $T_{fake}$ is holding the user's file $z$, Figure 1 (c). However, the user requested region is in the dotted enclosed region. Therefore $T_{fake}$ adds some delay to the Landmarks $L_1$ s.t. the location appears to be in $T_{real}$. Now consider the measurements from the Landmark $L_2$ and $L_3$. Since $T_{fake}$ is further away from $L_2$ and $L_3$ than $T_{real}$, the adversary $T_{fake}$ *somehow* needs to shorten the time $t_2$ and $t_3$. However, shortening time violates the law of physics. In such circumstances, the store-provider can collude with another server $T_{colluder}$ which is very close to $T_{real}$. $T_{fake}$ then sends a copy of the file to $T_{colluder}$ which in turn can answer any challenges from the landmark $L_2$ and $L_3$. This shortens the delay time from $t_2$ to $t'_2$ and $t_3$ to $t'_3$ for landmarks $L_2$ and $L_3$ respectively. The resulting dotted region now includes $T_{real}$. This shows that colluding helps in performing both lengthening and shortening strategy and makes it impossible for any location verification system to detect forgery. However, the probability that such a colluding server will be present in a user requested region is very small [1].

*Assumption 1:* It is therefore necessary to assume that the service provider will not collude with a server which lies within a short distance, $\varepsilon$, of $T_{real}$ s.t. $\varepsilon$ allows successful distance faking.

*Adversarial Model-II:* Due to *assumption 1*, we are left with one possible adversarial model. This is when the adversary has the ability to only lengthen the delay time in order to fake the location. The adversary can add appropriate delay by pinging each of the landmarks from the forged location. In that case, the adversary needs to control a machine near $T_{real}$. However, pings to landmarks from a machine not related to the geolocation algorithm may arouse suspicion [1]. We, therefore, make the following assumption in this model.

*Assumption 2:* We assume that the adversary does not know exactly what delay to add. So the adversary uses 2/3 the speed of light, $c$, as a lower-bound approximation for the network traffic speed to calculate delay [3].

We use CLVI scheme in *Adversarial Model-II* to verify the integrity and location of the file. In case of a location forgery CLVI identifies the adversary involved in the foul play. The scheme provides an approximation to the location of the adversary, $T_{fake}$.

### 2.2 Asserting Location Using Landmarks

Delay-based geo-location algorithm uses measurements of network delays to geo-locate the target IP. In our work, we use trilateration which converts the delay measurements to geographic distance constraints similar to [2]. The landmark $L_i$ needs to send challenges to other landmarks at some distances and record the corresponding delays. A graph of each landmark $L_i$ is then constructed with delay against distance Figure 3(a). A line $y = mx + c$ is constructed in such a way so that it is closest to, but below all the data points. We refer to this line $y_i = mx + c_i$ as the *best* line for landmark $L_{i.}$. The *best* line is *not* the same as the typical *best fit* line found using linear regression. The *best fit* line is above the *best* line as shown in Figure 3(a). Each landmark then can use its *best* line to convert any delay measurements into a geographic distance. It is easy to interpret from the figure that any distance found from the *best* line is an overestimation. Each landmark then constructs a circle, its location being the centroid, with radius set to the geographical distance found from its *best* line. Each landmark $L_i$ estimates that the target host $t$ is somewhere near its circumference $C_{it}$. Given $K = 3$ landmarks for trilateration, the target host $t$ has a collection of closed curves $C_t = \{C_{1t}, \ldots, C_{Kt}\}$. The intersection of all the closed curves or the special region $R$ provides the user with a continuous set of answers about the location of the target host $t$. To obtain a discrete point, the centroid of the region is found. A polygon or triangle is formed with vertices $v_0 = (x_0, y_0), \ldots, v_{N-1} = (x_{N-1}, y_{N-1})$. The centroid is $C_x = \frac{1}{6A}\sum_{i=0}^{n-1}(x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i)$ & $C_y = \frac{1}{6A}\sum_{i=0}^{n-1}(y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i)$ with area $A = |\frac{1}{2}\sum_{i=0}^{n-1} x_i y_{i+1} - x_{i+1} y_i|$. Figure 2 (a) is an example of how three landmarks uses trilateration to get the location of the target host.

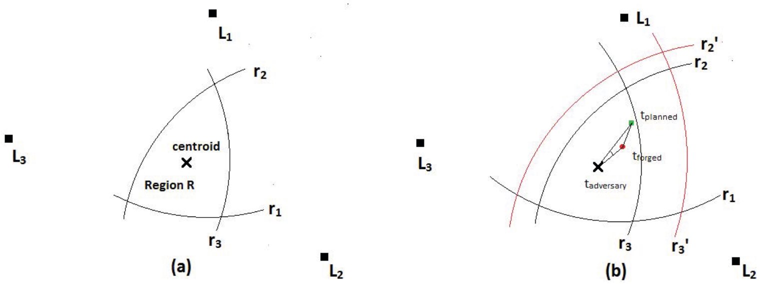


**Fig. 2.** (a) Geo-locating using Trilateration, (b) Location forgery

### 2.3 Overview of CLVI Scheme

Our CLVI scheme consists of four stages. In the first stage known as the *Store*, the user encodes his file $z$ and creates a corresponding file tag. The encoded file along with its tag is sent to the store-provider. The store-provider selects some servers by auctioning and sends the whole package to them. Now the user wants to verify both the integrity and the location of the file and identify adversary in case of forgery. He therefore, initiates the *Verify* stage, where the user employs a set of landmarks to verify whether copies of the encoded file actually exist in the claimed location. Notice that the user can ask the store-provider to provide a list of servers $S$ that are holding his file *z*. The landmarks then send challenges to the list of servers $S$ to verify whether they respond correctly. In the challenge-response protocol, each landmark send individual challenges to each target server and records the delay time. That means for each delay $d$ exist a pair of $(l \in L, s \in S)$ where the servers and landmarks can communicate with each other. In the third stage called *Locate* the delay time for each Landmark are translated to geographical distance using the *best* line. And eventually we get a region where the centroid is a probable location of a target server. The fourth stage verifies whether the target location has been forged. If forged, the scheme attempts to identify the location of the malicious adversary. The challenge response protocol used in our scheme is similar to [4,5]. We therefore do not emphasize much on the protocol. Simpler methods such as pinging can also be incorporated in CLVI.

**Stage 1: Store File**

*Encode:* The user or verifier encodes the file $M$ and obtains the encoded file $M'$. $M'$ is divided into $n$ blocks $m_1, m_2, \ldots, m_n \in \mathbb{Z}_p$, $p$ is a prime. A random number $\theta$ is chosen from $\mathbb{Z}_p$. For each block $\{m_i\}$ an authenticator $\{\sigma_i\}$ is created using a pseudorandom function(PRF) $f_{k_{prf}}$ i.e. $\sigma_i = f_{k_{prf}}(i) + \theta m_i$ *(where* $\theta \xleftarrow{r} \mathbb{Z}_p$*)*.

