

# **RECENT ADVANCES in ELECTRICAL ENGINEERING**

**Proceedings of the International Conference on Communications and  
Computers (CC 2015)**

**Proceedings of the International Conference on Circuits, Systems and  
Signal Processing (CSSP 2015)**

**Agios Nikolaos, Crete, Greece  
October 17-19, 2015**

# **RECENT ADVANCES in ELECTRICAL ENGINEERING**

**Proceedings of the International Conference on Communications and  
Computers (CC 2015)**

**Proceedings of the International Conference on Circuits, Systems and  
Signal Processing (CSSP 2015)**

**Agios Nikolaos, Crete, Greece  
October 17-19, 2015**

**Copyright © 2015, by the editors**

All the copyright of the present book belongs to the editors. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the editors.

All papers of the present volume were peer reviewed by no less than two independent reviewers. Acceptance was granted when both reviewers' recommendations were positive.

Series: Recent Advances in Electrical Engineering Series | 56

ISSN: 1790-5117

ISBN: 978-1-61804-351-1

# **RECENT ADVANCES in ELECTRICAL ENGINEERING**

**Proceedings of the International Conference on Communications and  
Computers (CC 2015)**

**Proceedings of the International Conference on Circuits, Systems and  
Signal Processing (CSSP 2015)**

**Agios Nikolaos, Crete, Greece  
October 17-19, 2015**



## Organizing Committee

### Editor:

Prof. Valeri Mladenov, Technical University of Sofia, Bulgaria

### Program Committee:

Prof. Lotfi Zadeh (IEEE Fellow, University of Berkeley, USA)  
Prof. Leon Chua (IEEE Fellow, University of Berkeley, USA)  
Prof. Michio Sugeno (RIKEN Brain Science Institute (RIKEN BSI), Japan)  
Prof. Demetri Terzopoulos (IEEE Fellow, ACM Fellow, UCLA, USA)  
Prof. Georgios B. Giannakis (IEEE Fellow, University of Minnesota, USA)  
Prof. George Vachtsevanos (Georgia Institute of Technology, USA)  
Prof. Abraham Bers (IEEE Fellow, MIT, USA)  
Prof. Brian Barsky (IEEE Fellow, University of Berkeley, USA)  
Prof. Aggelos Katsaggelos (IEEE Fellow, Northwestern University, USA)  
Prof. Josef Sifakis (Turing Award 2007, CNRS/Verimag, France)  
Prof. Hisashi Kobayashi (Princeton University, USA)  
Prof. Kinshuk (Fellow IEEE, Massey Univ. New Zealand),  
Prof. Leonid Kazovsky (Stanford University, USA)  
Prof. Narsingh Deo (IEEE Fellow, ACM Fellow, University of Central Florida, USA)  
Prof. Kamisetty Rao (Fellow IEEE, Univ. of Texas at Arlington, USA)  
Prof. Anastassios Venetsanopoulos (Fellow IEEE, University of Toronto, Canada)  
Prof. Steven Collicott (Purdue University, West Lafayette, IN, USA)  
Prof. Nikolaos Paragios (Ecole Centrale Paris, France)  
Prof. Nikolaos G. Bourbakis (IEEE Fellow, Wright State University, USA)  
Prof. Stamatios Kartalopoulos (IEEE Fellow, University of Oklahoma, USA)  
Prof. Irwin Sandberg (IEEE Fellow, University of Texas at Austin, USA),  
Prof. Michael Sebek (IEEE Fellow, Czech Technical University in Prague, Czech Republic)  
Prof. Hashem Akbari (University of California, Berkeley, USA)  
Prof. Yuriy S. Shmaliy, (IEEE Fellow, The University of Guanajuato, Mexico)  
Prof. Lei Xu (IEEE Fellow, Chinese University of Hong Kong, Hong Kong)  
Prof. Paul E. Dimotakis (California Institute of Technology Pasadena, USA)  
Prof. Martin Pelikan (UMSL, USA)  
Prof. Patrick Wang (MIT, USA)  
Prof. Wasfy B Mikhael (IEEE Fellow, University of Central Florida Orlando, USA)  
Prof. Sunil Das (IEEE Fellow, University of Ottawa, Canada)  
Prof. Panos Pardalos (University of Florida, USA)  
Prof. Nikolaos D. Katopodes (University of Michigan, USA)  
Prof. Bimal K. Bose (Life Fellow of IEEE, University of Tennessee, Knoxville, USA)  
Prof. Janusz Kacprzyk (IEEE Fellow, Polish Academy of Sciences, Poland)  
Prof. Sidney Burrus (IEEE Fellow, Rice University, USA)  
Prof. Biswa N. Datta (IEEE Fellow, Northern Illinois University, USA)  
Prof. Mihai Putinar (University of California at Santa Barbara, USA)  
Prof. Włodzisław Duch (Nicolaus Copernicus University, Poland)  
Prof. Tadeusz Kaczorek (IEEE Fellow, Warsaw University of Technology, Poland)  
Prof. Michael N. Katehakis (Rutgers, The State University of New Jersey, USA)  
Prof. Pan Agathoklis (Univ. of Victoria, Canada)  
Prof. Martin van den Toorn (Delft University of Technology, The Netherlands)  
Prof. Malcolm J. Crocker (Distinguished University Prof., Auburn University, USA)  
Prof. Urszula Ledzewicz (Southern Illinois University, USA)  
Prof. Dimitri Kazakos, Dean, (Texas Southern University, USA)  
Prof. Ronald Yager (Iona College, USA)  
Prof. Athanassios Manikas (Imperial College, London, UK)  
Prof. Keith L. Clark (Imperial College, London, UK)

Prof. Argyris Varonides (Univ. of Scranton, USA)  
 Prof. S. Furfari (Direction Generale Energie et Transports, Brussels, EU)  
 Prof. Constantin Udriste (University Politehnica of Bucharest, ROMANIA)  
 Prof. Patrice Brault (Univ. Paris-sud, France)  
 Prof. Jim Cunningham (Imperial College London, UK)  
 Prof. Philippe Ben-Abdallah (Ecole Polytechnique de l'Universite de Nantes, France)  
 Prof. Ichiro Hagiwara, (Tokyo Institute of Technology, Japan)  
 Prof. Andris Buikis (Latvian Academy of Science. Latvia)  
 Prof. Akshai Aggarwal (University of Windsor, Canada)  
 Prof. Ulrich Albrecht (Auburn University, USA)  
 Prof. Alexey L Sadovski (IEEE Fellow, Texas A&M University, USA)  
 Prof. Amedeo Andreotti (University of Naples, Italy)  
 Prof. Ryszard S. Choras (University of Technology and Life Sciences Bydgoszcz, Poland)  
 Prof. Remi Leandre (Universite de Bourgogne, Dijon, France)  
 Prof. Moustapha Diaby (University of Connecticut, USA)  
 Prof. Brian McCartin (New York University, USA)  
 Prof. Charles Long (Prof. Emeritus University of Wisconsin, USA)  
 Prof. Marvin Goldstein (NASA Glenn Research Center, USA)  
 Prof. Ron Goldman (Rice University, USA)  
 Prof. Milivoje M. Kostic (Northern Illinois University, USA)  
 Prof. Helmut Jaberg (University of Technology Graz, Austria)  
 Prof. Ardeshir Anjomani (The University of Texas at Arlington, USA)  
 Prof. Heinz Ulbrich (Technical University Munich, Germany)  
 Prof. Sesh Commuri (University of Oklahoma, USA)  
 Prof. Nicolas Galanis (Universite de Sherbrooke, Canada)  
 Prof. S. H. Sohrab (Northwestern University, USA)  
 Prof. Rui J. P. de Figueiredo (University of California, USA)  
 Prof. Valeri Mladenov (Technical University of Sofia, Bulgaria)  
 Prof. Hiroshi Sakaki (Meisei University, Tokyo, Japan)  
 Prof. Zoran S. Bojkovic (Technical University of Belgrade, Serbia)  
 Prof. Emira Maljevic (Technical University of Belgrade, Serbia)  
 Prof. Kazuhiko Tsuda (University of Tsukuba, Tokyo, Japan)  
 Prof. Milan Stork (University of West Bohemia , Czech Republic)  
 Prof. C. G. Helms (University of Athens, Greece)  
 Prof. Lajos Barna (Budapest University of Technology and Economics, Hungary)  
 Prof. Nobuoki Mano (Meisei University, Tokyo, Japan)  
 Prof. Nobuo Nakajima (The University of Electro-Communications, Tokyo, Japan)  
 Prof. Victor-Emil Neagoe (Polytechnic University of Bucharest, Romania)  
 Prof. P. Vanderstraeten (Brussels Institute for Environmental Management, Belgium)  
 Prof. Annaliese Bischoff (University of Massachusetts, Amherst, USA)  
 Prof. Virgil Tiponut (Politehnica University of Timisoara, Romania)  
 Prof. Andrei Kolyshkin (Riga Technical University, Latvia)  
 Prof. Fumiaki Imado (Shinshu University, Japan)  
 Prof. Yan Wu (Georgia Southern University, USA)  
 Prof. Daniel N. Riahi (University of Texas-Pan American, USA)  
 Prof. Alexander Grebennikov (Autonomous University of Puebla, Mexico)  
 Prof. Bennie F. L. Ward (Baylor University, TX, USA)  
 Prof. Guennadi A. Kouzaev (Norwegian University of Science and Technology, Norway)  
 Prof. Geoff Skinner (The University of Newcastle, Australia)  
 Prof. Hamido Fujita (Iwate Prefectural University(IPU), Japan)  
 Prof. Francesco Muzi (University of L'Aquila, Italy)  
 Prof. Claudio Rossi (University of Siena, Italy)  
 Prof. Sergey B. Leonov (Joint Institute for High Temperature Russian Academy of Science, Russia)  
 Prof. Lili He (San Jose State University, USA)

Prof. M. Nasseh Tabrizi (East Carolina University, USA)  
Prof. Alaa Eldin Fahmy (University Of Calgary, Canada)  
Prof. Gh. Pascovici (University of Koeln, Germany)  
Prof. Pier Paolo Delsanto (Politecnico of Torino, Italy)  
Prof. Radu Munteanu (Rector of the Technical University of Cluj-Napoca, Romania)  
Prof. Ioan Dumitrache (Politehnica University of Bucharest, Romania)  
Prof. Miquel Salgot (University of Barcelona, Spain)  
Prof. Amaury A. Caballero (Florida International University, USA)  
Prof. Maria I. Garcia-Planas (Universitat Politecnica de Catalunya, Spain)  
Prof. Petar Popivanov (Bulgarian Academy of Sciences, Bulgaria)  
Prof. Alexander Gegov (University of Portsmouth, UK)  
Prof. Lin Feng (Nanyang Technological University, Singapore)  
Prof. Colin Fyfe (University of the West of Scotland, UK)  
Prof. Zhaohui Luo (Univ of London, UK)  
Prof. Wolfgang Wenzel (Institute for Nanotechnology, Germany)  
Prof. Weilian Su (Naval Postgraduate School, USA)  
Prof. Phillip G. Bradford (The University of Alabama, USA)  
Prof. Ray Hefferlin (Southern Adventist University, TN, USA)  
Prof. Hamid Abachi (Monash University, Australia)  
Prof. Karlheinz Spindler (Fachhochschule Wiesbaden, Germany)  
Prof. Josef Boercsoek (Universitat Kassel, Germany)  
Prof. Eyad H. Abed (University of Maryland, Maryland, USA)  
Prof. F. Castanie (TeSA, Toulouse, France)  
Prof. Robert K. L. Gay (Nanyang Technological University, Singapore)  
Prof. Andrzej Ordys (Kingston University, UK)  
Prof. T Bott (The University of Birmingham, UK)  
Prof. T.-W. Lee (Arizona State University, AZ, USA)  
Prof. Le Yi Wang (Wayne State University, Detroit, USA)  
Prof. Suresh P. Sethi (University of Texas at Dallas, USA)  
Prof. Hartmut Hillmer (University of Kassel, Germany)  
Prof. Bram Van Putten (Wageningen University, The Netherlands)  
Prof. Alexander Iomin (Technion - Israel Institute of Technology, Israel)  
Prof. Roberto San Jose (Technical University of Madrid, Spain)  
Prof. Minvydas Ragulskis (Kaunas University of Technology, Lithuania)  
Prof. Arun Kulkarni (The University of Texas at Tyler, USA)  
Prof. Joydeep Mitra (New Mexico State University, USA)  
Prof. S. Y. Chen, (Zhejiang University of Technology, China and University of Hamburg, Germany)  
Prof. Duc Nguyen (Old Dominion University, Norfolk, USA)  
Prof. Tuan Pham (James Cook University, Townsville, Australia)  
Prof. Jiri Klima (Technical Faculty of CZU in Prague, Czech Republic)  
Prof. Rossella Cancelliere (University of Torino, Italy)  
Prof. Christian Bouquegneau (Faculty Polytechnique de Mons, Belgium)  
Prof. Wladyslaw Mielczarski (Technical University of Lodz, Poland)  
Prof. Ibrahim Hassan (Concordia University, Montreal, Quebec, Canada)  
Prof. James F. Frenzel (University of Idaho, USA)  
Prof. Vilem Srovnal, (Technical University of Ostrava, Czech Republic)  
Prof. J. M. Giron-Sierra (Universidad Complutense de Madrid, Spain)  
Prof. Walter Dosch (University of Luebeck, Germany)  
Prof. Rudolf Freund (Vienna University of Technology, Austria)  
Prof. Erich Schmidt (Vienna University of Technology, Austria)  
Prof. Alessandro Genco (University of Palermo, Italy)  
Prof. Martin Lopez Morales (Technical University of Monterey, Mexico)  
Prof. Vladimir Damgov (Bulgarian Academy of Sciences, Bulgaria)

## Additional Reviewers

Abelha Antonio	Universidade do Minho, Portugal
Angel F. Tenorio	Universidad Pablo de Olavide, Spain
Miguel Carriegos	Universidad de Leon, Spain
Alejandro Fuentes-Penna	Universidad Autónoma del Estado de Hidalgo, Mexico
Bazil Taha Ahmed	Universidad Autonoma de Madrid, Spain
M. Javed Khan	Tuskegee University, AL, USA
Moran Wang	Tsinghua University, China
Kazuhiko Natori	Toho University, Japan
Zhong-Jie Han	Tianjin University, China
Genqi Xu	Tianjin University, China
James Vance	The University of Virginia's College at Wise, VA, USA
Jose Flores	The University of South Dakota, SD, USA
Valeri Mladenov	Technical University of Sofia, Bulgaria
Minhui Yan	Shanghai Maritime University, China
Francesco Zirilli	Sapienza Universita di Roma, Italy
Tetsuya Shimamura	Saitama University, Japan
Andrey Dmitriev	Russian Academy of Sciences, Russia
Dmitrijs Serdjuks	Riga Technical University, Latvia
George Barreto	Pontificia Universidad Javeriana, Colombia
Francesco Rotondo	Polytechnic of Bari University, Italy
Masaji Tanaka	Okayama University of Science, Japan
Imre Rudas	Obuda University, Budapest, Hungary
Ole Christian Boe	Norwegian Military Academy, Norway
Stavros Ponis	National Technical University of Athens, Greece
Frederic Kuznik	National Institute of Applied Sciences, Lyon, France
Jon Burley	Michigan State University, MI, USA
Eleazar Jimenez Serrano	Kyushu University, Japan
Hessam Ghasemnejad	Kingston University London, UK
Konstantin Volkov	Kingston University London, UK
Takuya Yamano	Kanagawa University, Japan
João Bastos	Instituto Superior de Engenharia do Porto, Portugal
José Carlos Metrôlho	Instituto Politecnico de Castelo Branco, Portugal
Philippe Dondon	Institut polytechnique de Bordeaux, France
Xiang Bai	Huazhong University of Science and Technology, China
Tetsuya Yoshida	Hokkaido University, Japan
Shinji Osada	Gifu University School of Medicine, Japan
Kei Eguchi	Fukuoka Institute of Technology, Japan
Yamagishi Hiromitsu	Ehime University, Japan
Santoso Wibowo	CQ University, Australia
Sorinel Oprisan	College of Charleston, CA, USA
Deolinda Rasteiro	Coimbra Institute of Engineering, Portugal
Matthias Buyle	Artesis Hogeschool Antwerpen, Belgium



## Table of Contents

<b>A Design and Implementation of Vertical SNS for Student Backpackers</b>	11
<i>Kyeonhoon Kwak, Minwoo Yun, Yangwon Lim, Hankyu Lim</i>	
<b>The Usage of Neural Networks Paradigm in the Prediction of Protein Secondary Structure</b>	14
<i>Hanan Hendy, Wael Khalifa, Mohamed Roushdy, Abdel Badeeh Salem</i>	
<b>Inter-VM Performance Interference Aware Static VM Consolidation Algorithms for Cloud-Based Data Centers</b>	18
<i>Young-Chul Shim</i>	
<b>A Study on the Intelligent Techniques of the ECG-Based Biometric Systems</b>	26
<i>M. Bassiouni, W.Khalefa, El-Sayed A. El-Dahshan, Abdel-Badeeh M. Salem</i>	
<b>Control Flow Hardening with Program Counter Encoding for ARM® Processor Architecture</b>	32
<i>Seho Park, Jongmin Lee, Yongsuk Lee, Gyungho Lee</i>	
<b>Wide-Area Damping Controller Design Aiming at Robust Performance of the Controller</b>	39
<i>M. Beiraghi, A. M. Ranjbar</i>	
<b>Developing a Web-Based Ontology for Expert Decision Support Systems</b>	47
<i>Saud Ben Mahmoud Mandourah, Abdel-Badeeh M. Salem</i>	
<b>Design of Smart Bridge Based on WSN for Efficient Measuring of Temperature, Strain and Humidity</b>	53
<i>Mohamed Bouyahi, Houria Rezig, Tahar Ezzedine</i>	
<b>OVA Heuristic Approach Based on Variable Neighborhood Search Method for Multiclass Support Vector Machine</b>	58
<i>B. Sidaoui, K. Sadouni</i>	
<b>An Image Fusion Algorithm Based on Wavelet Transform and Tensor Analysis</b>	65
<i>Yongping Zhang, Dechun Zheng, Zhongkun He</i>	
<b>Discrete Sliding Mode Control with Predictive Sliding function for Multivariable Systems</b>	69
<i>Houda Ben Mansour, Khadija Dehri, Ahmed Said Nouri</i>	
<b>Benign or Malignant Lesion Classification in Mammography Images Using the Adaptive Orthogonal Transformation and the Coefficients of the Correlation Matrix</b>	76
<i>Khalid El Fahssi, Abdelali Elmoufidi, Abdenbi Abenaou, Said Jai-Andaloussi, Abderrahim Sekkaki</i>	
<b>Modeling of Solar-Diesel Hybrid Central Heating System</b>	81
<i>Musa Abdalla</i>	

<b>Psychological Aspects of Life Values Preferences in the Context of Various University Environment</b>	88
<i>Věra Strnadová, Petr Voborník, Kateřina Provazníková</i>	
<b>Authors Index</b>	98

# A Design and Implementation of Vertical SNS for Student Backpackers

Kyeonhoon Kwak, Minwoo Yun\*, Yangwon Lim and Hankyu Lim(corresponding author)

Department of Multimedia Engineering, Department of Computer Engineering\*

Andong National University

Andong-si, Republic of Korea

rusgns123@naver.com, ymw9801@naver.com, limyw@anu.ac.kr, hklm@anu.ac.kr

**Abstract**—Following widespread of Smartphone, SNS that shares experience and information of each individual is rapidly growing. In particular, vertical SNS which shares a certain type of information emerged. This paper designed and materialized vertical SNS for college student backpackers. The application provides travelers with an opportunity to communicate as well as a function that can compute travel expenses, while following the user location. Web application was materialized using HTML5, while footprint function was materialized using Geo location which is one of the HTML5 function. When application is materialized using HTML5, materialization is possible regardless of the types of platform. This study is expected to contribute to the provision and sharing of diverse information that are helpful to college student backpackers.

**Keywords**—component; SNS, Vertical SNS, HTML5, backpacker

## I. INTRODUCTION

Gathered in public SNS such as Twitter, Facebook and Google plus, we upload our emotions and information in a form of text, image or videos and share them with others [1, 2, 3]. With abundant information in diverse fields, a huge amount of traffic is generated. In these days, vertical SNS such as Instagram and Pinterest that induce large traffic by focusing on a specific topic and style began to emerge [4]. Since their approach is focused on a specific issue such as hobby, culture and tourism, users can share more profound information. In particular, restaurant and accommodation information related to tourism are provided through diverse types of service. However, considering that the share of information is done in text, pictures and videos, more specific form of information regarding travel will be helpful. In this paper, a vertical SNS for college student backpackers was designed and materialized. In addition to sharing information based on previous vertical SNS, the new vertical SNS supports sharing of users' own travel records and other people's travel records travel using GPS information and it also supports the formation of community for finding travel companions. Reactive web technology of HTML5 and GPS API were used for the materialization which can be used in diverse Smart devices.

This work was supported by a grant from 2015 Seoul Accord project(R0613-15-1148) of MISP(Ministry of Science, ICT and Future Planning) and IITP(Institute for Information and Communication Technology Promotion).

## II. TYPE STYLE AND FONTS

### A. HTML5

HTML5 is a web programming language that was confirmed as a next-generation web standard. HTML5 expands previous HTML that was capable of displaying text and hyperlinks only such that it can present diverse applications including multimedia to provide them to users. As it can provide diverse functions such as audio, video, graphic processing and location information, the functions that can be processed on web have remarkably enhanced [5]. HTML5 is a technology aimed at web application rather than web contents. In particular, diverse functions are standardized and provided for rich interaction regardless of the platform. Moreover, realistic web application can be programmed thanks to the additional API function.

In this paper, web application was materialized using HTML5, while footprint function was materialized using Geo location which is one of the HTML5 functions. When application is materialized using HTML5, materialization is possible regardless of the types of platform.

### B. GPS(Global Positioning System)

GPS (Global Positioning System) is a satellite navigation system that computes the current location of the user by receiving signal from GPS satellite. It is largely used in navigation devices of aircrafts, ships, and automobiles and it is also being used in Smartphone, tablet PC, etc. these days. Here, the coordinate is obtained by calculating the distance between the GPS satellite and GPS receptor. If we have accurate location and distance of GPS satellite, three GPS satellites are sufficient to find a precise location [6].

In this paper, location information was easily materialized using Open API of Daum Map by making use of GPS technology [7]. Convenience and usability of the application can further be enhanced when the travelers share their own location and check the footprint of other travelers to add travel information and entertaining function to the application

### C. Questionnaire Survey Analysis

Figure 1 below is an excerpt from a survey result that was conducted offline on 64 college students.

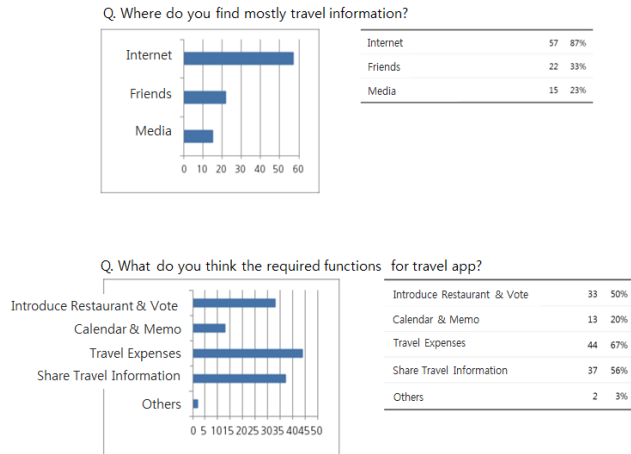


Fig. 1. Questionnaire Survey

According to the survey results, 35 respondents (53%) complained about insufficient travel information when going on a trip and 44 respondents (67%) pointed out the additional expenses. Moreover, 44 respondents (67%) said they wanted an additional function of travel expense calculation, while 37 (56%) hoped for an additional function of travel information share. Most of the college students showed strong interest in travel expenses and they also wanted shared information about the places where others visited or are planning to visit. This paper applied these results to the functional materialization of vertical SNS.

### III. DESIGN AND IMPLEMENTATION

#### A. Structured Design

Figure 2 shows a flowchart where there is a DB for storing and retrieving the travel information and user location and the functions of providing users with information are defined.

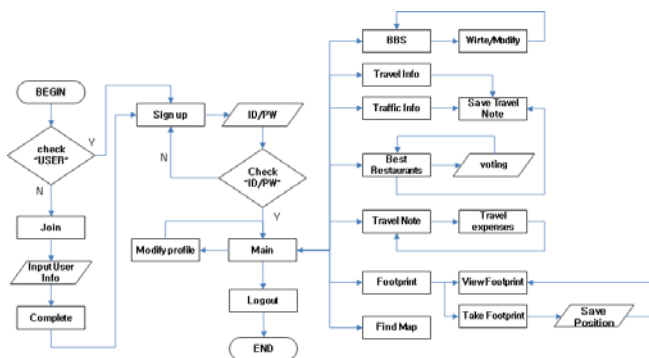
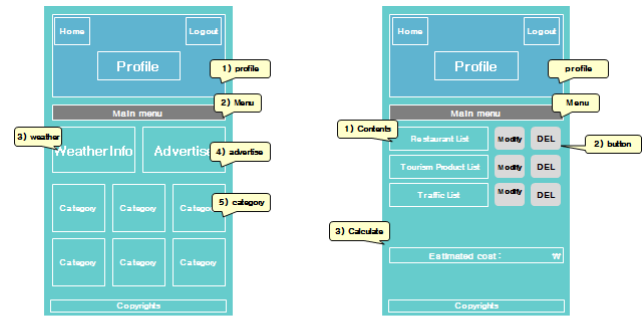


Fig. 2. Flowchart of vertical SNS

#### B. UI Design

Figure 3a below is a screenshot after login. There is a profile and background screen that can be changed by users and the main menus of the application are listed in a single row. There are six image buttons that enable easy movement

of the core functions of the application. Figure 3b shows the travel note page. The menus selected by the traveler are listed and the total expenditure is calculated at the bottom.



a. Main Page

b. Personal Page

Fig. 3. UI Design (Main Page and Persnal Page)

#### C. Implementation

In this paper, three functions including member registration and login, travel expenditure calculation function and footprint shot function were realized.

Figure 4 below shows the login and member registration page. The login page has a members-only area in the title part so that it can induce users' member registration. The member registration page uses minimum amount of contents and textbox to lessen the pressure of member registration.

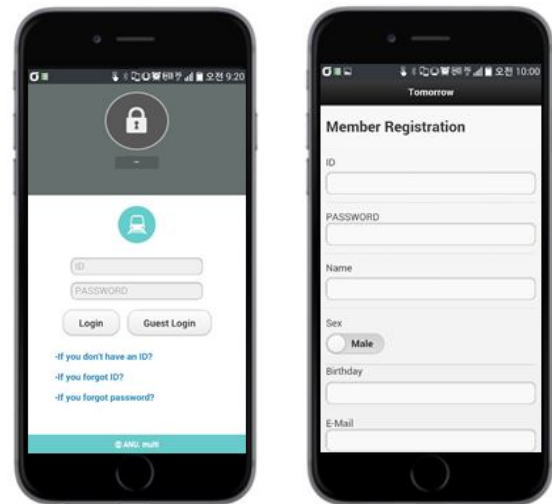


Fig. 4. login and member registration

Figure 5 below is a screenshot of famous restaurant bulletin board and user note of each location. In bulletin board, the users can look for the menu, prices and location of famous restaurants and they can select menu and contain them in their note. Information about the food such as pictures and brief introduction are given to the users who can give scores to the food. Moreover, they can check the foods and contain them in their travel note. Each user can have one

user note and it shows the items selected in the bulletin board. The users can check the menu they chose which can be revised or deleted. A total calculation fee is expressed at the bottom.



Fig. 5. Famous restaurant bulletin board and user note

Figure 6 below is a screenshot after running footprint function. In the screen, yellow marker shows the location checked by the application users, while blue marker indicates the current location of the user. The users can express their own location by clicking footprint store button. Footprint guarantees anonymity and the id information is not put into the DB.

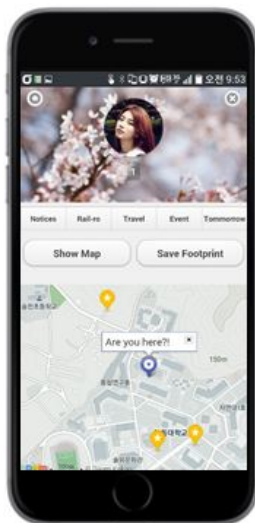


Fig. 6. Footprint function

#### IV. CONCLUSION

In this paper, vertical SNS service was designed and materialized for college student backpackers. The vertical SNS for sharing travel information was designed such that not only the share of basic travel information about backpacking, but also the real-time communication about the

travel expenses, visited spots and accompanying destination are possible. The users can search for other backpackers near them using GPS information and the system is also designed such that the users can make use of the information by the previous backpackers. Moreover, HTML5 technology was used for the materialization that can respond to diverse Smart devices.

Following recent spread of diverse Smartphone, people want to share information anywhere and at anytime. By facilitating communication of experience and information among the college student backpackers, this paper is expected to help the college students who are hesitant in going off on a journey so that they can grow out of fear for travel and hit the road.

#### REFERENCES

- [1] Richter, Alexander, and Michael Koch. "Functions of social networking services." Proc. Intl. Conf. on the Design of Cooperative Systems. pp.87-98, 2008.
- [2] Daegeun Kwon, ..., Yangwon Lim and Hankyu Lim, "Action Simulation Game Design using the Leap Motion Controller", 12<sup>th</sup> International Conference on Knowledge, Economy and Management Proceeding, pp381-391, Nov. 2014.
- [3] Hyeonpyo Hong, and Hankyu Lim, "Application of Smart Device Usage Control adopting Secure Coding for Hacking Protection", NNGT International Journal on Internet and Web Application, Vol. 2, pp1-8, Feb. 2015.
- [4] Changgu Kim, Junsu Jeong, Munsik Choo, Seokgi Hong, Chunghyun Lee, Yangwon Lim and Hankyu Lim, "Design and Implementation of Vertical SNS in the Big Data era", Proceeding of JCCIS Conference, pp109- 112, Nov. 2012.
- [5] <http://www.w3.org/TR/html5/>
- [6] [https://en.wikipedia.org/wiki/Global\\_Positioning\\_System](https://en.wikipedia.org/wiki/Global_Positioning_System)
- [7] <http://apis.map.daum.net/>

# The usage of Neural Networks Paradigm in the prediction of protein secondary structure

Hanan Hendy, Wael Khalifa, Mohamed Roushdy, Abdel Badeeh Salem

[hanan.hendy@cis.asu.edu.eg](mailto:hanan.hendy@cis.asu.edu.eg), [wael.khalifa@cis.asu.edu.eg](mailto:wael.khalifa@cis.asu.edu.eg), [mroushdy@cis.asu.edu.eg](mailto:mroushdy@cis.asu.edu.eg),  
[absalem@cis.asu.edu.eg](mailto:absalem@cis.asu.edu.eg)

Computer Science Department, Faculty of Computer and Information Sciences,  
Ain-Shams University, Cairo, Egypt

**Abstract**—The process of predicting the secondary structure of protein is crucial in understanding the protein functionality. It is very important in understanding the protein functions and in diagnosing any disorder. This paper shows the usage of neural networks paradigm in the prediction process. It discusses the difference between applying feed forward and back propagation algorithms on the prediction accuracy. The results show that the highest accuracy is reached when presenting the primary sequence in binary format and use Feed-Forward network. The accuracy reached around 86% when predicting Beta strands or sheets only. Moreover, it was seen that predicting alpha and beta each alone and combining the results didn't show significant accuracy enhancement.

**Keywords**—Artificial Neural Networks, Bioinformatics, Machine Learning, Protein Secondary Structure Prediction.

## I. INTRODUCTION

Protein secondary and tertiary structures prediction is a very important process this is because secondary structure prediction helps in finding protein disorder and figure out diseases. Moreover, accurate prediction of secondary structure helps in predicting the tertiary structure which is important and difficult as well. Protein forms its secondary structure when the amino acids (primary basic elements) bond together forming one of three shapes. The secondary structure can be alpha helix, beta sheets/strands or loops. The goal of the secondary structure prediction is to know the bonds that the amino acids form, which will help scientist take the correct actions in case of finding any disorder.