*Store:* The store-provider runs auction to select the eligible servers $S$ in the requested region *R*. The encoded file $M'$ along with the file tag and authenticators $\{\sigma_i\}$ get stored in the server.

**Stage 2: Verify Response**

*Query:* The user queries the store-provider to provide a list of server $S$ that is holding the file *M'*.

*Challenge:* The employed landmark initiates a challenge-response protocol between the trusted landmarks and the servers. Each Landmark $l$ sends a challenge C containing $w$ random indices $i$ from $\{1, \ldots, n\}$ and $w$ random coefficients $v_i$ from $\mathbb{Z}_p$ i.e. $C = \{(i, v_i)\}$ and outputs the time the challenge is sent as $t_c$.

*Response:* The server on receiving a challenge C computes a response pair $r = (\sigma, \mu)$, $\sigma = \sum_{(i,v_i)\in Q} v_i \sigma_i$ and $\mu = \sum_{(i,v_i)\in Q} v_i m_i$ and sends them back to the landmarks.

*Verification:* The landmark on receiving the response $r$ outputs time $t_r$. It checks if $r$ is correct as follows $\sigma = \sum_{(i,v_i)\in Q} v_i f_{k_{prf}}(i) + \theta\mu$. If incorrect, the response pair and corresponding $(t_c, t_r)$ is discarded and a new challenge is initiated.

**Stage 3: Locate Target Server**

*DelayTranslation*: This step takes into $(l, s, t_c, t_r)$ for each landmark and outputs the corresponding delay *d.* This delay is then transferred into geographical distance using the *best* line of the landmark *l.*

*TargetLocate:* Then using trilateration as mentioned previously, a region $R$ is constructed where the centroid gives the location of the target server *s.* We then add $s$ to a region vector $L$. This is repeated for the rest of the servers claiming to hold the file. If $L \subseteq R$, then our verification is successful. Otherwise, the client's requested file was not stored in the requested region.

**Stage 4: Detect Forgery & Identify Adversary**

*ForgeryDetect:* If for any target server s, the area of the region R crosses a certain threshold $th_A$, then a location forgery takes place.

*Identify Adversary:* The radii of some of the arcs from the corresponding landmarks are then decreased until the area of the enclosed region $R \leq th_A$. The centroid of this newly created region $R$ is the approximate location of the adversarial server which took part in the delay lengthening attack.

## 3 Detecting Location Forgery

We consider *Adversarial Model-II* and provide overview on how a location forgery can be detected. Subsequently, we attempt to find the adversary who took part in location forgery. Figure 2 (b) is an example of an adversary $t_{adversary}$ trying to deceive the user into thinking that his file is being held at location $t_{planned}$. The arcs labelled $r_1$, $r_2$ and $r_3$ belong to landmarks $L_1$, $L_2$ and $L_3$ respectively. The region enclosed by these three arcs gives us a continuous space of points where the server $t_{adversary}$ must lie. The centroid of the region is the closest estimation to the location of $t_{adversary}$.

In order to fake the location the server $t_{adverary}$ must shift the centroid of the region towards $t_{planned}$. Therefore, $t_{adversary}$ acts maliciously and increases the delay time of landmarks $L_2$ and $L_3$ by certain amounts. The adversary knows the latitude and the longitude of the $t_{plannand}$ and thus can calculate the distance between itself and $t_{planned}$. Then taking into consideration 2/3 of the speed of light, he can infer the amount of delay that needs to be added for each landmark $L_2$ and $L_3$ After adding the delays, the radii $r_2$ and $r_3$ are now increased to $r'_2$ and $r'_3$ as shown in Figure 2(b). The arc $r_1$ can remain unchanged since the direction of the movement of the centroid does not require any change to $r_1$. This results into a larger feasible region with a centroid $t_{forged}$. If this large area of the region enclosed by $r_1$, $r'_2$ and $r'_3$ crosses an experimental threshold $th_A$ value the user is alerted of a location forgery.

However the adversary $t_{adversary}$ was not able to forge the location properly, Figure 2(b). The adversary planned to move the location to $t_{planned}$ but turned out that using the delay calculated considering 2/3 of the speed of light, his best effort to shift the location is $t_{forged}$. We assume that the adversary will always attempt to keep $\delta = R - th_A$ as small as possible to avoid detection. This consequently implies that the arc opposite of the direction of movement must always remain unchanged e.g. $r_1$ in the above case. This leads to an error distance $\triangle_e^d$ $(t_{forged}\ to\ t_{planned})$ and an error angle $\triangle_e^\theta$ $(\angle t_{planned} t_{adversary} t_{forged})$.

## 4 Approximating Adversary's Location

It is possible for the user to take legal action against the store-provider if he is successful at detecting a location forgery. However, the issue can be made stronger if the user is able to identify the adversary who participated in the forgery. Let us consider the scenario in Figure 2(b) where the adversary $t_{adversary}$ makes its location appear to be in $t_{forged}$. In the process it increases the feasible region $R$. Now, in order to identify the adversary $t_{adversary}$, the user first needs to identify the increased radii e.g. $r'_2$ & $r'_3$. Then those radii can be decreased until $R \leq th_A$. At that point the centroid of the reduce region $R'$ must provide an approximate to $t_{adversary}$. We now generalize the notion of keeping one of the arcs constant while decreasing the other two radii. $t_{adversary}$ is the actual server holding the file and responding to the protocol. The attempt of lengthening the time using 2/3 of the speed of light and moving the location to $t_{planned}$ infers that $t_{forged}$ is bound to be located somewhere between $t_{adversary}$ and $t_{planned}$. We denote $\theta(t_f \rightarrow t_p)$ as the bearing of $t_{planned}$ relative to $t_{forged}$. This angle measured in a clockwise direction from the north line will then be used to identify the arc of the landmarks that need to remain fixed. We find the landmark with a bearing of $\theta(t_f \rightarrow t_p) \pm \triangle$ relative to $t_{forged}$. $\triangle$ is increased from 0 incrementally until the first landmark is sighted. The arc of the sighted landmark is then kept fixed. Intuitively, given a single point $p$ on the circumference of a circle, if the centroid of the circle is shifted with $p$ fixed then all other points on the circumference of the circle is bound to get displaced. Assuming that a successful attacker will always try to minimize the aforementioned $\triangle_e^d$ and $\triangle_e^\theta$, the other two radii is decreased s.t. the new centroid $c'$ of $R'$ is collinear with $t_{planned}$ and $t_{forged}$. Larger $\triangle_e^\theta$ will thus result into $c'$ being located further away from $t_{adversary}$. To minimize the effect of $\triangle_e^\theta$, multiple trilaterations are executed using a combination of the available landmarks. More formally, for L landmarks we carry out $\binom{L}{3}$ trilateration. This provides us with $C = c'_1,\ c'_2, \ldots,\ c'_{\binom{L}{3}}$ centroids. We use K-means clustering on $C$ and choose K using the elbow method [6]. In an elbow method the cost function, $J$, is first plotted against the number of clusters, $K$. The "elbow" point $p$ is a point on the plot s.t. for $K<p$, $J$ decreases relatively rapidly to when $K \geq p$. $K$ is then set to $p$. The centroid of the densest cluster out of the $K$ is then considered as the location of $t_{adversary}$. If more than one such centroid exists, the result is discarded and the identification process is executed again.