Neural networks [1] were used as the machine learning technique for the secondary structure prediction process. The first use was by Qian and Terrence [2]. They worked on a network with 17 input groups having 21 units per group, 40 hidden units and three output units. They reached accuracy of about 64.3%. Later, Chandonia et al [3] used some additional info in the prediction process that helped them reach an accuracy of 73.9% for class prediction. Then a more advanced neural networks were then proposed by Pollastri et al [4]. They used bidirectional recurrent neural network and their accuracy reached 78%. Recently a combined method was introduced by Y. Wei et al [5] in which the prediction is accomplished using

the value from multiple predictors, these values are then combined to find out the likelihood of the amino acid sequence. Its accuracy reached 83.04%.

In this paper, section one presents a brief introduction about the usage of neural networks in the protein structures prediction. Section II shows the data pre-processing in order to be ready for the learning process. Later, section III shows the implementation of the neural network and the results obtained by applying different neural networks hierarchies showing the best accuracy. Finally, the conclusion is presented showing what can be done in the future to enhance the accuracy.

## II. DATA PREPROCESSING

This paper is based on a data set extracted from the PDB - Protein Data Bank-[6]. Specifically the data set used is the CB513 [7] which contains 513 file each representing a single protein sequence. Each file contains data about a single protein sequence. Converting separate files having protein structures to data ready for machine learning process (matrix like form) is a hard task. To start any machine learning algorithm you need to have a specific input and output. Preprocessing steps are shown briefly in fig. 1 then will be explained in detail. Later a simple example will be explained.



Fig. 1 - Block diagram for protein data preparation

### A. Dealing with raw files

The raw file is represented in FASTA format [8] which is a text-based format for representing either nucleotide sequences or peptide sequences, in which nucleotides or amino acids are represented using single-letter codes. The file contains multiple data but only the following is needed:

Table 1 - Data of protein raw

RES	Represents the protein residue (primary structure).
DSSP	Represents the secondary structure from the DSSP database.

### B. Primary and secondary structures separation:

The first step done on these files is collecting all the primary structures and the secondary structures together in two separate files. Then all the commas that separates the characters of the sequence are removed and replace the in the secondary structures with a coil (C).

### C. Encoding primary and secondary structures:

The next step after having the primary and secondary structures separated is to encode them. The encoding step includes assigning a numeric value for each letter of the input. First the primary structure is represented in one of 20 main amino acids. The ambiguous amino acids like Asx (B), Glx (Z) [9] are dealt with as a single value. So the encoding of primary structure and secondary structures are shown in table 2 and 3 respectively. After finishing this step, we have two files one having all encoded primary structures and the other having all encoded secondary structures as explained above.

Table 2 - Encoding of protein primary structure

Amino Acid	Letter Code	Encoding Value
Alanine	A	01
Cysteine	C	02
Aspartic acid	D	03
Glutamic acid	E	04
Phenylalanine	F	05
Glycine	G	06
Histidine	H	07
Isoleucine	I	08
Lysine	K	09
Leucine	L	10
Methionine	M	11
Asparagine	N	12
Proline	P	13
Glutamine	Q	14
Arginine	R	15
Serine	S	16
Threonine	T	17
Valine	V	18
Tryptophan	W	19
Tyrosine	Y	20
Asparagine	B	21
Glutamine	Z	
Leucine	J	
Unspecified	X	

Table 3 - Encoding of protein secondary structure

Secondary Structure	Letter Code	Encoding Value
Alpha Helix (3-turn helix)	G	01
Alpha Helix (4-turn helix)	H	
Alpha Helix (5-turn helix)	I	

Beta Sheet	E	02
Others (bridge)	B	03
Others (bend)	S	
Others (hydrogen bond)	T	
Others (coil)	C	

### D. Prepare input and output for machine learning phase

Having encoded values for input and output which is primary and secondary structures respectively is not enough to start the machine learning process. The problem needs to be more specific and aligned. Since the sequences are not of equal length and the output structure sequence may have the three secondary structures (alpha helix, beta sheet and coils), a generalization is needed in which the problem is seen as inputting a sequence and deciding whether it's alpha helix, beta sheet or coil. How will this be done? By choosing a number which will specify the number of amino acids from the primary sequence that will be considered a single input. This number is a constant through the learning and testing phases. The corresponding output will be a number representing alpha helix, beta sheet and coil and it will be the one at the mid index of the chosen constant. For example, we'll choose the number 7 which means that the first 7 amino acids will be the input and the output will be the secondary structure mapped to index 4 (as if the amino acid number 4 is being checked when it's in a sequence of 7 amino acids).

This step will be clearer later when explained with an example. In this stage, all primary sequences are combined together with "00" separator also the same is done for the secondary sequences. Then the primary sequence is divided with the value chosen before (i.e. 7). So the first seven are taken then the first amino acids are dropped to take the next and so on as shown below. When a sub sequence is found to have "00" in the mid place (i.e. 4<sup>th</sup> place), the sequence is discarded.

15170302200612181215080317170601160209170109130

15170302200612181215080317170601160209170109130

15170302200612181215080317170601160209170109130

Then the output for each is taken by skipping the first 3 numbers as the first to have a secondary output will be the 4<sup>th</sup>. Fig. 2 represents an example of the steps explained above.

Raw File	P: RES:A,P,A,F,S,V,S,P,A,S,G,A,S,D,G,Q,S,V,S,V P: RES:T,P,A,F,N,K,P,K,V,E,L,H,V,H S: DSSP: , ,E,E,E,E,E, , ,S,S, , ,S,S, ,E,E,E,E S: DSSP: , ,S, , ,S, ,E,E,E,E,E,E,E
Remove commas	P: APAFSVSPASGASDGQSVSV TPAFNKPKVELHVH S: CCEEEEECCSSCCSSCEEEEE CCSCCSCEEEEEEE
Encode	P: 0113010516181613011606011603061416181618 1713010512091309180410071807 S: 030302020202020303030303030303020202 03030303030303020202020202
Combine with 00	P: 0113010516181613011606011603061416181618 001713010512091309180410071807 S: 030302020202020303030303030303020202 0003030303030303020202020202
Split by odd number to make the input and output matrices In this example, 9 will be used. The matrices are column based	
Input Matrix	
01 13 01 05 16 18 16 13 01 16 06 01 16 03 06 14 16 18 18 00 17 13 01 05 12 09 13 01 05 16 18 16 13 01 16 06 01 16 03 06 14 16 18 18 00 17 13 01 05 12 09 13 01 05 16 18 16 13 01 16 06 01 16 03 06 14 16 18 18 00 17 13 01 05 12 09 13 09 05 16 18 16 13 01 16 06 01 16 03 06 14 16 18 18 00 17 13 01 05 12 09 13 09 18 <b>16 18 16 13 01 16 06 01 16 03 06 14 16 18 18 17 13 01 05 12 09 13 09 18 04</b> 18 16 13 01 16 06 01 16 03 06 14 16 18 18 00 13 01 05 12 09 13 09 18 04 10 16 13 01 16 06 01 16 03 06 14 16 18 18 00 17 01 05 12 09 13 09 18 04 10 07 13 01 16 06 01 16 03 06 14 16 18 18 00 17 13 05 12 09 13 09 18 04 10 07 18 01 16 06 01 16 03 06 14 16 18 18 00 17 13 01 12 09 13 09 18 04 10 07 18 07	
Output Matrix	
00000000000000000000000000000000 11100000000011110000000111 000111111111000011111111000	

Fig. 2 - Real example for data pre processing  
P and S refers to primary and secondary structure files respectively.

### III. IMPLEMENTATION AND RESULTS

Artificial neural network is chosen to be the machine learning tool used to predict the protein secondary structure. *MTLAB R2013a* neural network tool box is used to implement different neural networks and use different architectures. Implementation varies according to specific commuting parameters as shown in table 4. This resulted in 128 different trial.

Table 4 - Implementation computation parameters

Computing Parameter	Different Values
Neural Network Architecture	Feed Forward Back propagation.
Hierarchy	Single hidden layer with different number of neurons.  Multiple hidden layer with different number of neurons
Number of input neurons	The input file that represents the primary structures was split to form the input matrix using different numbers that vary from 17 to 31. This decides the number of amino acids that will be given to the neural network at once}
Input format	Integer input: such that amino acids are encoded as numbers from 1 to 21  Binary input: such that amino acids are encoded as binary values (0 and 1). Each amino acid is encoded to a string of 0's and 1 at the index of the amino acid. For example: sequence 0307 will be encoded to 0010000000000000000000 0000001000000000000000
Prediction Output	Predict alpha and beta together Predict alpha only Predict beta only

Results will be discussed below showing the most and least error rates. Moreover, comparing results among different architectures is shown. It is clearly seen that the highest error rate results when predicting both alpha and beta together using input in number format ranging from 50% to 56% with accuracy about 50% as shown in table 5. While the lowest error rate results when predicting Beta structures alone and representing the input in binary format. Error rate in this case ranged from 14% to 18% with accuracy about 85% as shown in table 6. Table 7 shows the average performance when predicting both alpha and beta using binary format. Error rate average is 37% with accuracy 63%. These results exceeds the accuracy of John et al [10] whose results had an accuracy ranging from 62.3% to 73.9%. Also exceeds those of Qian and Terrence [2], Chandonia et al [3] and Pollastri et al [4].



Table 5 - Accuracy when predicting alpha and beta using binary format (Worst accuracy)

	Back propagation network				Feed forward network			
Number of input nodes	17	19	25	31	17	19	25	31
Number of hidden layers								
10	45.04	45.81	45.39	45.27	47.76	48.99	49.00	47.57
3	44.38	45.14	45.49	45.30	46.16	46.58	46.24	46.37
10 3	44.93	45.05	44.12	45.54	48.18	49.04	49.06	47.59
3 10	45.17	45.37	45.23	45.26	48.59	46.45	46.57	46.43

Table 6 - Accuracy when predicting beta using binary format (Best accuracy)

	Back propagation network				Feed forward network			
Number of input nodes	17	19	25	31	17	19	25	31
Number of hidden layers								
10	81.86	81.68	81.99	78.72	83.08	83.35	83.73	82.7
3	81.79	81.57	82.08	81.01	82.36	82.10	82.57	82.36
10 3	81.84	80.82	81.76	81.95	85.07	83.69	85.56	83.44
3 10	81.97	81.91	82.09	81.88	82.09	82.06	82.17	83.03

Table 7 - Accuracy when predicting alpha and beta using binary format (average accuracy)

	Back propagation network				Feed forward network			
Number of input nodes	17	19	25	31	17	19	25	31
Number of hidden layers								
10	63.70	63.81	64.14	64.01	65.8	64.62	65.75	65.82
3	58.28	58.94	63.49	58.54	64.09	59.34	64.4	59.45
10 3	58.89	63.54	63.51	63.54	65.8	65.57	67.84	66.86
3 10	44.12	62.96	58.80	58.78	64.17	60.21	64.66	64.53

#### IV. CONCLUSION AND FUTURE WORK

As discussed all through the paper, artificial neural networks were used to predict the protein secondary structure. The method showed varying accuracy. It is clearly seen that the method reached an accuracy of 86% which is considered significant with the use of Neural Network. This increase in the accuracy is due to the following:

- Unifying the length of the input while having different protein primary structure lengths.
- Encoding the input and the output in binary format. This made the network either take the whole component or discard it which is different that the case of numeric encoding.

Having reached 86% accuracy is not enough. Further enhancements can be done to increase the accuracy. For example, we can combine the results from more than one

network and take the highest weight or use other encoding techniques. Moreover, the trend in the prediction, is to use more than one machine learning technique and combine the results. By doing this, a better accuracy can be reached that will help later on in the tertiary structure prediction.

#### REFERENCES

- [1] S. Haykin, Neural Networks and Learning Machines, Ch. 1. Pearson International Edition, 2009.
- [2] Ning Qian and T. J. Sejnowski, "Predicting the secondary structure of globular proteins using neural network models," Journal of Molecular Biology, vol. 202, no. 4, p. 865–884, August 1988.
- [3] John-Marc Chandonia and Martin Karplus, "Neural networks for secondary structure and structural class predictions," Protein Science, vol. 4, no. 2, p. 275–285, February 1995.
- [4] Gianluca Pollastri, Darisz Przybylski, Burkhard Rost and Pierre Baldi, "Improving the Prediction of Protein Secondary Structure in Three and Eight Classes Using Recurrent Neural Networks and Profiles," PubMed, pp. 228-235, May 2002.
- [5] Y. Wei, J. Thompson and C. A. Floudas, "CONCORD: a consensus method for protein secondary structure prediction via mixed integer linear optimization," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science, pp. 831-850, November 2011.
- [6] Rcsb protein data bank." <http://www.rcsb.org/pdb/home/home.do>. [Accessed: July - 2015].
- [7] "Cuff and barton data set." <http://comp.chem.nottingham.ac.uk/disspred/datasets/CB513>[Accessed: July - 2015].
- [8] Fasta format." [https://en.wikipedia.org/wiki/FASTA\\_format](https://en.wikipedia.org/wiki/FASTA_format). [Accessed: June - 2015].
- [9] Amino acids." [https://en.wikipedia.org/wiki/Amino\\_acid](https://en.wikipedia.org/wiki/Amino_acid). [Accessed: May - 2015].
- [10] J.-M. Chandonia and M. Karplus, \Neural networks for secondary structure and structural class prediction," Cambridge university press, 1995.

# Inter-VM Performance Interference Aware Static VM Consolidation Algorithms for Cloud-Based Data Centers

Young-Chul Shim

**Abstract**—Energy efficiency in data centers is a very important issue and getting growing attention from researchers. One approach to reduce energy consumption is to allocate virtual machines (VMs) to physical machines (PMs) in such a way that the number of idle PMs is maximized. Approaches of this kind are called static VM consolidation methods. Idle PMs can be put into an energy-saving sleep mode, in which PMs consume significantly lower energy than in the normal operation mode. But if too many VMs are packed into a single PM, the performance interference among VMs can cause significant slowdown to jobs. In this paper we introduce new static VM consolidation algorithms which noticeably reduce energy consumption but do not incur too much performance degradation on jobs due to inter-VM interference. We analyze proposed algorithms through simulation

**Keywords**—Cloud-based data centers, Energy-efficiency, Inter-VM performance interference, SLA violation, Static VM consolidation algorithm.

## I. INTRODUCTION

SINCE the emergence of the concept of cloud computing, it is getting more widely adopted and deployed in the IT industry sector and receiving more attention from computer scientists and engineers. NIST defines cloud computing to be a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [1]. Examples of well-known cloud computing systems include Amazon EC2, Microsoft Azure, Google AppEngine, and Rackspace Cloudservers [2],[3].

Virtualization is an essential mechanism in providing the computing-as-a-service vision of cloud-based data centers. By providing physical resource sharing, fault isolation, security isolation, and live migration, virtualization allows diverse applications to run in isolated environments through creating multiple virtual machines (VMs) on shared hardware platforms [4]. When a user rents VMs to run an application on them, he specifies the amount of resources allocated to each VM by

stating required CPU cycles, memory, and I/O bandwidth. VM monitor (VMM) or hypervisor manages and multiplexes access to the physical resources, maintaining isolation between VMs at all times. As the physical resources are virtualized, several VMs, each of which is self-contained with its own operating system, can execute on a physical machine (PM) [5].

Scheduling in cloud-based data centers is a problem of mapping VMs to PMs. Of course, the mapping should be done in such a way that the resources needed by all VMs running on a PM should not exceed the capacity of that PM so that the PM may not be overloaded and the performance of its VMs may not be degraded.

Another important issue in the scheduling problem is the minimization of the energy consumed in the cloud-based data center. The mapping should be performed in such a way that the total amount of energy consumed by all the PMs running all the VMs should be minimized. It is reported that worldwide, the data centers use about 30 billion kilowatts of electricity, roughly equivalent to the output of 30 nuclear power plants [6] and this number is expected to grow 12% a year. This excessive use of energy in data centers not only raises the operation cost of data centers heavily but also creates environmental issues such as air pollution. Therefore, minimizing energy consumption in data centers is becoming a more and more important issue.

In many literatures, the power consumption of a server is modeled in its simplest form as follows:

$$P(u) = P_{min} + (P_{max} - P_{min}) \times u$$

where  $u$  is the CPU utilization,  $P_{min}$  is the power consumed when the server is idle,  $P_{max}$  is the power consumed when the server is fully utilized. In this paper we use two terminologies, PMs and servers, interchangeably. For most of all the servers available in the current market, it is shown that  $P_{min}$  is almost 50% or more of  $P_{max}$  [7]. We can think of turning off a server when it becomes idle. But the procedure of turning off and then on a server takes too much of time and, therefore, cannot be thought of as a practical solution. But some researchers developed a power-saving mechanism in which an idle server can be rapidly put into a sleep mode and consume just 10% of  $P_{min}$ . When a job arrives at the server, it can be put back into the normal operation mode rapidly [8]. With this kind of power-saving mechanism, a server consumes power with the

This work was supported by 2015 Hongik University Research Fund.  
Young-Chul Shim is with Department of Computer Engineering at Hongik University, Seoul, Rep. of Korea (phone: 82-2-320-1695, fax: 82-2-332-1653, e-mail: shim@cs.hongik.ac.kr).

above power model when it has one or more jobs to run but can consume very little power when it has no job.

With the above observation, if we run all the VMs on the smallest number of PMs and, therefore, maximize the number of idle PMs, we can greatly reduce the amount of energy consumed in data centers. This feature of scheduling, packing VMs into a small number of PMs, is called VM consolidation. It is known that VM consolidation is classified into (i) static consolidation, (ii) semi-static consolidation, and (iii) dynamic consolidation [9]. In static consolidation the sizing and placement of VMs on PMs is determined statically when a job arrives and does not change over a period of time. Semi-static consolidation attempts to take advantage of medium to long term workload variations by periodically re-sizing workloads and relocating them on target PMs. Semi-static consolidation typically takes advantage of intra-week variations or intra-month variations. Semi-static consolidation is performed by re-sizing and relocating workloads once a week or once a month. Dynamic VM consolidation extends the idea of semi-static consolidation even further by consolidating workloads daily or multiple times on the same day. Consolidations on such a frequent basis cannot be performed using VM relocation due to downtime required for VM relocation and, therefore, needs live VM re-sizing and live VM migration. While semi-static consolidation is suitable for very long jobs running for several weeks or months and dynamic consolidation is suitable for from long to very long jobs running for several days, static consolidation will be suitable for short jobs running for a couple of hours.

In this paper we will address the problem of static consolidation. This problem has been studied extensively by many researchers as will be described in the next section on related works. But most of their works have the following problems.

- 1) They assume that the performance of a VM is not affected by other VMs running on the same PM. But recent research results, which will be summarized in the next section, show that current virtualization technology does not provide perfect performance isolation. This imperfect performance isolation will have the effect of making job completion time longer than calculated assuming perfect performance isolation and in some cases can cause service level agreement (SLA) violations for some jobs.
- 2) Most of researchers consider only CPU utilization. But there can be various kinds of jobs. They can be CPU-bound, Memory-bound, or IO-bound. For CPU-bound jobs, CPU utilization is an important factor to consider but for IO-bound jobs, IO bandwidth utilization is an important factor. To be able to handle mix of various kinds of jobs, utilization of resources including CPU, memory, and IO should be considered simultaneously.
- 3) Most of researchers assume that a server has one core and a server can run as many a number of VMs as possible unless the total resource requirements of all VMs on that server exceed the capacity of the server. But in reality a server has

many cores and the number of VMs that can be allocated on a core is limited, for example 2 in Amazon EC2.

In this paper we introduce new static VM consolidation algorithms considering (i) imperfect performance isolation of virtualization technology, (ii) mix of various kinds of jobs, and (iii) servers consisting of multiple cores and the limitation on the number of VMs for a core. The proposed algorithms are analyzed using the two metrics, the idle PM ratio and the Service Level Agreement (SLA) violation ratio, through simulation

The rest of the paper is organized as follows. Section 2 summarizes related works. Section 3 describes how much performance degradation a VM may suffer due to imperfect performance isolation when it is run with other VMs on the same PM. Section 4 explains the proposed static VM consolidation algorithm. Section 5 presents the simulation results and is followed by the conclusion in Section 6.

## II. RELATED WORKS

The VM consolidation problem is quite often formulated as a bin packing problem which can be described as follows. Given  $n$  items and  $m$  bins with

$$w_j = \text{weight of item } j, \\ c = \text{capacity of each bin},$$

assign each item to one bin so that the total weight of the items in each bin does not exceed  $c$  and the number of bins used is minimum [10]. This problem is known to be NP-hard and several approximate algorithms are introduced: First-Fit (FF), Best-Fit (BF), First-Fit Decreasing (FFD), and Best-Fit Decreasing (BFD) algorithms. We assume that items and bins are indexed. The FF algorithm considers the items according to increasing indices and assigns each item to the lowest indexed bin which it fits. The BF algorithm is obtained from FF by assigning the current item to the feasible bin having the smallest residual capacity. If items are sorted in the non-increasing order of their weight and then FF and BF algorithms are applied, the resulting algorithms are called FFD and BFD algorithms, respectively. Making an item,  $w_j$ , a bin, and  $c$  correspond to a VM, the resource requirements of VM <sub>$j$</sub> , a PM, and the resource capacity of a PM, respectively, the static VM consolidation problem exactly becomes a bin packing problem. When jobs, each of which requires one VM, arrive one at a time, either FF or BF algorithms can be used. When many jobs, each of which requires one or more VMs, are allocated to PMs simultaneously, FF, BF, FFD, or BFD algorithms can be applied. In this paper we address the situation where jobs, each of which is run on a VM, arrive one at a time. So we use terms, VM and a job, interchangeably in this paper.

Eucalyptus which is open source software for building AWS-compatible clouds, provides the following VM allocation algorithms: a greedy algorithm, a round-robin algorithm, and a power save algorithm [11]. The greedy algorithm is a FF

algorithm explained above. The round robin algorithm mainly focuses distributing the load equally to all the nodes and is outside our interest. The power save algorithm optimizes the power consumption by turning off PMs which are not currently used.

Lin et al proposed a hybrid approach which combines FF and dynamic round robin algorithms [12]. The dynamic round-robin algorithm uses two rules to help consolidate VMs. The first rule is that if a VM has finished and there are still other VMs hosted on the same PM, this PM will accept no more new VMs. Such PMs are referred to as being in the retiring state. The second rule is that if a PM is in the retiring state for a sufficiently long period of time, instead of waiting for the residing VMs to finish, the PM will be forced to migrate the rest of the VMs to other PMs, and be shutdown after the migration finishes. The hybrid approach uses FF during rush hours and uses dynamic round-robin during non-rush hours. This approach is a combination of static and dynamic VM consolidation methods.

[13] describes two energy saving static VM consolidation algorithms: ECTC and MaxUtil. When a new job, which is to be run on a new VM, arrives, a cost function is computed for each PM and the PM which has the lowest cost is selected to run the VM. [14] introduces a modified BFD algorithm which allocates each VM to a PM that provides the least increase of power consumption due to this allocation. The algorithm is used to allocate multiple VMs for multiple jobs simultaneously. The algorithms in [13] and [14] assume that the exact completion time of a job is known a priori and is not affected by others VMs collocated on the same PM.

While all of the above works focus only one resource type, mostly CPU, there are VM allocation and migration algorithms which deal with multiple resource types simultaneously. The Dynamic Integrated Resource Scheduling algorithm in [15], the VectorDot algorithm in [16], and ESWCT and ELMWCT algorithms in [17] take multiple resource types, including CPU, Memory, and IO, into consideration simultaneously. But the goal of these algorithms is to balance resource utilization among PMs, which is outside our interest.

None of the works that we described above consider the performance interference among VMs running on the same PM. Many researchers have studied this performance interference issue [18]-[20]. They take various types of applications, CPU-intensive, Memory-intensive, IO-intensive, and Mixed, run them on a PM together, and study how much performance degradation applications of one type cause to applications of another type. Pu et al in [21] states that network I/O applications are becoming dominating workloads in most cloud-based data centers and studies the performance interference among multiple VMs running on the same hardware platform with the focus on network I/O processing. Kang et al in [22] especially focus on the performance impact of contention in a last-level shared cache (LLC). Through measurement they find that the degree of impact on performance degradation heavily depends on the LLC reference ratio of applications and suggests a throughput-maximizing VM consolidation policy based on this

observation. We use the results of these works in designing and analyzing the static VM allocation algorithms.

### III. PERFORMANCE INTERFERENCE AMONG JOBS

In this section we describe how much performance degradation one job is causing to another job running in the same PM. Each job is executed in a separate VM. We consider three types of jobs as follows.

- 1) CPU-bound jobs: Consist of arithmetic operations mostly and require little IO operations. Because their code size is not large and they do not need large size data structures, their requirement for cache is modest. When they are run alone, their CPU utilization on a core is higher than 95%.
- 2) Memory-bound jobs: Consist of arithmetic operations mostly and require little IO operations like CPU-bound jobs. Basically they read data from a large size array, perform arithmetic operations on them, and write the results to another large size array. We set the size of arrays to be couple of times larger than the LLC size and, therefore, their requirement for cache is very heavy. When they are run alone, their CPU utilization on a core is higher than 90%.
- 3) IO-bound jobs: Consists of disk IO operations mostly, make little use of arithmetic or logical operations, and incurs little burden on CPU. Basically they read data from a large size file and write them back to another file, repeatedly. If the IO bandwidth of a PM is  $B$  Gbps and there are  $N$  cores in a PM, we let the bandwidth requirement of an IO-bound job be  $B/N$  Gbps. When they are run alone, their CPU utilization on a core is 1~2%.

To explain the performance interference between jobs running on the same PM, we use the metric called the *slowdown*. First the completion time of *job1* is measured without any other jobs in the PM. Then we measure the completion time of *job1* while *job2* is running in the same PM. *Job2* can be run on the same core or a different core.  $Slowdown(job1@job2)$  is defined to be the ratio of the second value to the first value and is greater than 1. The larger this value gets, the larger performance degradation *job1* suffers from *job2*.

We analyzed the values from [12]-[16] and performed experiments ourselves to measure the slowdown values for various combinations of jobs. The results are shown in Table 1 and Table 2. The element at the row- $i$  and the column- $j$  represents the value of  $slowdown((row-i \text{ type job})@(column-j \text{ type job}))$ . Table 1 gives slowdown values when two jobs are run on the same core while Table 2 presents slowdown values when two jobs are run on different cores on one PM. From Table 1 we see that slowdown values between CPU-bound jobs and Memory-bound jobs are greater than 2. This means that the job completion time increases more than two times longer. This is because CPU and Memory-bound jobs compete heavily for the CPU resource and, moreover, Memory-bound jobs contend excessively for cache against other jobs. We see around 10% slowdown increase between Memory and IO-bound jobs. This

is because (1) Memory-bound jobs' heavy demand on cache hurts the performance of IO-bound jobs and (2) IO-bound jobs' frequent IO requests incurs frequent service requests on the Dom0, which deteriorate the performance of other jobs. From Table 2 we see that the slowdown values become much lower when two jobs are run on different cores. One case that requires special mention is when a Memory-bound job is run against another Memory-bound job. In this case two Memory-bound jobs contend heavily for LLC and the slowdown value becomes 1.25, which is much higher than other cases in Table 2.

Table 1. Interference between jobs executed on the same core

Against	@ CPU	@ Memory	@ IO
CPU	> 2	> 2	1.05
Memory	> 2	> 2	1.11
IO	1.02	1.09	1.11

Table 2. Interference between jobs executed on different cores

Against	@ CPU	@ Memory	@ IO
CPU	1.02	1.04	1.02
Memory	1.05	1.25	1.05
IO	1.01	1.03	1.05

#### IV. PROPOSED INTERFERENCE AWARE STATIC VM CONSOLIDATION ALGORITHMS

In this section we describe the proposed inter-VM performance interference aware static VM consolidation algorithms. First we explain interference aware versions of random, first fit and best fit algorithms, called IARD (Interference Aware RandOm), IAFF (Interference Aware First Fit), and IABF (Interference Aware Best Fit) algorithms respectively. Then we explain fourth algorithm called IAMBF (Interference Aware Modified Best Fit).

All the proposed algorithms are based upon the following two policies. These policies are based on the fact that up to 2 VMs can be run on a core.

- 1) One CPU or Memory-bound job at a core: At most one CPU or Memory-bound job can be allocated at a core. According to Table 1, if two CPU or Memory-bound jobs are allocated to one, their contention for CPU resource is excessive and their job completion times become longer more than two times, which may cause the requested SLA to be violated.
- 2) One IO-bound job at a core: If a two IO-bound jobs are allocated to one core, the CPU resource of that core is almost wasted because the CPU utilization of an IO-bound job is assumed to be 1~2% in this paper. We also assume that CPU, Memory, and IO-bound jobs are created with equal probability. Therefore, for a PM which has  $N$  cores and can run maximum  $2N$  VMs, there is no need to allocate more than  $N$  IO-bound jobs to one PM. For a core to which one IO-bound job has been allocated, 98~99% of CPU resource remains available and one CPU or Memory-bound job can be

assigned to it without incurring much performance interference. Likewise, to a core to which one CPU or Memory-bound job has been allocated, one IO-bound job can be assigned additionally.

With these two policies, Figure 1 is the algorithm to check whether a PM can accommodate a new job. For the algorithms we assume that there are  $NoOfPM$  PMs in the cloud and each PM is indexed as  $PM_i$  where  $0 < i < NoOfPM$ .

```

IsPMAvailableForJob ( $PM_i$ , JobType) {
    if (JobType is either CPU or Memory-bound)
        if (total number of CPU or Memory-bound jobs
            allocated to the PM is less than N)
            return (True);
        else
            return (False);
    else /* JobType is IO-bound */
        if (total number of IO-bound jobs allocated
            to the PM is less than N)
            return (True);
        else
            return (False);
}

```

Figure 1. Algorithm to check whether a PM is available for a new job

And Figure 2 shows the algorithm to select a core within a PM to run a new job when that PM is decided to be available for that job. The algorithm attempts to find an idle core first. This is because jobs running on different cores interfere with each other less than those running on the same core as can be seen Table 1 and 2.

```

SelectCoreForNewJob ( $PM_i$ , JobType) {
    if (PM has an idle core)
        return id of any idle core;
    else
        if (JobType is CPU or Memory-bound)
            return id of any core running one IO-bound job;
        else /* JobType is IO-bound */
            return id of any core running CPU or
                Memory-bound job;
}

```

Figure 2. Algorithm to select a core to run a new job

In IARD, a PM is selected randomly and checked for availability with the algorithm in Figure 1. If available, a core is selected with the algorithm in Figure 2. Otherwise, another PM is selected randomly and this process is repeated until an available PM is found.

```

IAFF (JobType) {
    for ( $i = 0$ ;  $i < NoOfPM$ ;  $i++$ )
        if (IsPMAvailableForJob( $PM_i$ , JobType))
            return ( $PM_i$ );
}

```

Figure 3. IAFF algorithm

Figure 3 shows the IAFF algorithm. IAFF checks PMs in order until an available PM is found. Basically it finds the first PM which can accommodate the new job.