## 5 System Evaluation

We carried out experiments to evaluate our proposed CLVI scheme. We simulated the landmarks and servers using Planet-lab nodes [7]. A total of 17 nodes spread across North America(13) and Europe (4) were used. These nodes can act as landmarks as well as servers. In the *Learning* phase for all landmarks we constructed the *best* lines which act as the delay-to-distance functions. In the *Locate* phase we considered one of the landmarks as the target server and initiated our challenge-response protocol from 3 other landmarks. The protocol outputs the delay for each landmark and the target server. These delays are then converted to geographical distance so that each landmark can construct a circle around it. The union of all the circles is the feasible region where the target server may lie. The centroid of the region is considered to be the location of the server. In the *Forgery Detection* phase, we compare whether the estimated area of the region has crossed a certain threshold. If crossed, we flag the location as a fake location and execute the last phase called *Adversary Identification.* The goal is to find a location, which is very near to the adversary who participated in lengthening the delay time.

### 5.1 Details of the Experiment

**Learning Phase.** The objective of this phase is to assign each landmark a *best* line function. We start off with the 12 landmarks where each one of them acts as a server with all other nodes acting as landmarks. Each landmark issued a total of 100 challenges to the server. This mean each landmark $l \in L$ will have 100 delay values $(d_1, d_2, .. d_{100})$ for a particular target server $t$ which is located at a distance $h$ from the landmark $l$. We then constructed a scatter plot where the y-axis represents the delay or RTT and the x-axis represents the distance from the landmark similar to Figure 3(a). The geographical distance is calculated using Haversine formula [8]. It calculates the distance between two points on a sphere from their longitudinal and latitudinal values.

**Locate Phase.** Once we get the delay functions for each landmark, we select a target server and initiate a challenge-response protocol. 5 planet lab nodes were chosen as the target servers and were kept isolated and did not participate in the learning phase. Figure 3(b) shows a target server *US12* and the landmarks *US17, US18, US15, US7* surrounding it. *US18, US15, US7* initiates the challenge-response protocol and the corresponding delays are recorded. The landmarks now can use their *best* line to calculate the Haversine distance to the target server. Figure 3(c) shows the radii in 2 dimensional Cartesian co-ordinates system. The latitudes and the longitudes are converted into corresponding x and y values and then the circles are plotted using MATLAB. Notice that the circles overlap over a certain region and the centroid of this region is an estimate to the location of US12 i.e. the target server.

**Forgery Detection Phase.** It is critical to setup a threshold $th_A$ on the region area $R$ which when crossed will alert the user of a location forgery. The mean area of the region from the experiment was found to be 190,459 $KM^2$ and the mean difference in distance between the actual target server and the centroid of the region is 127.34 KM. $th_A$ is set to the *upper outer fence* of the box plot using interquartile range [9] $Q3 + 3 * (Q3 - Q1) \approx 250,000KM^2$ (around the size of Utah) where $Q3$ & $Q1$ are the upper and lower quartiles respectively. Any region exceeding the threshold would be considered as a region having the forged location. Let us consider that US12, from the previous example, is an adversary who wants to deceive the user in believing that the user's file is stored in some location at Nebraska as shown in Figure 3(b) by a light blue arrow. The adversary has access to the Haversine distance, $\alpha$, from US 12 to the location at Nebraska. It calculates the amount of delay $d'$ to be added to each landmark $l$ considering the speed of network traffic as 2/3 the speed of light. $d' = \frac{2\times\alpha}{\frac{2}{3}\times c}$.

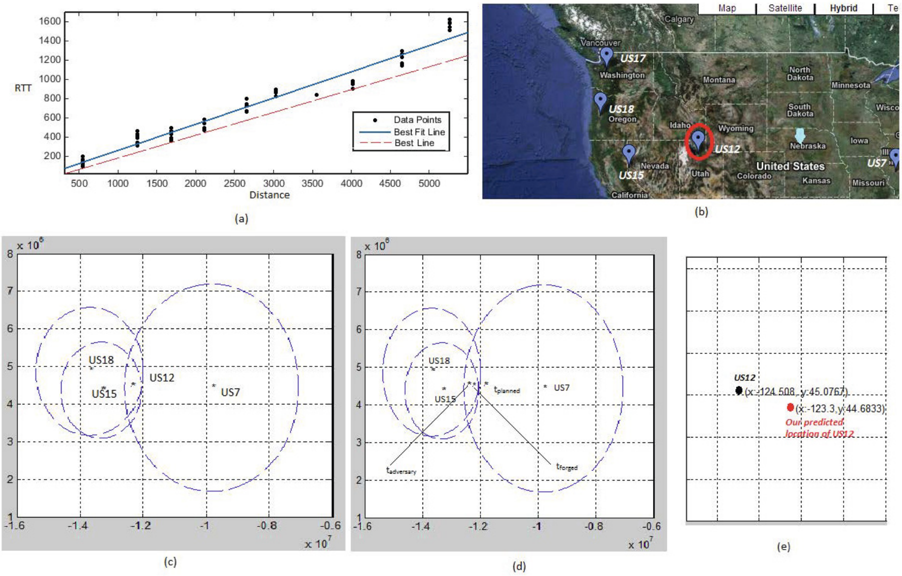


**Fig. 3.** (a) Distance-to-delay function, (b) Target server US12 surrounded by few landmarks, (c) Trilateration to locate US12, (d) Location forged, (e) Approximate adversary location.

US18, US15 and US7 are the landmarks that surrounds the adversary US12. To fake the location to Nebraska, US12 needs to shift its location rightwards, so it adds delay to the original delay of US18-US12, $\triangle_{d+d'}^{US18-US12}$, and US15-US12, $\triangle_{d+d'}^{US18-US12}$. The adversary is not able to shorten the delay of US7-US12. Using the new delays the landmarks constructs a region and finds out the centroid of the region. This time the area of the feasible region becomes 322,905 KM$^2$ and ultimately crosses the threshold that we had previously set. The user is then alerted that a location forgery has taken place and he records the forged location, i.e. the centroid of the region as $t_{forged}$, Figure 3(d).

**Adversary Identification Phase.** CLVI initiates this phase on detection of location forgery. It recognizes the radii that need to be decreased. CLVI now attempts to identify the adversary US12. It increments $\triangle$ in $\theta(t_f \rightarrow t_p) \pm \triangle$ and infers that the radius of Landmark US7 must remain unchanged. It reduces the other two radii until $th_A \leq 250,000$ and s.t. the new centroid is collinear with $t_{planned}$, and $t_{forged}$. This process is repeated $\binom{12}{3}$ times to generate vector $C$. K-means clustering is applied to $C$ with K fixed using elbow method. The densest cluster centroid is the closest point to the adversary as shown Figure 3(e). The error in the estimation of the adversary location is 169.32 KM i.e. the difference between $t_{identified}$ and t$_{adversary}$. Around 82% of the time a location forgery was detected. Targets were moved to 20 forged locations and the mean difference in distance is found to be approximately 440 KM. This implies that an adversary carrying out a forgery attack can be located with some approximation.

## 6 Related Work

There is considerable amount of work on delay-based geolocation such as statistical [10], learning based geolocation [11]. Our approach of translating a delay to distance and using trilateration is similar to [2]. [1] evaluate attacks on both delay-based geolocation and topology-aware techniques. However, the authors [1] do not provide any technique for identifying the adversary involved in the attack. Similarly, [12] provides a survey of geolocation technique and their applicability in the presence of an adversarial target. However, they emphasize on the use of external sources of information such as (IP, DNS) for geolocating. We used a challenge-response protocol which is similar to [4,5]. These challenge-response protocols allow a user to verify efficiently the existence and the integrity of a file. Variants of this protocol have later been used in many researches [13]. Our CLVI scheme can incorporate any protocol that is capable of constructing the distance-to-delay function and therefore, is independent of the protocol used.

## 7 Conclusion

In this paper, we have proposed a scheme CLVI which enable the user in a distributed storage system such as the cloud to verify the integrity and location of the file storage. We considered both non-adversarial and adversarial environment and showed that CLVI is applicable in all those environments. Additionally, our scheme provides technique through which user will be able to detect location forgery and identify the adversary. CLVI can aid the user in taking legal actions against store-providers who violate the Service Level Agreement.

## References

1. Gill, P., Ganjali, Y., Wong, B., Lie, D.: Dude, where's that ip?: circumventing measurement-based ip geolocation. In: Proceedings of the 19th USENIX conference on Security, p. 16. USENIX Association (2010)
2. Gueye, B., Ziviani, A., Crovella, M., Fdida, S.: Constraint-based geolocation of internet hosts. IEEE/ACM Transactions on Networking 14(6), 1219–1232 (2006)
3. Katz-Bassett, E., John, J.P., Krishnamurthy, A., Wetherall, D., Anderson, T., Chawathe, Y.: Towards ip geolocation using delay and topology measurements. In: Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement, pp. 71–84. ACM (2006)
4. Shacham, H., Waters, B.: Compact proofs of retrievability. In: Pieprzyk, J. (ed.) ASIACRYPT 2008. LNCS, vol. 5350, pp. 90–107. Springer, Heidelberg (2008)
5. Juels, A., Kaliski Jr, B.S.: Pors: Proofs of retrievability for large files. In: Proceedings of the 14th ACM Conference on Computer and Communications Security, pp. 584–597. ACM Press, New York (2007)

6. Tibshirani, R., Walther, G.: Cluster validation by prediction strength. Journal of Computational and Graphical Statistics 14(3), 511–528 (2005)
7. Planet-lab, `http://www.planet-lab.org/` (last accessed June 29, 2014)
8. Needle, C.L., Catarino, R.: Evaluating the effect of real-time closures on cod targeting. ICES Journal of Marine Science: Journal du Conseil, fsr092 (2011)
9. Box plot rule, `http://www.itl.nist.gov/div898/handbook/` (last accessed 29 June 2014)
10. Youn, I., Mark, B.L., Richards, D.: Statistical geolocation of internet hosts. In: Proceedings of 18th Internatonal Conference on Computer Communications and Networks, ICCCN 2009, pp. 1–6. IEEE (2009)
11. Eriksson, B., Barford, P., Sommers, J., Nowak, R.: A learning-based approach for IP geolocation. In: Krishnamurthy, A., Plattner, B. (eds.) PAM 2010. LNCS, vol. 6032, pp. 171–180. Springer, Heidelberg (2010)
12. Muir, J.A., Oorschot, P.C.V.: Internet geolocation: Evasion and counterevasion. ACM Computing Surveys (CSUR) 42(1), 4 (2009)
13. Benson, K., Dowsley, R., Shacham, H.: Do you know where your cloud files are? In: Proceedings of the 3rd ACM Workshop on Cloud Computing Security Workshop, pp. 73–82. ACM (2011)

# AutiSay: A Mobile Communication Tool for Autistic Individuals

Nyuk Hiong Voon[1], Siti Nor Bazilah[2], Abdullah Maidin[1],
Halina Jumaat[1], and Muhammad Zulfadhli Ahmad[1]

[1] School of Computing and Informatics,
Institut Teknologi Brunei, Jalan Tungku Link, Gadong, Brunei Darussalam
[2] School of Business,
Institut Teknologi Brunei, Jalan Tungku Link, Gadong, Brunei Darussalam
`jennifer.voon@itb.edu.bn`

**Abstract.** Children diagnosed with Autism Spectrum Disorders (ASD) often have social communication difficulties and/or display restrictive and repetitive behaviours. This paper focuses on the aspect of the social communication difficulties and the design of a mobile application to be used as a communication tool between the autistic child and his/her caregivers. The advances in mobile and ubiquitous technologies provide an opportunity to develop a communication tool to improve the social communication of the autistic child and therefore alleviate the quality of life for the child and those around him/her. Relevant design requirements were identified and incorporated into an application that is highly customizable to the individual child. The application is now available on the Apple app store and as its usage widen, the team intends to collect feedback to continue to refine the application.

**Keywords:** Mobile application, autism spectrum disorders, autism, assistive technologies for persons with disabilities, social communication difficulties, computer applications, design considerations for ASD applications, apps.

## 1 Introduction

According to Gartner, the device market is forecasted to grow 6.9% in 2014 and 5.9% in 2015. This translates to approximately 2.5 billion units shipped in 2014. Tablet is shipping at an increase of 38.6 % while desktop PC is at a decline of 6.6% from 2013 [1]. This forecast shows that 2014 is the year that the number of tablets shipped is increasing much faster in proportion to desktop devices.