Figure 4 shows the IABF algorithm. We assume that a PM has  $N$  cores. PMs are sorted in the decreasing order of the number of jobs running on them. PMs are checked from the highest priority machines to the lowest priority machines. PMs with  $2N-1$  jobs have the highest priority and idle PMs are the lowest priority machines. Because of the constraint that up to 2 VMs can be run on a core and 2 policies stated in the beginning of this section, we claim that metrics such as CPU utilization or IO utilization do not play a significant role in determining the priority of PMs. Instead we simply use the number of jobs (or VMs) running on a PM to determine the priority. This simple metric is sufficient in our environment in which all the PMs have the same resource capacity but more complex metrics may need to be adopted if this hardware homogeneity assumption is not valid.

```
IABF (JobType) {
  for (i = 2N-1; i >= 0; i--) {
    S is the set of PMs having i jobs;
    for each PMj in S {
      if (IsPMAvailableForJob(PMj, JobType))
        return (PMj)
    }
  }
}
```

Figure 4. IABF algorithm

From Table 2 we see that the Memory-bound job running on a core slows down other jobs running on other cores non-trivially and especially Memory-bound jobs much heavily. From this observation we need to curb the number of Memory-bound jobs simultaneously running on a PM and we choose to limit up to  $N/2$  Memory-bound jobs on a PM. This additional policy can be implemented by modifying the algorithm in Figure 1 as that in Figure 5. If the IABF algorithm uses this new PM availability checking algorithm, the resulting BF algorithm is called an IAMBF algorithm.

```
IsPMAvailableForJob (PMj, JobType) {
  if (JobType is CPU-bound)
    if (total number of CPU or Memory-bound jobs
        allocated to the PM is less than N)
      return (True);
  else
    return (False);
  else if (JobType is Memory-bound)
    if (total number of CPU or Memory-bound jobs
        allocated to the PM is less than N and
        total number of Memory-bound jobs allocated
        to the PM is less than N/2)
      return (True);
  else
    return (False);
  else /* JobType is IO-bound */
    if (total number of IO-bound jobs allocated
```

```
    to the PM is less than N)
      return (True);
  else
    return (False);
}
```

Figure 5. Modified algorithm to check whether a PM is available for a new job

## V. EXPERIMENTAL RESULTS

In this section we first explain simulation environments for analyzing the performance of proposed algorithms. Then we present the experimental results.

For simulation we consider a data center consisting of 32 homogeneous PMs. Each PM has 2 cores and each core can run up to 2 VMs. Therefore at maximum 128 VMs can run simultaneously. As explained in Section 3, we consider three types of jobs: CPU-bound, Memory-bound, and IO-bound. Each job is executed on one VM. Jobs of these three types are created with equal probability. The size of a job is defined to be the optimal execution time of the job under no inter-VM interference from other jobs. We assume that the inter-arrival time and the size of jobs follow the exponential distribution.

From the simulation we measure two metrics: (1) ratio of idle PMs and (2) ratio of SLA violations. The ratio of idle PMs is obtained as follows. First the total idle time of each PM during the simulation is calculated and this per-PM idle time is summed up for all the PMs in the data center. Then this sum is divided by the number, (simulation time)  $\times$  (total number of PMs). The ratio of idle PMs represents the energy efficiency of the simulated algorithm and takes the value from 0 to 1. The larger this ratio gets, the more idle PMs there are in the data center during the simulation.

The actual completion time of a job during the simulation is defined to be the time between the arrival time of the job and the finish time of the job. If a job is allocated to a core as soon as it arrives, the actual completion time of the job consists of only its actual execution time. But if a job has to wait in the job queue for a core until it is allocated, the actual completion time consists of waiting time and actual execution time. Moreover the execution time can become longer than its ideal execution time due to the performance interference from other jobs running in the same PM. The slowdown of a job is defined to be the ratio of the actual completion time to the ideal execution time. The slowdown has a value greater than or equal to 1. If the slowdown value gets higher than some threshold, we say that the SLA of the job is violated. In this paper we choose 1.25 for this threshold. To calculate the ratio of SLA violation, we count the number of completed jobs whose SLA has been violated and then divide this count value by the total number of completed jobs during the simulation. The SLA violation ratio gets value from 0 to 1. The smaller this ratio gets, the more reliable the consolidation algorithm becomes.

For the simulation we vary the load level of a data center. The load level value reflects how many and how large jobs are submitted to the data center. The load level is calculated as

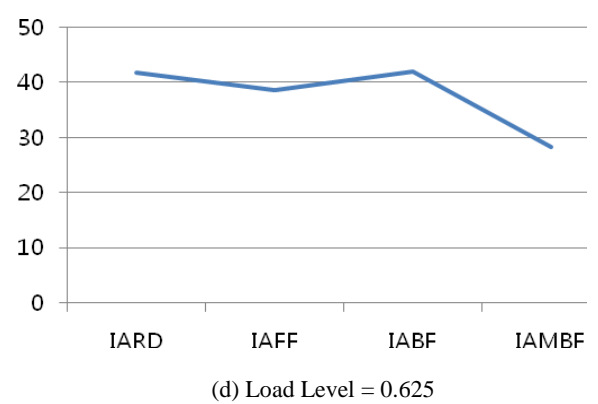
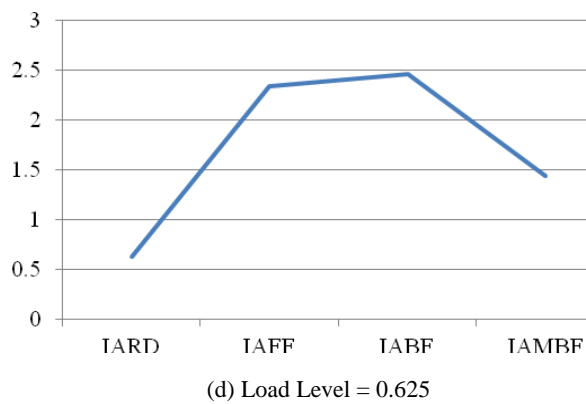
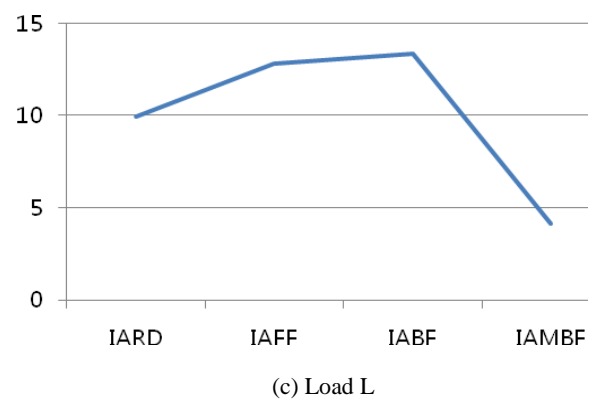
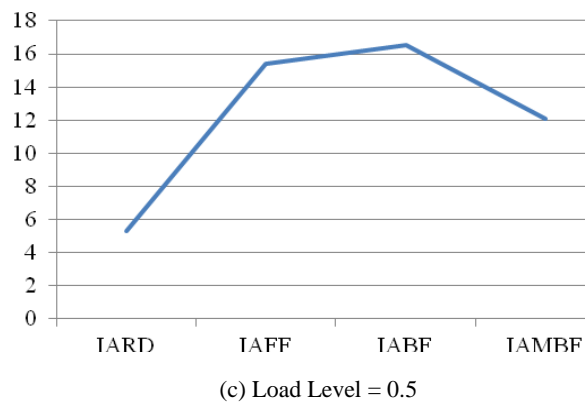
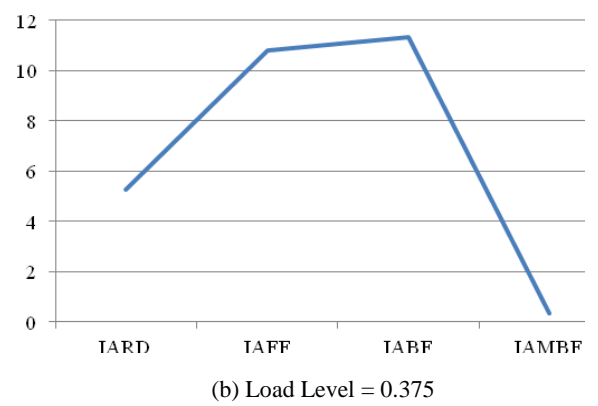
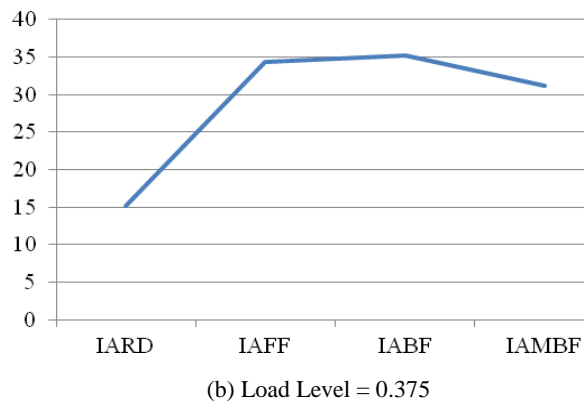
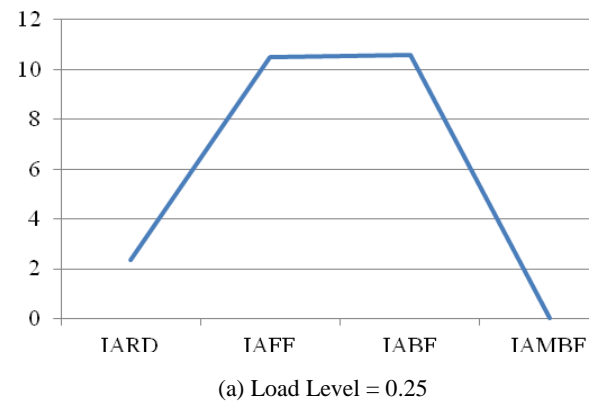
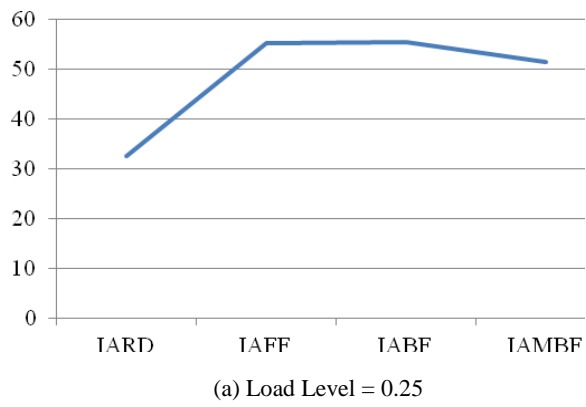


Figure 6. The ratio of idle PMs (in %)

Figure 7. The ratio of SLA violation (in %)

follows.

$$\frac{((\text{average job size}) \times (\text{average number of jobs per sec}))}{((\text{number of PMs}) \times (\text{number of cores per PM}) \times 2)}$$

In this paper we consider four load levels: 0.25, 0.375, 0.5, 0.625. The load level 0.25 means that the data center is lightly loaded while 0.625 means that the data center is heavily loaded.

Figure 6 shows the idle PM ratio of the four VM consolidation algorithms under varying load levels. Under any load levels, IABF has the highest idle PM ratio and IAFF has slightly lower number. IAMBF's idle PM ratio value is lower than IABF and IAFF but not significantly lower. IARD has much lower idle PM ratio than any other algorithms. This result is quite obvious because IABF, IAFF, and IAMBF algorithms try to allocate a new job to an already packed PM and try to maximize the number of idle PMs. The IAMBF algorithm shows slightly lower performance in terms of energy efficiency because this algorithm avoids the situation when too many Memory-bound jobs are allocated to one PM and, therefore, many of jobs on this PM suffer from the SLA violation.

Figure 7 presents the ratio of SLA violation of 4 VM consolidation algorithms under varying load levels. Under any load levels, IAMBF has much lower SLA violation ratio than any other algorithms because it tries to reduce performance interference among VMs at a PM. Under low to medium load levels, IARD's ratio is much higher than IAMBF. But its SLA violation ratio is much lower than IAFF and IABF, which is made possible at the expense of using more PMs than IAFF and IABF. Under the heavy load level, the SLA violation ratios of IARD, IAFF, and IABF are similar and all of them are much higher than IAMBF. Even for IAMBF the SLA violation ratio rises sharply as the load level is rises and becomes 0.28 when the load level is 0.625. This is because as the load level rises, new jobs find it more difficult to find PMs which can accept them and, therefore, their waiting time in the queue increases.

We want to minimize the energy consumption in a data center but at the same time we do not want the job completion time to be delayed excessively to save energy. From simulation results in Figures 6 and 7, we conclude that the IAMBF VM consolidation algorithm significantly saves energy in a data center without incurring too much performance degradation of jobs due to inter-VM interference.

## VI. CONCLUSION

In this paper we addressed a static VM consolidation problem in a cloud-based data center and proposed VM consolidation algorithms satisfying the following two conflicting goals: (1) minimize the energy consumption in a data center and (2) reduce the performance degradation on jobs due to inter-VM interference among VMs allocated to the same PM. The proposed algorithms are based upon the following assumptions: (1) a job can be CPU, Memory, or IO-bound, (2) each job arrives one at a time, is short-lived at most couple of hours, and is run on one VM, and (3) all the PMs in a data center are

homogeneous in terms of their hardware capacity, each PM consists of multiple core, and each core can run up to 2 VMs.

We modeled the static VM consolidation problem as a bin packing problem and developed four algorithms: interference-aware random, interference-aware first fit, interference-aware best fit, and interference-aware modified best fit. These algorithms are variations of random, first fit and best fit algorithms and try to reduce performance degradation due to inter-VM interference. The proposed algorithms were analyzed through simulation. Their performance was compared using two metrics: the ratio of idle PMs and the ratio of SLA violations. The interference-aware modified best fit algorithms exhibited the lowest ratio of SLA violations with slightly lower ratio of idle PMs than interference-aware first fit and interference-aware best fit algorithms.

To measure the energy efficiency we just counted the number of idle PMs in this paper, but we will analyze the energy efficiency of proposed algorithms using the energy consumption model described in Section 1 in the near future. We also plan to relax the assumptions made in this paper and consider (1) long-lived jobs whose execution may require multiple cooperating VMs and (2) a heterogeneous data center which consists of PMs of different hardware capacities.

## REFERENCES

- [1] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," *NIST Special Publication 800-15*.
- [2] A. Li, X. Yang, S. Kandula, and M. Zhang, "Cloud-Cmp: Comparing Public cloud Providers," *ACM Internet Measurement Conference*, 2010.
- [3] G. Lu and W. Zeng, "cloud Computing Survey," *Applied Mechanics and Materials*, vol. 530-531, 2014, pp. 650-661.
- [4] X. Pu, L. Liu, Y. Mei, S. Sivathanu, Y. Koh, and C. Pu, "Who is Your Neighbor: Net I/O Performance Interference in Virtualized Clouds," *IEEE Transactions on Services Computing*, vol. 6, no. 3, 2013, pp.314-329.
- [5] M. Mishra, A. Das, P. Kukarni, and A. Sahoo, "Dynamic Resource Management Using Virtual Machine Migrations," *IEEE Communication Magazine*, Sep. 2012, pp. 34-40.
- [6] J. Glanz, "Power, Pollution and the Internet," *The New York Times*, Sep. 22, 2012. Available: <http://www.nytimes.com>.
- [7] L. A. Barroso and U. Hoesle, "The Case for Energy-Proportional Computing," *IEEE Computer*, Dec. 2007, pp. 33-37.
- [8] D. Meisner, B. T. Gold, T. F. Wenisch, "PowerNap: Eliminating Server Idle Power," *ACM ASPLOS*, 2009, pp. 205-216.
- [9] A. Verma, J. Bagrodia, and V. Jaiswal, "Virtual Machine Consolidation in the Wild," *ACM Middleware*, 2014, pp. 313-324.
- [10] S. Martello and P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*, John Wiley and Sons Ltd., 1990.
- [11] S. Subramanian et al., "An Adaptive Algorithm for Dynamic Priority based Virtual Machine Scheduling in Cloud," *Int. Journal of Computer Science Issues*, vol. 9, no. 2, pp. 397-402, 2012.
- [12] C.-C. Lin, P. Liu, and J.-J. Wu, "Energy-Efficient Virtual Machine Provision Algorithms for Cloud Systems," *IEEE Int. Conf. on Utility and Cloud Computing*, pp. 81-88, 2011.
- [13] Y. C. Lee and A. Y. Zomaya, "Energy efficient Utilization of Resources in Cloud Computing Systems," *Springer Journal of Supercomputing*, vol. 60, pp. 268-280, 2012.
- [14] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-Aware Resource Allocation Heuristics for Efficient Management of Data Centers for Cloud Computing," *Future Generation Compute Systems*, vol. 28, pp. 755-768, 2012.
- [15] W. Tian et al, "Dynamic and Integrated Load-Balancing Scheduling Algorithm for Cloud Data Centers," *China Communications*, vol. 8, no. 6, pp. 117-126, 2011.



- [16] A. Singh, M. Korupolu, and D. Mohapatra, "Server-Storage Virtualization: Integration and Load Balancing in Data Centers," *ACM/IEEE Conf. on Supercomputing*, 2008.
- [17] H. Li, J. Wang, J. Peng, J. Wang, and T. Liu, "Energy-Aware Scheduling Scheme Using Workload-Aware Consolidation Technique in Cloud Data Center," *China Communications*, pp. 114-124, 2013.
- [18] Y. Koh, R. Knauerhase, P. Brett, M. Bowman, Z. Wen, and C. Pu, "An Analysis of Performance Interference Effects in Virtual Environments," *IEEE Int. Symp. On Performance Analysis of Systems & Software*, pp. 200-209, 2007.
- [19] F. Oh, H. S. Kim, H. Eom, and H. Y. Yeom, "Enabling Consolidation and Scaling Down to Provide Power Management for Cloud Computing," *3rd USENIX Conf. on Hot Topics in Cloud Computing*, 2011.
- [20] S. Govindan, J. Liu, A. Kansal, and A. Sivasubramaniam, "Cuanta: Quantifying Effects of Shared On-Chip Resource Interference for Consolidated Virtual Machines," *SOCC*, 2011.
- [21] X. Pu et al., "Who is Your Neighbor: Net I/O Performance Interference in Virtualized Clouds," *IEEE Transactions on Services Computing*, vol. 6, no. 3, pp. 314-329, 2013.
- [22] S. Kang, S.-G. Kim, H. Eom, and H. Y. Yeom, "Toward Workload-Aware Virtual Machine Consolidation on Cloud Platforms," *ICUIMC*, 2012.

# A study on the Intelligent Techniques of the ECG-based Biometric Systems

M. Bassiouni, W.Khalefa, El-Sayed A. El-Dahshan , and Abdel-Badeeh. M. Salem.

**Abstract**— Electrocardiogram (ECG) signal reflects the recorded cardiac electrical activity of an individual over time; it has been emerged as an important biometric trait. It has the advantages of being a unique aliveness indicator as it is difficult to spoofed and falsified. In this paper, we will present a comprehensive survey on the employment of ECG in biometric systems. An overview of ECG, its techniques and methods followed by a series of studies are presented. We will demonstrate the most datasets used in ECG and also the devices needed to capture them. ECG based biometric systems can be fiducial or non-fiducial according to the utilized features. In particular, we categorize the methodologies based on features extraction, reduction and classification schemes. Finally, we present a comparative analysis of the authentication performance of ECG biometric systems.

**Keywords**— Biometric, Electrocardiogram, Machine Learning, Preprocessing, Feature Extraction, Feature Reduction, Classification.

## I. INTRODUCTION

With the rapid development of transportation, communication, and network technology in modern society, the scope of human activities is broadening, while the importance of identification becomes increasingly prominent. Traditional identification methods are generally divided into two kinds: items that people remember, such as user names, passwords, etc. and items that people own, such as keys, identification cards, etc. The two types both have limitations in that they may be easily forgotten and lost. These traditional identification techniques cannot meet the higher security demands of a highly modernized society. Biological recognition technology refers to individual identification based on unique physiological or behavioral characteristics of the human body [1].

M.Bassiouni; Egyptian E-Learning University (EELU), 33 El-meshah Street, Eldoki; Postal Code: 11261; El-Giza; Egypt; (e-mail: mbassiouni@eelu.edu.eg)

W.Khalefa; Faculty of Computer and Information Science; Ain Shams University; Abbassia; Cairo; Egypt (e-mail: wael.khalifa@gmail.com)

El-Sayed A. El-Dahshan ; Egyptian E-Learning University (EELU), 33 El-meshah Street, Eldoki; Postal Code: 11261; El-Giza; Egypt; (e-mail: Sayed Eldahshan @eelu.edu.eg )

A.M. Salam; Faculty of Computer and Information Science; Ain Shams University; Abbassia; Cairo; Egypt (e-mail: absalem@cis.asu.edu.eg )

The main objective of this study is to identify the most commonly used methods and techniques used for ECG as a biometric identification system determining the datasets used in biometric identification, the features extraction and classification methods used. We will determine the most efficient techniques that lead to a high identification rate and accuracy and number of subjects used for identification. To achieve this objective the rest of the paper is organized as follows. Section II gives a general methodology of using ECG as a biometric. Section III presents some of the recent works in ECG. Section IV presents the discussion; Section V presents the conclusion and future work and Section VI presents the references

## II. GENERAL METHODOLOGY OF ECG BIOMETRIC

The ECG signal measures the change in electrical potential over time. The trace of each heartbeat consists of three complexes: P, R, S, T, and U. these complexes are defined by the fiducial that is the peak of each complex. Fig. 1 documents the commonly used medical science ECG fiducials.

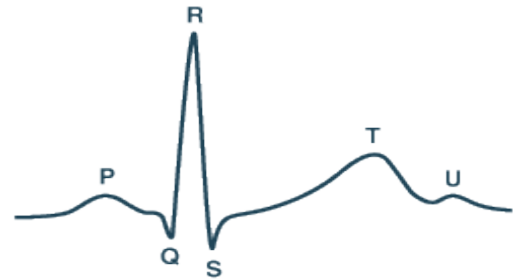


Fig. 1 Ideal ECG signal containing P, Q, R, S, T, U peaks  
ECG biometric recognition system consists of a set of stages starting from data acquisition, preprocessing, using fiducial or non fiducial approach, feature extraction, feature reduction and classification as shown in Fig. 2.

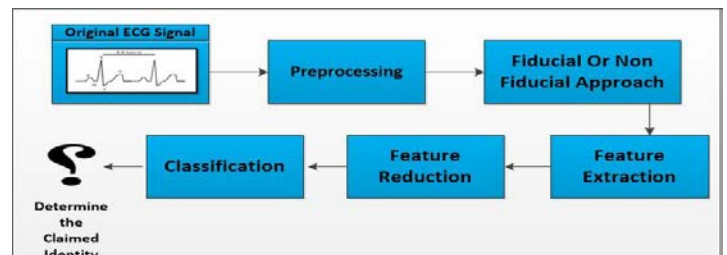


Fig.2 ECG Biometric Recognition System stages

### A. Data Acquisition

There are two methods to obtain ECG signals:

First method is using a device to capture ECG. There are different devices to capture ECG, like heal force color portable ECG monitor with ECG leads cables it is a compact, potable handy device for patients to record 30 seconds of ECG data anytime or anywhere, comfortably, with or without electrodes and the measure ECG data can be transferred to PC, laptop. Another device is mini portable heart health care device micro ECG recorder it have different recording mode, quick mode, event mode, holter mode and monitor mode, lead connection limb lead (I, II, III), chest lead (V1, V3, V5) and sampling rate 100/200/400sps, and can be user adjustable, input band width 0.05Hz – 150 Hz and can be user adjustable and the heart rate range from 30bpm~250bpm and the threshold of the heart rate alarm is adjustable

Second is collecting database for an ECG signal from the internet and using them for performing the biometric system stages. There are some data sets available for identification of ECG, and the most commonly used First a database called MIT-BIH Arrhythmia Database it contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. Twenty-three recordings were chosen at random from a set of 4000 24-hour ambulatory ECG recordings collected from a mixed population of inpatients (about 60%) and outpatients (about 40%) at Boston's Beth Israel Hospital; the remaining 25 recordings were selected from the same set to include less common but clinically significant arrhythmias that would not be well-represented in a small random sample.

Second a database called PTB this database contains 549 records from 290 subjects (aged 17 to 87, mean 57.2; 209 men, mean age 55.5, and 81 women, mean age 61.6; ages were not recorded for 1 female and 14 male subjects). Each subject is represented by one to five records. There are no subjects numbered 124, 132, 134, or 161. Each record includes 15 simultaneously measured signals: the conventional 12 leads (i, ii, iii, avr, avl, avf, v1, v2, v3, v4, v5, v6) together with the 3 Frank lead ECGs (vx, vy, vz). Each signal is digitized at 1000 samples per second, with 16 bit resolution over a range of  $\pm 16.384$  mV.

Third a database called Fantasia it contains twenty young (21 - 34 years old) and twenty elderly (68 - 85 years old) rigorously-screened healthy subjects underwent 120 minutes of continuous supine resting while continuous electrocardiographic (ECG), and respiration signals were collected; in half of each group, the recordings also include an uncalibrated continuous non-invasive blood pressure signal. Each subgroup of subjects includes equal numbers of men and women. The continuous ECG, respiration, and (where available) blood pressure signals were digitized at 250 Hz. Each heartbeat was annotated using an automated arrhythmia

detection algorithm, and each beat annotation was verified by visual inspection.

Last database is called ECG-ID it contains 310 ECG recordings, obtained from 90 persons. Each recording contains: ECG lead I, recorded for 20 seconds, digitized at 500 Hz with 12-bit resolution over a nominal  $\pm 10$  mV range; the records were obtained from volunteers (44 men and 46 women aged from 13 to 75 years who were students, colleagues, and friends of the author). The number of records for each person varies from 2 (collected during one day) to 20 (collected periodically over 6 months).

### B. Preprocessing

Preprocessing is called de-noising of signal in which the noise is removed from the original signal. (ECG) signals they will be preprocessed in order to remove baseline wander, dc shift, power-line noise, high frequency interference and also if the signal is not captured well due to human error or movement of the stethoscope during taking the signal. Most of these noises are removed using low pass filter, band stop and pass filters, high pass filters; also wavelets play an important role in removing these noises to produce filtered signals.

### C. Fiducial or Not Fiducial approach

The existing ECG-based biometric system can be categorized into fiducial or non-fiducial systems according to the utilized approach to feature extraction. The fiducial approach requires the detection of fiducial points from heartbeat in an ECG trace. These fiducial points allow us to produce fiducial features represent the temporal and amplitude distances between fiducial points along with angle features. On the other hand, non-fiducial approaches usually operate in the frequency domain (ex: wavelet, discrete cosine transform (DCT) ...), and they have the advantage of relaxing the detection process to include only the R peak, which is considered the easiest point to detect due to its strong sharpness, and for some approaches, no detection is needed at all.

### D. Feature Extraction and Reduction

The most important process after collecting the database and performing pre-processing is the feature extraction and feature reduction process. We select the most discriminant values in the features. If these features were large enough so it is reduced to small amount of data. Most of the techniques used in feature extraction and reduction for ECG are based on non fiducial approaches such as, Wavelet Transform Techniques, Rough sets, Cross-correlation, Auto correlation and fiducial approaches that is based on selecting 11 or 15 or 19 or 36 features from the ECG signal representing distance, amplitudes and angles features, including also RR interval, also QRS and P and T delineation methods and also some morphology and segmentation techniques to select the ECG features as shown in the survey in Table.1.

### E. Classification

The extracted features are compared against the stored templates to generate match scores. In a heart-based biometric system, the number of matching data between the input and the template feature sets is determined and a match score

reported. The match score may be moderated by the quality of the presented biometric data. Most of the techniques used in classification for ECG are Vector Quantization, Gaussian Mixture Mode (GMM), Support Vector Machine (SVM), Artificial Neural Network (ANN), Radial Bias Function Neural Network (RBFNN), Euclidean Distance (ED), and K-nearest neighbor (KNN), template matching, Similarity Distance and Hidden Markov Models.

### III. ANALYSIS OF TECHNIQUES USED IN ECG RECOGNITION

Table. 1 Machine learning techniques used in the ECG identification

Author	Database Size	Feature Set	Classification	Accuracy
<b>Tatiana, et. Al, 2005 [2]</b>	Data set of 90 persons ECG-ID Database Age from (13 – 75) 195 Train 115 Test	Wavelet + (PCA) Principle Component Analysis)	Linear Discriminant analysis (LDA) and Majority Vote Classifier	96%
<b>Steven, et. al, 2005 [3]</b>	29 Individual Male and Female Age from (22 – 48)	In total 15 features were extracted from each heartbeat	Standard linear discriminant analysis.	According to the sensor location 100% (neck and chest) Within Anxiety state 97% (low, high stress) Between Anxiety state 98% (low, high stress)
<b>Saechia, et.al, 2005 [4]</b>	MIT-BIH Database Sample size = 35	Fourier Transform (FT)	Multilayer Perceptron (MLP) Back Propagation (BP)	97.15%
<b>Plataniotis, et.al, 2006 [5]</b>	PTB Database No of subjects = 14	Auto Correlation and Discrete cosine transform (AC/DCT)	Normalized Euclidean distance Normalized Gaussian log likelihood	100% for both Normalized ED Normalized Gaussian log likelihood
<b>Molina, et.al, 2007 [6]</b>	Sample size = 10	Depends on the R-R interval segmentation using Morphology Getting Amplitude and length normalization	Similarity score compared with a threshold $\theta$	98%
<b>Yongjin, et.al, 2007 [7]</b>	PTB Database No of subjects = 294 MIT-BIH Database No of subjects = 13	Two Approaches 1 <sup>st</sup> is 15 Temporal features 6 Amplitude features PCA, LDA 2 <sup>nd</sup> (AC / DCT)	ED Nearest Neighbor (NN)	Accuracy of (95.5%) with (PTB + PCA), (93%) with (PTB + LDA) and (98%) with (MIT+PCA), (98%) with (MIT+LDA) and (94%) with AC (DCT).
<b>Adrian, et.al, 2008 [8]</b>	No of subjects = 50 45 male, 5 female Ages from (18 – 40)	PQRST complexes were automatically detected using the multiplication of backward differences algorithm. Correlation coefficients were computed between PQRST complexes	Three different quantitative measures: percent residual difference (PRD) Correlation Coefficient (CC) Wavelet distance measure(WDIST)	PRD = 70% CCORR = 80% WDIST = 89%
<b>Yogendra, et.al, 2008 [9]</b>	A test set of 250 ECG recordings prepared from 50 subjects ECG from Physionet	19 features based on time intervals, amplitudes and angles.	Template Matching	99%

<b>Fatemian, et.al, 2009 [10]</b>	PTB Database No of subjects = 294 MIT-BIH Database No of subjects = 13	QRS detection and delineation of T and P wave then PCA LDA	Template Correlation	99.61%
<b>Boumbarov, et.al, 2009 [11]</b>	No of subject = 9 20 complete cardiac cycles for each.	linear subspace projection using PCA and improving the separability by LDA	RBFNN	86.1%
<b>Can Ye, et.al, 2010 [12]</b>	MITBIH Arrhythmias Database, MIT-BIH Normal Sinus Rhythm Database and Long-Term ST Database	Wavelet Transform (WT) and Independent Component Analysis (ICA) methods are applied to extract morphological features.	SVM RBFNN	99.6%
<b>Jun Shen, et.al, 2011 [13]</b>	PTB Database No of subjects = 13 MIT-BIH Database No of subjects = 14 Self-collected = 15	Piecewise Linear Representation (PLR) is used to keep important information of an ECG signal segment	(DTW) Dynamic Time Warping	100%
<b>Khairul, et.al, 2012 [14]</b>	No of subjects = 30 obtained from a non-invasive measurement called the Revitus ECG module	With Normalized QRS complex  Without Normalized QRS Complex	MLP	With Normalized QRS complex it has an accuracy of 96.1% Without Normalized QRS complex it has an accuracy of 93.4%
<b>Manal, et.al, 2013 [15]</b>	PTB Database Starting from 13,25,50, 75 and to 90 subjects	1 <sup>st</sup> approach Wavelet and reduced coefficients to examine the utilization of QT and QRS intervals and RR interval 2 <sup>nd</sup> approach (AC / DCT)	RBFNN	1 <sup>st</sup> approach 100% for RR from (13 to 90) 100% for QT For 13 and 83.3% for 25,50,70,90, 100% for QT For 13 and 83.3% for 25,50,70 for 90 66.67% 2 <sup>nd</sup> approach 100%
<b>Vuksanovic, et.al, 2014 [16]</b>	Training set = 13 subjects Test set includes another same 13 subjects.	QRS detection, various temporal, amplitude and AR coefficients are extracted	ANN	Accuracy using width is 55.56% , Amplitude 98.89% With Amplitude and Ar coefficients is 95% Amplitude, Ar, Time, Width is 100%
<b>X. Tang1, et.al, 2014 [17]</b>	MIT-BIH arrhythmia	Rough set	Quantum Neural Network	91.7%
<b>Tiago, et.al, 2015 [18]</b>	ECG data were collected from 63 subjects during two data-recording sessions separated by six months (Time Instance 1, T1, and Time Instance 2, T2).	Morphology and Segmentation of RR heartbeats	KNN classifier, using the mean wave's Euclidean distances	95.2% for the First Test 90.2% for the Second Test
<b>Sandeep, et.al, 2015 [19]</b>	MIT-BIH	A Multitask learning approach i in which feature extraction and classifier design are carried out simultaneously	KNN	94.5%



#### IV. DISCUSSION

According to the survey, we have presented novel biometric identification techniques that are based on heart sounds. The most techniques used were based fiducial and non fiducial approach. Fiducial features represent the temporal and amplitude distances between fiducial points along with angle features. Hence, they require the detection of 11 fiducial points from each heartbeat: three peak points (P, R and T), two valleys (Q and S) and the six onsets and offsets for the three heartbeat waves.