The implication is that the choice of device for work, social and communication will be a hand-held mobile device in some form. Mobile applications are on the increase and the IOS App store and the Google Play dominate the volume in applications. To date, IOS App store has over 1.2 million applications [2] and according to AppBrain, the number of applications on Google Play is about 1.3 million [3].

S. Phon-Amnuaisuk and T.-W. Au (eds.), *Computational Intelligence in Information Systems,*
Advances in Intelligent Systems and Computing 331, DOI: 10.1007/978-3-319-13153-5_34

Arguably, mobile applications have become platforms for social change, education, lifestyle and health to some extent. This should not be a surprise as current mobile devices such as smart phones and tablets are equipped with rich-features which encourage the development of applications to improve the quality of human life.

More mobile applications on health-related issues are appearing on the landscape such as applications for monitoring sleep apnea, diabetes, education and communication applications for Autism Spectrum Disorder (ASD) and other health-related conditions. In particular, the "smartness" of mobile tablets is useful for creating applications for ASD. For example, such applications could take advantage of the built-in camera, video and voice recorder features to capture a mother's instructions to her autistic child.

The focus of this paper is to discuss the design and the development of Autisay, a mobile application which can be used as a communication tool between autistic individuals and their caregivers, parents, teachers and therapists. The application takes advantage of a features-rich iPad tablet to create a social communication tool that has potential to be customised for different languages and cultural norms. It is also hoped that such a tool would be adopted with ease, due to the recent advances in mobile and ubiquitous technologies.

The paper is organised as follows. Section 2 specifies the background research and literature review in ASD. Section 3 describes related work in mobile applications for supporting ASD. Section 4 describes the design direction for the AutiSay application. Section 5 describes the overview of the AutiSay application including its features and system architecture. Section 6 describes AutiSay's future perspectives and Section 7 provides the conclusion.

## 2 Background Research

The term Autism Spectrum Disorder (ASD) is generally used to describe a group of complex disorders of brain development [4]. In the USA, one of the tools for ASD diagnosis is based on the guidelines of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) published in May 2013 [5].

The previous manual, DSM-IV recognised autism disorders as three distinct subtypes: autistic disorder, Asperger's disorder, or pervasive developmental disorder not otherwise specified (PDD-NOS). The new DSM-5 diagnostic manual have merged all the distinct autism disorders into one umbrella diagnosis, collectively referred to as Autism Spectrum Disorder (ASD). Furthermore it has also created a new diagnosis for Social Communication Difficulties (SCD). SCD describes individuals who have social communication difficulties without the repetitive behaviours or restrictive interests typical of autistic individuals [6]. It should be noted that SCD is not included in the ASD domain as it defines a group of individuals with related but separate symptoms. DSM-5 is a significant departure from DSM-IV as it has narrowed the autism disorders to two specific domains: social communication domain and restrictive interest and repetitive behaviours domain. Clearly the revised criteria under DSM-5 would have an impact on clinical adaptation of diagnostic practices and documentation to fit the criteria. This would also impact on the type of mobile applications developed for ASD, as developers would have to incorporate features to fit the criteria.

Recent statistics indicate that ASD is prevalent world-wide. According to the U.S.A Centers for Disease Control and Prevention (2010), global prevalence of autism has increased twentyfold to thirtyfold from studies conducted in the late 1960s to early 1970s [7]. It is unclear if the increase is due to a broader definition of ASD and/or better skills in diagnosis based on the criteria of DSM-IV, DSM-5 and other accepted diagnostic methods. However an increase in the number of individuals with ASD cannot be ruled out. Table 1 illustrates autism prevalence among some countries.

**Table 1.** Autism prevalence among some countries

| Country | Autism Rate (Children) | Data source |
|---|---|---|
| China | 1:909 | Peking University Health Science Centre (2005 estimate) [8]. |
| U.S.A. | 1:68 | U.S Centers for Disease Control and Prevention ADDM Study 2010 [7]. |
| South Korea | 1:38 | The American journal of psychiatry [9]. |
| Malaysia | 1:625 | Ministry of Health, Malaysia [10]. |
| United Kingdom | 1:263 (boys), 1:1250 (girls) | BMJ Open [11]. |
| India | 1:250 | Autism Society of India [12].<br>The Times of India [13]. |
| Brunei | 1:150 to 1:1000 | The Society for the Management of Autism-Related Issues - Training, Education and Resources (SMARTER), Brunei [14].<br>Child Development Center (CDC), Ministry of Education, Brunei Darussalam [15]. |

Children diagnosed with autism have in common an impediment in social interaction. This is characterised by social communication difficulties and restrictive and repetitive behaviours. However the severity of the impairments and behaviours differ respectively in each child. The DSM-5 has classified three levels of severity for diagnosis [6].

In terms of communication difficulties, children with autism tend to delay speaking and learning to use gestures. Some can experience significant language delays and will be nonverbal or nearly nonverbal. Even when they start to develop speech, they may have difficulty combining words into meaningful sentences [16]. There are several systems in place to help autistic individuals to learn to communicate such as PECS (Picture Exchange Communication System) [17]. Children using PECS are taught to use pictures to communicate their needs, to understand their environment and to learn new behaviours [18]. PECS is also incorporated in the TEACCH® Autism Program, a clinical, training, and research program based at the University of

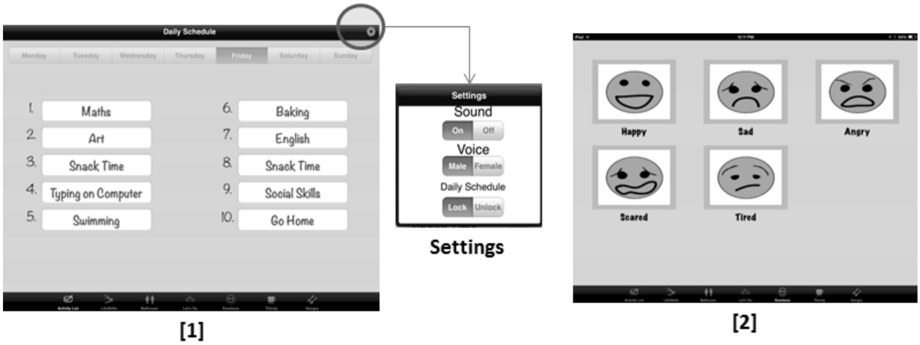
North Carolina at Chapel Hill. TEACCH® is an acronym for the "Treatment and Education of Autistic and Related Communication Handicapped Children".

TEACCH® was established as a state-wide program by the North Caroline Legislature in 1972 [19] and has become a model for other programs around the world [20]. Studies indicate that TEACCH® has some success as an effective intervention for autism [21][22]. The findings of how PECS is incorporated into successful programs such as TEACCH® is a crucial design consideration for the AutiSay application.