Thus, the efficiency of the fiducial approach significantly relies on the accuracy of the fiducial detection process, which is a big challenge by itself especially for the onsets and the offsets points, since they are susceptible to error and there is no universally acknowledged rule for defining exactly where the wave boundaries lie.

On the other hand, non-fiducial based approaches usually investigate the ECG spectra. Only the R peak is needed for such approaches and for some of them; no detection is needed at all. However, non-fiducial approaches usually result in a high dimension feature space (hundreds of coefficients), which in turn has its limitation.

Some of these techniques achieved a high accuracy even reached to 100% accuracy using QRS detection and delineation, P and T wave detection and delineation followed by LDA and PCA and Piecewise linear representation. But most of them share the problem that the evaluation is carried over small databases, making the results obtained difficult to generalize. In fiducial approaches entropy and energy of the ECG signal can be introduced to improve the features obtained, also a combined approach using fiducial and non fiducial approach must be introduced deeply as it can have a lot of computational overload but it will improve the accuracy. Most of the previously mentioned studies evaluated the biometric system performance regardless of many combined effectual factors, such as age, database scale, race, and gender and disease status. It is showed from the studies that there are two different classification approaches. The first approach is based on the concept of similarity; such as template matching and KNN are used in my studies. Second one is to construct decision boundaries by optimizing certain error criterion. Examples are ANN, RDF and SVM. Also the recognition performance could be more efficient if fusion between different classifiers is introduced.

#### V. CONCLUSION

This review discussed on one of the most extensively studied medical biometric systems, the ECG. The ECG signal measures the change in electrical potential over time. We presented in this paper several intelligent techniques that were used in user identification systems based on ECG. Each technique is explained with its data set used. The best accuracy was achieved by Jun Shen, et.al, using dynamic time wrapping, Can Ye using support vector machine and radial

bias function and S.Zahra Fatemian, et.al, using template correction. In the future, we plan to work on a new ECG system that will use a large dataset. The system will use existing dataset for user identification.

#### VI. REFERENCES

1. A. ROSS, A.K. JAIN, "Human Recognition using biometrics: an Overview," *Annals of Telecommunications*, vol. 62, Jan/Feb 2007, pp. 11-35.
2. A.P. Nemirko, T.S. Lugovaya, "Biometric human identification based on electrocardiogram," *Proc. XII-th Russian Conference on Mathematical Methods of Pattern Recognition*, Moscow, MAK Press, 2005, pp. 387-390.
3. A. Israel, M. Irvine, A. Cheng, D. Wiederhold, K. Wiederhold, "ECG to identify individuals," *Pattern Recognition* 38, (2005), pp. 133 – 142.
4. S. Saechia, J. Koseeyaporn, P. Wardkein, "Human Identification System Based ECG Signal," *TENCON IEEE Region 10 Melbourne, Qld*, 2005, pp. 1 - 4.
5. N. Plataniotis, D. Hatzinakos, K.M. Lee, "ECG biometric recognition without fiducial detection," *Proc. of Biometrics Symposiums (BSYM)*, Baltimore, Maryland, USA, 2006.
6. G.G. Molina, F. Bruekers, C. Presura, "Morphological synthesis of ECG signals for person authentication," *15th European Signal Proc. Conf. Poznan, Poland*, September 3-7, 2007.
7. Y. Wang, F. Agraftioti, D. Hatzinakos, N. Plataniotis. "Analysis of Human Electrocardiogram for Biometric Recognition" *Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing* Volume 2008,
8. D.C. Chan, M. Hamdy, A. Badre, V. Badee, "Wavelet distance measure for person identification using electrocardiograms," *Instrumentation and Measurement, IEEE Transactions on* 57(2), pp.248 –253.
9. Y. Singh, P. Gupta, "ECG to individual identification," *"2nd IEEE Int. Conf. on Biometrics: Theory, Applications and Systems 2008"*.
10. S. Fatemian, D. Hatzinakos, "A new ECG feature extractor for biometric recognition," *16th International Conference on Digital Signal Processing*, 2009, pp. 1 –6.
11. O. Boumbarov, Y. Velchev, S. Sokolov, "ECG personal identification in subspaces using radial basis neural networks," *IEEE Int. Workshop on Intelligent Data Acquisition and Advanced Computing Systems*, 2009, pp. 446 –451.
12. C. Ye, M.T. Coimbra, B. Vijaya, "Arrhythmia Detection and Classification using Morphological and Dynamic Features of ECG Signals," *32nd Annual International Conference of the IEEE EMBS Buenos Aires, Argentina*, August 31 - September 4, 2010.
13. J. Shen, S. Bao, "The PLR-DTW Method for ECG Based Biometric Identification," *33rd Annual International Conference of the IEEE EMBS Boston, Massachusetts USA*, August 30 - September 3, 2011.
14. K. Azami, I. Khalil, M. Smole, "ECG Biometric Recognition in Different Physiological Conditions using

- Robust Normalized QRS". Complexes.cinc.org Computing in Cardiology 2012; vol.39, pp.97-100.
15. M. Tantawi, K. Revett, A.M. Salem, M.F. Tolba." A wavelet feature extraction method for electrocardiogram (ECG)-based biometric recognition". © Springer-Verlag London 2013.
  16. B. Vuksanovic , M. Alhamdi. "Analysis of Human Electrocardiogram for Biometric Recognition Using Analytic and AR Modeling Extracted Parameters". *International Journal of Information and Electronics Engineering, Vol. 4, No. 6, November 2014.*
  17. X. Tan , LShu. "Classification of Electrocardiogram signal with RS and Quantum neural networks". *International Journal of Multimedia and Ubiquitous Engineering Vol.9, No.2 (2014), pp.363-372.*
  18. T. Araújo, N. Nunes, H. Gamboa, A. Fred. "Generic Biometry Algorithm Based on Signal Morphology Information," © Springer International Publishing Switzerland 2015
  19. S. Gutta, Qi. Cheng, "Joint Feature Extraction and Classifier Design for ECG Based Biometric Recognition," *IEEE Journal of Biomedical and Health Informatics, 2015.*

**Mahmoud Bassiouni** is a Teaching assistant at Egyptian E-learning University, received his B.Sc. in computer science from Ain shams University Cairo, Egypt in 2012. He received a premaster in computer science from Ain shams University Cairo, Egypt 2013. He is currently enrolled in master degree in computer science from Ain shams University Cairo, Egypt. Research interest: Biometrics, machine learning, Artificial Intelligence

**Dr. Wael Khalifa** is a lecturer at the Computer Science Department, Faculty of computer and Information Sciences Ain Shams University. Most recently he finished is PhD in biometrics. Research Interest: Bio-informatics and e-Health Medical Informatics, Biometrics

**El-Sayed A. El-Dahshan** received his B.Sc. in Physics and Computer Science from Ain Shams University Cairo, Egypt in 1986. He received a post graduate diploma in electronics from Ain Shams University, Cairo Egypt 1988. In 1990 he received his M.Sc. in the microwaves area from Ain Shams University Cairo, Egypt. He received his Ph.D. degree in thin films technology 1998 (cooperation system-Scientific Channel) between Claustahl-Zeller Field Teschneche Universtate Germany and Ain Shams University Egypt). He is currently an assistant professor of industry electronics at Faculty of Science-Ain Shams University. His research interests include wavelet theory and its applications in the fields of signal and image processing, as well as optimization techniques based on machine learning and soft computing

**Prof Dr. Abdel-Badeh M Salem** is a professor emeritus of Computer Science since September 2007 till now. He was a former Vice Dean of the Faculty of Computer and Information Sciences at Ain Shams University, Cairo-Egypt (1996-2007). He was a professor of Computer Science at Faculty of Science, Ain Shams University from 1989 to 1996. He was a Director of Scientific Computing Center attain Shams University (1984-1990). His research includes intelligent computing, expert systems, medical informatics, and intelligent e-learning technologies. He is author for many books and co-author of 15 books in English and Arabic languages. He has published around 200 papers in refereed journals and conference proceedings in these areas. He has been involved in more than 200 conferences and workshops as an Int. Program Committee, organizer and Session Chair. He is author and co-author of 15 Books in English and Arabic Languages. He is a member of the Editorial Board of many journals.

# Control Flow Hardening with Program Counter Encoding for ARM<sup>®</sup> Processor Architecture

Seho Park, Jongmin Lee, Yongsuk Lee, and Gyungho Lee

**Abstract**— Software operations are prone to *semantic gap*, which refers to potential difference between intended operations described in software and actual operations done by processor. Attacks that compromise program control flows, which results in the semantic gap, are a major attack type in modern software attacks. Many recent software protection schemes focus on protecting program control flows. But even without challenging the protection scope of the schemes, most of them suffer not only from the deployment difficulties but also from critical performance issues. This paper uses a program counter (PC) encoding technique (*PC-Encoding*) to harden program control flows under ARM<sup>®</sup> processor architecture. The PC-Encoding directly encodes control flow target addresses that will load into the PC. It is simple and intuitive and to implement and incurs little overhead. Encoding the control flow target addresses can minimize the semantic gap by preventing potential compromises of the control flows. This paper describes our efforts of implementing PC-Encoding to harden portable binary in ELF (Executable and Linkable Format). Our LLVM (low level virtual machine) based PC-Encoding compiler provides control flow protection with little overhead for programs running under Arm<sup>®</sup> processor architecture.

**Keywords**— *Key-Words:* Compiler, Control Flow Integrity, PC-Encoding, Software Security, ARM<sup>®</sup>.

## I. INTRODUCTION

Software security has become increasingly important concern with the prevalent use of the Internet. With a looming popularity of the Internet of Things (IoT), the concern becomes more prevalent in every aspect of modern life. Various security techniques have been researched and developed at the software level, and some of these have actually been adopted in practice. However, attackers are able to continue to find vulnerabilities, and the attacks succeed more or less in the same way as used to be but with more alarming rates. The number of vulnerabilities utilized for attacks is increasing, and the risks inherent in the security vulnerabilities have become even more serious.

A major portion of the vulnerabilities allows memory overwrite, which in turn causes program control transfer in a

way the programmer did not intend. This *semantic gap*, i.e., potential difference between intended operations described in software and actual operations done by processor can exist as long as there is a vulnerability allowing a memory overwrite.

There have been various techniques for preventing arbitrary memory overwrite along with schemes for preventing the activation of injected attack code, e.g., many variations of DEP (Data Execution Prevention) [13],[18]. However, software attacks utilizing memory overwrite keep appearing. A recent sophisticated attack method is ROP (Return Oriented Programming), a code reuse attack that does not need to inject attack code but needs a memory overwrite to start an attack with the code already existing in the memory [2],[3],[4],[7],[8]. Note that a memory overwrite is essential not only in traditional code injection attacks but also in recent code reuse attacks to divert program control flows from the intended operations described in software.

In order to preserve the control flow from software attacks, we need to enforce the integrity of control flows by examining the destinations of all control flow transfer instructions to see whether they are legitimate or not. Conceptually one can generate a control flow graph (CFG) for a program under protection and check at run time to see if program execution actually follows the CFG. Even without contemplating the accuracy of the CFG, a couple of issues arise from this conceptual control flow validation scenario in terms of the granularity of the CFG along with representing and storing the CFG. System call level CFG has been utilized in early days for the sake of reducing the overhead of representing and storing the CFG and also of accessing the CFG at run time [9]. However, such a coarse grain CFG often fails to catch a compromised control flow because there may exist many hundred or thousand instructions exist between the calls.

The CFI (Control flow integrity) [1],[16],[23] is one of the early proposals to check and validate each control flow transfer at fine grain machine instruction level; force the destinations of all indirect branch instructions to be checked. The full CFG with a pair of IDs for all indirect branch instructions, branch instruction address ID and its target address ID, is generated by binary patch, and each control flow transfer is checked at run time per the CFG. However, even if the CFG is limited to statically linked procedures, it is very difficult to draw a full complete CFG in practice and will cause a serious performance overhead for representing and accessing the CFG. Basic

This work was supported in part by the National Research Foundation (NRF 2015R1A2A01003242).

Seho Park is with the Dept. of Computer Science and Engineering, Korea University, Seoul, Korea (e-mail: pssshh6@gmail.com).

Jongmin Lee is with the Dept. of Computer Science and Engineering, Korea University, Seoul, Korea (e-mail: flackekd@korea.ac.kr).

Yongsuk Lee is with the Dept. of Computer Science and Engineering, Korea University, Seoul, Korea (e-mail: duchi@korea.ac.kr).

Gyungho Lee is with the Dept. of Computer Science and Engineering, Korea University, Seoul, Korea (e-mail: ghlee@korea.ac.kr).



implementation of the CFI utilizes a fact that code region is safe from memory overwrite and also coalesces the instruction addresses together and also the target addresses together to make the CFG representation compact and fast to access. Also there have been some CFI implementation methods with looser notion in order to improve the ease of deployment and less performance overhead [24],[25]. However, the coalesced ID for the addresses and looser notion of the CFG will naturally increase the semantic gap potential, reducing the robustness of the protection [10].

*PC-Encoding* can protect the general indirect branch instructions as the CFI does but without generating the CFG [11],[12],[17]. Instead, the PC-Encoding encodes the target address at its definition and requires a proper decode of the target address for a legitimate control flow transfer. PC-Encoding has little room for the semantic gap because the CFG is not generated as an approximation of what is described in the program but the program itself acts as the CFG. And its checking of the legitimacy for each control flow transfer needs just a couple of machine instructions without the need of accessing the memory for the CFG. It can be with minimum performance degradation.

This paper describes our efforts to implement the PC-Encoding to harden portable binary in ELF (Executable and Linkable Format). Our LLVM (low level virtual machine) based PC-Encoding compiler provides control flow protection with little overhead for programs running under ARM® processor architecture [5]. Our implementation suggests that the PC-Encoding fits well for a relatively simple architecture environment of ARM® processors with low performance degradation.

Section II discusses the basics of PC-Encoding. The process of applying the PC-Encoding for *ARM-Linux-elf* binary is in Section III. Section IV presents the performance test results from Gem5 simulator. This paper closes in Section V with a discussion on the limitations of our current implementation of PC-Encoding for ARM® processors.

## II. BACKGROUNDS

Program control-flow is dictated by program data loaded to the program counter at runtime, which we call the *PC-bound data*. The basic idea of the *PC-Encoding* is to check the integrity of the PC-bound data. The PC-Encoding ensures the integrity of program control flow by protecting the PC-bound data. The PC-Encoding encodes the PC-bound data at their definition and decode at their use with a secret key. The PC-Encoding makes a good solution to harden program control flow for embedded platforms because it has little performance penalty. Data encoding/decoding computation can be made to be simple. Simple encoding/decoding operation that can be done in one or two cycles may be employed, e.g. exclusive-or. Also, there is no compatibility issue. It doesn't need to change memory layout. Moreover it can cooperate well with not hardened binary.

The PC-Encoding encodes destination of indirect jump instructions such as return addresses on stack, function addresses on GOT, function pointers, or exception handlers. But it is impossible to overwrite hard coded destination in direct jump. Therefore, direct jumps are excluded from scope of the PC-Encoding. Fig. 1(a) shows normal code parts in object file. Memory overwrite attack can occur during execution time of the *Some code*. If an attacker contaminates the PC-bound data during the *Some code* execution, the “*mov pc, {PC-bound data}*” instruction will take the control flow to illegal location. However, Fig. 1(b) shows hardened case. The PC-bound data is already in encoded state by the Key during the *Some code* execution, so attacker can't overwrite PC-bound data with intended value. Consequently, the Key is a value that the attacker must know for a successful attack.