## 3 Related Work

It should be expected that the prevalence of autism worldwide and its increased awareness, would lead to a proliferation of autistic-related applications. Indeed, a search on the Apple app store shows an encouraging number of autistic-related applications in various categories, ranging from games, education, skills and communication. At the time of writing this paper, the "Autism Apps" autism locator tool for the iPad is able to locate approximately over 60 applications in the communication category (USA app store) [23]. Several of these applications are similar to the AutiSay application, as they are based on the PECS system and have the ability to incorporate sounds as well. For example, the PictureCanTalk application has some in-built pictures and ability to create personalised picture catalogue. Each picture may have a sound associated with it, created in one of two ways, (i) personalised recording or (ii) synthesized voice through the "Text to Speech" feature [24]. AutisMate is another example of an application that has incorporated PECS and in addition, is highly customizable to allow personal videos, audios and images. This customization has similarity to Autisay, however it is relatively expensive in comparison to AutiSay [25].

The EducateMe application is an early beta version of the AutiSay application. EducateMe was published on the Apple app store in July 2013 and available for free [26]. As of November 2013, it had over 500 downloads worldwide [27]. This beta version is no longer available for downloads. The design of the EducateMe application was influenced by the methodology employed by the SMARTER Autism Center in Brunei Darussalam (www.smarterbrunei.org). Hence its features include an "Activity List" where parents and autistic individuals can plan a seven-day activity list, simply by ticking the activity for the respective day. Other features include "Lifeskills" such as "Brushing Teeth". "Lifeskills" are presented in picture format. Touching a life-skill will bring the user to a process on how to perform the particular life skill step-by-step. Other features are "Bathroom", "Let's Go", "Emotions", "Thirsty" and "Hungry". These features are preloaded pictures that the autistic individual is familiar with and he/she can touch on the picture to show his/her needs at that time. If the sound feature is turned on, the touch will also trigger off a voice describing the picture. Some of the pictures preloaded in this application are courtesy of the SMARTER organisation [28]. Fig. 1 is a screen shot of the EducateMe application showing the main screen and the "Emotions" template that pops up when the "Emotions" item is selected at the bottom of the main screen. A child can touch one of these icons to express his feelings. A voice will also describe the feeling if the sound feature is turned on.



1. Maths
2. Art
3. Snack Time
4. Typing on Computer
5. Swimming
6. Baking
7. English
8. Snack Time
9. Social Skills
10. Go Home

Settings
Sound: On / Off
Voice: Male / Female
Daily Schedule: Lock / Unlock

**Fig. 1.** Main screen of the EducateMe application [1]. Selecting the Emotions menu at the bottom of the main screen opens up the Emotions template [2]. A child touches one of the faces to express his / her feelings.

## 4 Design Direction for the AutiSay Application

The beta version, the EducateMe application was the precursor and was tested at the SMARTER organisation. The test involved a trainer using the application to guide the activities of a child over the course of a day. It was interactive as the child used the application to communicate his needs to the trainer and the trainer used the application to show and guide the child on the step-by-step process for performing an activity. The technical aspects of the application were reviewed by peers at the Institut Teknologi Brunei (www.itb.edu.bn) and by industry practitioners working at the BAG Networks (Brunei Accenture Group), a leading IT consulting and outsourcing services provider in Brunei Darussalam (www.bagnetworks.com.bn) [29]. Based on the feedback of the test conducted at the SMARTER organisation and the technical review, the following were highlighted:

- The "Activity List" feature is text based and the consensus is that pictures would better serve the needs of the autistic child.
- For the other features, pictures were preloaded and are not customizable to individual requirements.
- Voices are preloaded and not customizable to individual requirements. Hence the application is only catering for one language, English. However the option to select a male or female voice is a plus.
- The placement of the features at the bottom of the screen is not as clear.
- Categories of the activities were mixed with the menu section. For example, Bathroom, a life skill was not placed in the Life Skills category and was placed as a menu item on its own.
- There was only one login mode and hence a child can easily reset the settings.
- The application is over 70MB in size, and this could be an issue during downloads.

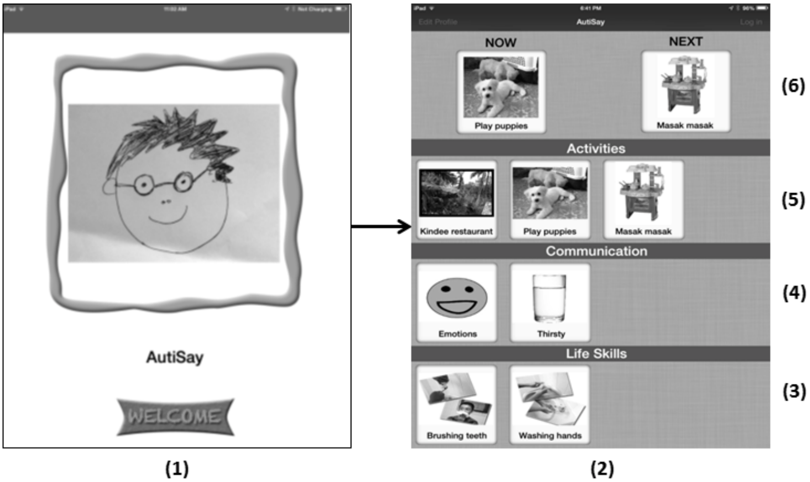
Taking the feedback into account, the AutiSay application was designed for an omnipresent user-friendly device to meet the requirements. Also, since the development team members were not experts in the field of autism, the design of the AutiSay application is informed by the review of the previous literature and the guidelines and norms of established practices for the treatment of autism for communicating, such as PECS. It is also influenced by the norms and practices adopted by the SMARTER organisation. In itself, SMARTER has met international standards and is a leading regional body providing world class quality service [30].

## 5 Description of the AutiSay Application

The proposed solution, AutiSay is developed for the autistic child who has difficulty in communicating verbally. Since the degree of severity differs from child to child, the system is designed to be flexible, and to easily capture context and contents relevant to the individual child. Hence, the application has a high degree of customizability, which is one of the most important features of the application.

### 5.1 Features Overview

The opening screen of the AutiSay application is a customised image of the individual child using the application. From the onset, this customization image allows the child to take ownership of the application. After the child presses the Welcome button, the application opens up onto the main screen; this displays a framework of three features that are customizable: (1) Life skills (2) Communication (3) Activities. Refer to Fig 2. for the opening and main screens of the AutiSay application.



(1) (2)

**Fig. 2.** The opening screen (1) allows for customisation of a picture of a child, therefore making the child take ownership of the application at the onset. Once the "Welcome" button is touched, the main screen (2) opens up to a selection of three features: Life Skills (3), Communication (4), and Activities (5). Selected activities by the child are shown in the Activity frame (6).

The following describes the features available on the AutiSay application:

- Life skills feature
  The purpose of the Life skills feature is to categorise a group of skills that are deemed necessary for the learner to acquire and perform on a day to day routine. Such skills would include brushing teeth and choosing and wearing clothes. Apart from using the feature to guide the autistic child to perform a prescribed life skill, new life skills can be added according to the development of the child. When adding a new skill, the Life skills feature is customizable to allow the caregiver (administrator) to define a life skill and to input the step-by-step task associated to performing that life skill. Each step is augmented with pictures, text and voice; all of which are customizable to individual requirements. Basically this means that picture and text are configurable with the individual's language and cultural norms and nuances.
- Communication feature
  The Communication feature is a category to group emotions (feelings) and needs. The autistic child can use the feature to express his/her feelings or to communicate what he/she needs at a moment in time. When the child touches on one of the icons, and if the voice feature is recorded, then a voice will verbally express that emotion or need. Apart from assisting the child to express emotions and/or needs, it can also be used by the child to learn to express new emotions or needs. For example, if a caregiver identifies an emotion such as "I am feeling peaceful", which is a relatively new expression for the child, the caregiver can define this emotion pictorially on the template and record the speaking voice of the emotion. The child can then use the feature to express this new emotion.
- Activities feature
  The Activities feature is special in the sense that it gives control to the individual child to communicate with the caregivers that he/she would like to do some activities that are not part of the daily routine. The application allow caregivers and parents to define activities on a list, hence, there is parental control on the type of "outside" activities decided by the child. This feature has potential for both parents and child to decide on a fun activity as a reward when the child shows the right behaviour. For example, if the child likes to play with the puppies, then this can be defined as an activity. Hence every time the child wants to play with the puppies, he/she can touch on a picture of the puppies and it becomes an activity-to-do next. The feature only allows two activities to be scheduled at any one time. The rationale for the limitation is that the child needs to learn discipline over the daily routines and to be rewarded accordingly. Credit for the incorporation of this feature is accorded to SMARTER, Brunei Darussalam (www.smarterbrunei.org).
- Login feature
  There are two login modes for the application (1) parental login and (2) child login. Customization for pictures, audio recordings and defining activities are only available at the parental login mode. The child login mode is to use the application as a communication tool and for step-by-step guidance for performing the life skills. The child option also allows the child to decide on the "outside" activities selected from a list which is customized by the parent or caregiver. This is the schedule for the "Now-Activity" and the "Next-Activity".

### 5.2 System Architecture Overview

The AutiSay application is developed for the hand-held smart device Apple iPad and requires iOS 6.1 or later. This device is equipped with a built-in camera, microphone and Wi-Fi capabilities which the application needs for customization and accessing the app store. The lowest range of the iPad would also have ample disk space to accommodate the images and audio recordings. The application is written in Objective-C, the native programming language for the Apple environment, and the IDE interface (development tool) used is Xcode.

Using the parental login mode, the AutiSay application uses the camera and audio recording facilities of the smart device to capture images and audios. As shown in Fig. 3, the captured images and audio recordings are stored in the SQLite database. SQLite is embedded and preinstalled in the IOS by Apple.

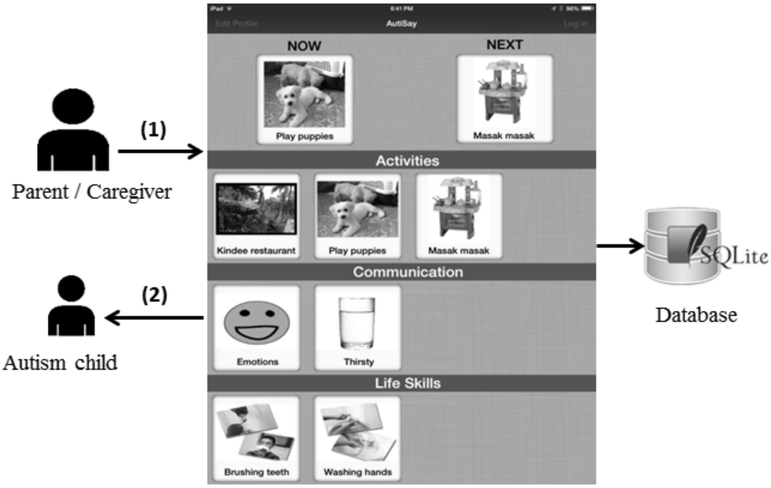


**Fig. 3.** Overview of the system architecture of AutiSay. The parent login mode is used to add, edit and amend the images and voices for the Life Skills, Communication and Activities features (1). These customised media are stored on the device using the SQLite database feature. The child uses the child login mode to communicate and learn (2).

Images are captured and stored in portable network graphics (png) format while sounds are in Apple's core audio format (caf). Captured images are captured in aspect fit, and scaled automatically by the application, to fit one of two image holders' dimensions during edit mode. For the bigger image holder, the image dimension is 430 by 445 pixels and the smaller image holder is 160 by 160 pixels. The bigger format is used in the step-by-step instructions under the Life Skills feature. All other features capture image with the smaller image holder dimension.

### 5.3 Benefits of the Application

The primary benefits of using the AutiSay application are that it is portable, highly customizable and relatively inexpensive. In terms of its portability, the application is available for a device that is small enough to be carried everywhere by the child and therefore can be used anytime and anywhere, as a communication tool to translate his/her needs to parents and caregivers. Furthermore, the application is designed to be highly customizable and therefore images and audios can be easily added and modified to meet individual requirements. This means that it allows for images and language nuances (spoken) that are unique to individual cultural norms, thereby giving the application a global reach. Finally, the application is relatively inexpensive as it is distributed on the Apple App Store for a relatively low price in comparison to similar applications.

## 6 Future Perspectives

Although the AutiSay application is focused on the autistic child, there is significant potential for the application to be scaled for adult autistic individuals. The images and audios can be replaced with contents that are more age-friendly for adults. For example, the Life Skills feature could include making a bank transaction or even taking a particular route to town. The possibility here is to have a version of the application oriented towards adult autistic individuals. Furthermore, the application can be easily configured to hold contents that are relevant for teaching and learning; thereby allowing caregivers and school educators to infuse contents that will guide the progress of the autistic child who is studying in the mainstream education system.

The application is currently available on the Apple app store only and the team is developing a version for the android environment, downloadable at the Google Play store.

## 7 Conclusions

AutiSay is designed as a tool for the autistic child to communicate with caregivers and vice versa. Much of the design is based on reviewed literature on the treatment for autism and also on the methodology employed by the SMARTER organisation which met international standards for providing services to autistic children. One of the most important features of the application is the easy customization of pictures and audio. This allows the application to be applicable for multiple languages and the nuances unique to respective cultures. While its test has been limited to the SMARTER organisation, it is hoped that the broader community would use the application, and continue to validate and extend it.

Apart from the application being used as a communication tool, it has potential to be used as a learning tool. Teachers can configure the application to be used for learning purposes, especially in creating the meaning of words or steps to perform an activity which are new to a child. There is potential for the application to be configured for adult individuals with autism. Furthermore, even though the application is designed for the Apple iPad, the team is developing an android version which will be distributed through the play store (Google Play).

**Acknowledgements.** The AutiSay application is the evolution of the EducateMe application, which started as a higher national diploma (HND) final year project. When the majority of the original student developers decided to further their degree education at the same university, they took on the challenge to develop the EducateMe application further through the mentorship of the main author. The result is the AutiSay application.

## References

1. Gartner Inc. Gartner Says Worldwide Traditional PC, Tablet, Ultramobile and Mobile Phone Shipments Are On Pace to Grow 6.9 Percent in 2014 (press release, 2014), `http://www.gartner.com/newsroom/id/2692318` (accessed July 5, 2014)
2. Perez, S.: iTunes App Store Now Has 1.2 Million Apps, Has Seen 75 Billion Downloads To Date (2014), `http://techcrunch.com/2014/06/02/itunes-app-store-now-has-1-2-million-apps-has-seen-75-billion-downloads-to-date/` (accessed July 5, 2014)
3. AppBrain. Number of Android applications (2014), `http://www.appbrain.com/stats/number-of-android-apps` (accessed July 5, 2014)
4. Autism Speaks Inc. What is Autism?, `http://www.autismspeaks.org/what-autism` (accessed July 19, 2014)
5. American Psychiatric Association, DSM-5 Overview: The Future Manual (2014), `http://goo.gl/cNzt7e` (accessed July 19, 2014)
6. Autism Speaks Inc. DSM-5 Diagnostic Criteria, `http://www.autismspeaks.org/what-autism/diagnosis/dsm-5-diagnostic-criteria` (accessed July 19, 2014)
7. Centers for Disease Control and Prevention. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States (2010), `http://www.cdc.gov/mmwr/preview/mmwrhtml/ss6302a1.htm?s_cid=ss6302a1_w` (accessed July 20, 2014)
8. Zhang, X., Ji, C.: Autism and Mental Retardation of Young Children in China. Biomedical and Environmental Sciences 18, 334–340 (2005)
9. Kim, Y.S., et al.: Prevalence of Autism Spectrum Disorders in a Total Population Sample. AM J. Psychiatry (2011), `http://www.autismeurope.org/files/files/new-study-on-autism-prevalence.pdf` (accessed July 20, 2014)
10. Malaysia Health Technology Assessment Section (MaHTAS). Draft Management of Autism Spectrum Disorder in Children and Adolescents. Malaysia:MaHTAS, Ministry of Development Division, Ministry of Health Malaysia (2014), `http://goo.gl/pu2Y4V` (accessed July 20, 2014)
11. Taylor, B., et al.: Prevalence and incidence rates of autism in the UK: timr trend from 2004-2010 in children aged 8 years. BMJ Open 3(10) (2013), `http://www.bmjopen.bmj.com/content/3/10/e003219.full` (accessed July 20, 2014)
12. Autism Society of India, What is Autism? (2014), `http://autismsocietyofindia.org/?page_id=1043` (accessed July 22, 2014)

13. TNN, Autism is not a disease. The Times of India (April 2, 2013), `http://timesofindia.indiatimes.com/city/ahmedabad/Autism-is-not-a-disease/articleshow/19333810.cms` (accessed July 27, 2014)
14. Marilyn, P.: Shining spotlight on autism. Borneo Bulletin (April 3, 2009), `http://www.brusearch.com/news/39817` (accessed July 22, 2014)
15. Uditha Rohana RatnayakeThien, R.: One autism case filed every week. The Brunei Times (June 20, 2014), `http://www.bt.com.bn/news-national/2014/06/30/one-autism-case-filed-every-week` (accessed July 22, 2014)
16. Autism Speaks®. Symptoms. What are symptoms of autism? (2014), `http://www.autismspeaks.org/what-autism/symptoms` (accessed July 22, 2014)
17. The National Professional Development Center on Autism Spectrum Disorders. Evidence Base for Picture Exchange Communication System (PECS) (2010), `http://goo.gl/Sfu4iE` (accessed September 8, 2014)
18. National Autism Resources. The Picture Exchange Communication System (2014), `http://goo.gl/rVI7XQ` (accessed July 24, 2014)
19. Mesibov, G.B.: Division TEACCH: A collaborative model program for service delivery, training, and Research for people with autism and related communication handicaps. Routledge, New York (1996) Roberts, M.C. ed.
20. Autism Speaks®, TEACCH Autism Program (2014), `http://www.autismspeaks.org/what-autism/treatment/teacch` (accessed July 24, 2014)
21. Ozonoff, S., Cathcart, K.: Effectiveness of a home program intervention for young children with autism. Journal of autism and developmental disorders 28(1), 25–32 (1998)
22. Panerai, S., et al.: Benefits of the Treatment and Education of Autistic and Communication Handicapped Children (TEACCH) programme as compared with a non-specific approach. Journal of Intellectual Disability Research 46(4), 318–327 (2002)
23. TouchAutism. Autism Apps (version 2.4) (Mobile app) (July 2014), Available from: Apple AppStore (accessed July 25, 2014)
24. App Consult Pte Ltd. PictureCanTalk (Mobile app.) (July 2014), Available from: Apple AppStore (accessed July 24, 2014)
25. SpecialNeedsWare. AutisMate (Mobile app) (2014), Available from: Apple AppStore (accessed September 9, 2014)
26. Institut Teknologi Brunei. EducateMe (Mobile app) (2013), Available from: Apple AppStore (accessed November 9, 2013)
27. Bazilah, D.S.N., Jumaat, H.B., Maidin, A.B., Voon, N.H.: AutiSay (PowerPoint slides). Presented at APICTA 2014 Competition (Tertiary) at Cyperport Hong Kong (2013)
28. Society for Management of Autism Related issues in Training, Education and Resources, SMARTER (2013), (In-house flashcards). Brunei Darussalam: SMARTER. Website: `http://www.smarterbrunei.org`
29. Lim, M., Harbison, M.: Conversation with Melvin Lim and Mark Harbison at the BAG Networks Sdn. Bhd.'s Head Office (26 July, 2013)
30. SMARTER Brunei, Meaning & International Standards (2009), `http://www.smarterbrunei.org/index.php?option=com_content&view=article&id=39&Itemid=200075` (accessed July 30, 2014)

# Author Index