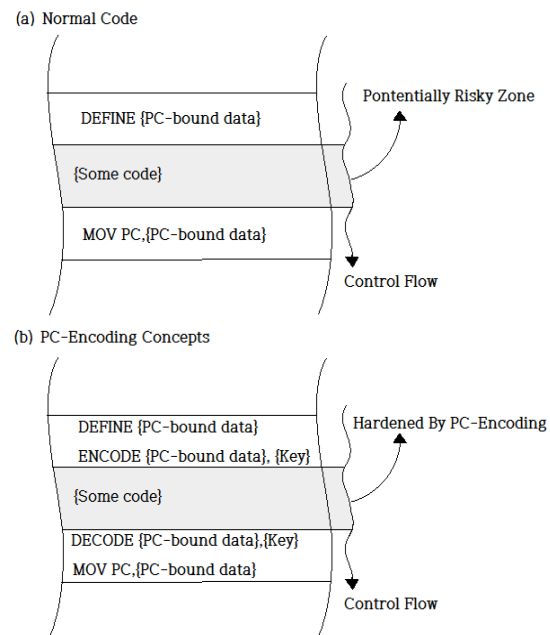


Fig. 1 PC-Encoding Concepts

In general, the key must be stored in a recoverable point at the time of verification. Typical places that can be mentioned as a point of storage is memory. However, if the key is stored in memory, the key itself can be vulnerable from the memory overwrite attack. This dilemma can be found in some protection schemes. e.g., *StackShield* copies the return address of the function to special purpose area, called *Global Ret Stack* or *Shadow RET Stack* [21]. And then it performs integrity check by comparing the two values at the epilogue of functions. However, the Shadow Ret Stack is also dynamically writable memory area. And it can still be a target of the memory overwrite attack. So some schemes are supported by low level software like kernel to ensure read-only property of the key storage. These solutions maintain the key storage as read-only area during a code execution under protection, and temporarily change storage permission to writeable at the time of defining PC-bound data, e.g., *mprotect()* function can change memory

permission dynamically, but this function also can be the victim of RTL attack[14],[19],[22],[26]. As a result, `mprotect()` function can become a more serious vulnerability. Moreover, frequent dynamic change of memory permission will cause critical performance penalty.

The value used as the key in this paper is the “self-address” of the PC-bound data. The *self-address* is the address of the memory location containing the PC-bound data. The self-address points a permanent space that is programmable but requires no additional storage to store it. And it is always maintained as a part of the value-address pair. Therefore it is possible to extract the key easily at the moment of decoding or encoding the target addresses. Also it is not possible to compromise the self-address with the memory overwrite attack, because address is not in the memory. Memory layout for different address spaces can be allocated with some level of randomness in ASLR environment [15],[20]. And the allocation order of function frames on the stack space can vary depending on the dynamic execution flow of the process.

One can apply the PC-Encoding to all indirect calls including call and returns. However, it is difficult to locate the exact locations of all indirect calls except a few stylized cases such as GOT entries, function returns and exception handlers. This paper focuses on the basic implementation under ARM® processor environment. With the fact that call/return pairs are the most frequent indirect branches, this paper focuses on encoding and decoding return addresses as our first attempt for realizing PC-Encoding compiler for ARM® processors.

Various high level languages such as C, C++, Fortran and Ada can benefit from PC-Encoding. PC-Encoding adds a few instructions into a binary executable in order to harden the binary executable. There are three ways one can add new instructions into a binary executable. The first is a runtime modification (*binary editing*). With the binary editing at run time it is possible to damage the functionality of the binary executable. Nevertheless, it may have the highest utility in the sense that it is independent from the languages the code is written. The second is a binary patch. Binary patch also is of a high utility, because it requires no source code. It is possible to insert a protection patch by using only the executable binary without depending on the type of high-level languages. However, relocation of the addresses due to the inserted code patch can be a difficult problem, because many clues for understanding the programmer's intent would have disappeared in the binary executable. Third is the modification at the compile time. Adding additional code of security mechanism at the compile step has a disadvantage, i.e., dependency on the type of high-level languages. In the worst case, it needs to modify compilers manually as many as the number of high-level languages used to extend the coverage of adoption. However, it can add a protection in a relatively reliable way by using an internally validated compiler library. This property leads to a guarantee of the functionality of the modified program.

In this paper, we have applied our PC-Encoding at the compile time. A compiler infrastructure, LLVM, was used to complement for the limitations mentioned above. The LLVM aims to operate independently from high level language and the architecture. The LLVM is with the *frontend* separate from code generation to eliminate the dependency with a specific high level language. The frontend converts a high level language code into an intermediate language code called the LLVM IR, e.g., the Clang is a typical frontend for the C language. By utilizing these features, codes written in various types of high level language can be easily hardened with the PC-Encoding.

In order to understand this paper, it is necessary to know the Arm® processor Architecture. The Arm® processor stores the return address at *lr*(Link Register) when it calls a function by *bl*(Branch Link) instruction. In the prologue of functions, *lr* and *fp* (frame pointer Register) are pushed into the stack. At the epilogue of functions, *fp* and *pc*(program counter register) are restored.

### III. IMPLEMENTATIONS

Fig.2 is implementation overview about generating hardened ARM-Linux-elf binary. This walkthrough consists of three steps. In step 1, it converts the code written in high level language to LLVM IR code. In step2, it creates the hardened Assembly file with LLVM IR file. The LLC is a tool that can convert LLVM IR files into a specified architecture assembly file. Also it is possible to adjust options to make other architecture's assembly file. In this paper it make assembly file for ARM® processors. In Step 3, it is possible to determine whether the appropriate patches or collect the number of added instructions by analyzing the assembly file. And then, Step 3 convert assembly file to binary object file through ARM®

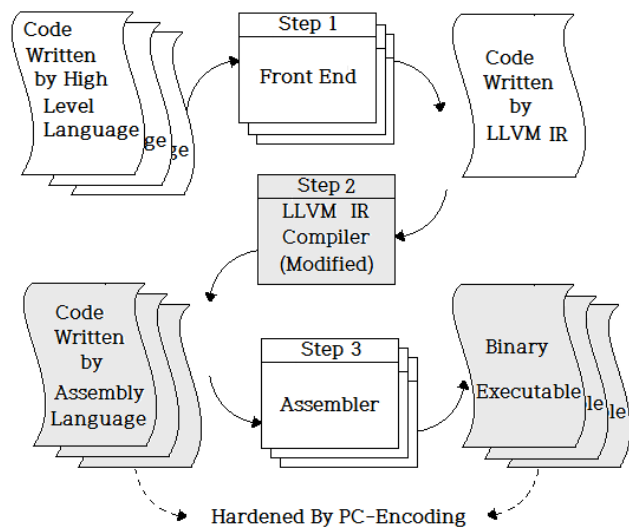


Fig. 2 Implementation overview

processor Assembler.

In step2, the implementation uses a modified LLC, named *LLC.PCE*. The *LLC.PCE* will apply the PC-Encoding to LLVM IR file regardless of the high level language. Among the LLVM source code, the *ARMFrameLowering* class is responsible for code generation of the function frame in the ARM® environment. The *emitPrologue* and the *emitEpilogue* functions in the *ARMFrameLowering* Class was edited for the PC-Encoding.

Prologue and epilogue that the *gcc* compiler generates in general for the ARM® environment is in the below.

*Prologue:*

```
push {fp,lr} //Save frame pointer and return
add sp,sp,$n //Allocate space for local variables
```

*Epilogue:*

```
mov sp,fp //Move stack pointer to stack base
pop {fp,pc} //Restore frame pointer and return
```

In above code, the *lr* register is stored at stack in prologue. It is mean that *lr* register can be overwritten by memory overwrite attack. The PC-Encoding can insert encode and decode instructions for protect *lr*.

*Prologue:*

```
eor lr,lr,sp //Encode return address
push {fp,lr} //Save frame pointer and return
add sp,sp,$n //Allocate space for local variables
```

*Epilogue:*

```
mov sp,fp //Move stack pointer to stack base
ldr lr,[sp,#4] //Load return address at lr
eor lr,lr,sp //Decode return address
str lr,sp //Save decoded return address
pop {fp,pc} //Restore fp and return
```

In the above code, four instructions are added for the PC-Encoding. In this case, it is impossible to avoid memory access at the decode, because “*pop {fp,pc}*” is to take place in atomic fashion. The return address moves to the PC register directly. The PC-Encoding must decode return address before this action. Thus decode process must be accompanied by a memory access. To solve this problem, we use the LLVM features. In fact, compilers do not always need to make the form of a prologue and epilogue shown in the above. The LLVM compiler can create a different form of function prologue and epilogue as in the below.

*Prologue:*

```
push lr //save return address
add sp,sp,$n //Allocate space for local variables
```

*Epilogue:*

```
add sp, sp, $n //Mov stack pointer to base
pop lr //Restore return address
mov pc,lr //Return to caller
```

In this code, there is no *fp* register. Offsets are calculated based on the *sp* instead of using *fp* when accessing local variables or function parameters. And the return address is restored at *lr* before it moves into *pc*. By using this moment, it is possible to remove a memory access in decode process. The PC-Encoding example implemented in this consideration is as follows.

*Prologue:*

```
eor lr,lr,sp // Encode return address
push lr // Save encoded return address
sub sp,sp,$n //Allocate space
```

*Epilogue:*

```
add sp,sp,$n //Move stack pointer to base
pop lr //Restore return encoded address
eor lr,lr,sp //Decode return address
mov pc,lr //Return to caller
```

Two added *eor* instructions have no memory reference. As a result, we are able to protect return process of a function with just two of register-to-register instructions. Encoding the target addresses can avoid unintended control transfer causing the semantic gap because it allows the control transfer only to a target address legitimately decoded.

In fact, at the function exit, the main intention of a programmer is just a “*return to the caller*”. The status of the registers and the memory are not always in programmer’s consideration. However, it is possible to manipulate the return address using a frame pointer overflow. Compilers and its library do not perform action to correct this potential semantic gap source. Since *fp* register is also pushed to stack, it can be exploited. So both *lr* and *fp* registers must be encoded for the full protection of return process in the *gcc* version.

The PC-Encoding implementation depends on trustworthy low layer software like the OS kernel. If the code were contaminated by an arbitrary memory overwrite attack, it is not possible to protect the PC-bound data with PC-Encoding. But it does not mean that the PC-Encoding needs special functions like the *mprotect()*.

#### IV. PERFORMANCE

Comparing the efficiency of PC-Encoding with the well-known CFI [1] is illustrating to see the performance overhead potential. One can patch the code with ARM® processor assembly to have the CFI enforced as in the below.

*caller:*

```
bl callee //Call the callee
b next //Bypass the ID
```

```

[AABBCCDD] //block ID
next:
.
callee:
    push lr          //Save return address
    sub sp,sp,$n     //Allocate space for local variables
    .
    add sp,sp,$n     //Move stack pointer to base
    pop lr          //Restore return address
    ldr r3,[lr,#4]   //Obtain ID of return block
    cmp r3,AABBCCDDh //Check validation
    bne error_label  //Error handling
    mov pc,lr        //Return to caller(next)

```

When implementing the CFI on ARM® processor, the caller needs one and the callee needs three additional instructions. One of them is a memory reference instruction, and the other one is conditional branch instruction. Memory reference instruction will take longer than the others, and the conditional branch instruction also requires more cycles than a normal command requires. Consequently, it will causes the relatively large performance degrade. Table.1 shows the number of

	added instruction	of memory access	conditional branch
CFI	4	1	1
PC-Encoding	2	0	0

Table. 1 Number of instructions in CFI and PC-Encoding

instructions between the CFI and the PC-Encoding.

ROP (Return Oriented Programming) [19] is a latest powerful code reuse attack that can bypass the defenses such as DEP, ASLR and ASCII-Armor [26]. Attackers can gather fixed position code parts, called gadget. These fixed position address can be selected with not containing NULL byte. The Attacker can performs a desired operation without executing the code on the NX(Non-Executable) area using these gadgets. The overwrite attack and advanced RTL attack that jumps to fixed PLT entries use some the ROP gadgets, and it allows the attacker seize control of the system more easily.

First of all, these advanced memory overwrite attack should start from the contamination of one PC-bound data. Canary and ASCII-Armor can take effect against the linear memory overwrite attack (smashing attack). But there are many ways to overwrite the PC-bound data. Arbitrary overwrite attack can bypass the defense such as StackGuard [6] because the guard is separate from the return address. The PC-encoding protects the target directly, i.e., the target address and its protection is integrated together as a single data: it is still a useful protection whether the attacker uses a point for indirect overwrite or does straight smashing. Thus, the PC-Encoding can prevent the first intended jump using the target PC-bound data. ROP attack sequence occurring after overwriting the first PC-bound data is blocked by the PC-Encoding environment.

The PC-Encoding provides additional benefits against ROP. The PC-Encoding can eliminate the ROP gadgets by inserting decode instruction in front of the return instructions. Typical gadgets in a pattern of *pop-pop-ret* will be transformed into a *pop-pop-eor-ret* pattern, and thus attackers should be able to guess the key in order to use the gadgets. In Intel x86 architecture, inserting instruction into the gadgets may cause a side effect of unexpected instructions that can be exploited for ROP attack, because unaligned instructions different from the programmer's intention can be fetched and executed with Intel x86 architecture. But only the 4byte aligned instructions are allowed to execute with ARM® processor. It means that inserting instructions to eliminate the gadget can be more reliable and clear solution with ARM® architecture.

This paper assumes the following instruction sequences as potential gadgets.

```

{Several instructions}; b ;
{Several instructions}; bl ;
{Several instructions}; bx ;
{Several instructions}; blx ;
{Several instructions}; bxj ;
{Several instructions}; pop {pc,...};
{Several instructions}; mov pc, {reg/mem};

```

Binaries compiled by the LLVM rarely have the "*pop {pc,...}*" instruction, because the LLVM compiler does not use "*pop {pc,...}*" instruction as a return instruction. Therefore, the PC-Encoding removes most of the gadgets that use the "*mov pc,{reg/mem}*" instruction. However, "*b,bl,bx,blx,bxj*" gadgets are not eliminated by the current PC-Encoding implementation reported in this paper. The implementation in this paper covers only return address case. So attacks using other PC-bound data can still valid. Also this narrow protection scope causes un-elimination of "*b,bl,bx,blx,bxj*" gadgets. But it is an implementation problem rather than a fundamental disadvantage inherent in the concept of PC-Encoding. Our future implementation will be able to handle all the indirect branches as done in our PC-Encoding compiler for Intel x86

name	Simulated instructions (million)		Overhead (instruction)	Overhead (tick)
	normal	PC-Encoding		
mcf	8953	9128	1.95%	1.93%
sjeng	33823	34170	1.02%	1.00%
name	Input		Added instructions	
			Encode	Decode
bzip2	dryer.jpg		110	120
mcf	test/input.in		24	24
sjeng	test/test.txt		141	144

Table. 2 Performance result

[11], [17].

Table.2 shows the performance results with SPEC2006 benchmark binaries hardened by the PC-Encoding. The simulations are performed with default SE mode of Gem5 simulator [27]. The SE mode of the Gem5 can simulate not all but most system calls of Linux. It shows the performance overhead is under 2%. Also shown in Table.2 is the number of encode/decode instructions added by our PC-Encoding compiler. The numbers of encode and decode instructions are not the same, because a must-terminate epilogue terminates the process directly without executing return instruction, and a function may have multiple epilogue upon branch instructions. Unfortunately, the SE mode of Gem5 does not simulate the all Linux system calls. Only a few of the binaries can be simulated properly on the Gem5. For example, the bzip2, software for compression, needs to call a *utime()*, a system call not included in the Gem5. So the bzip2 execution on Gem5 cannot be finished properly.

## V. CONCLUSION

This paper has shown a guide to make practical PC-Encoding compiler implementation under Arm® processor architecture. Arm® processors have become the most popular embedded and application processor architecture, widely adopted in embedded devices including the smart-phones. Based on our experience of building PC-Encoding compiler for Intel x86 architecture [11],[17], we are implementing PC-Encoding compiler for ARM® processor. By utilizing the features of Arm® processors, our implementation protects the return addresses without inserting a protection code accessing the memory. Therefore, the protection is with minimal performance overhead.

In our implementation presented in this paper, the key for the PC-Encoding is the self-address, i.e., the location of the instruction defining PC-bound data. It allows a low overhead implementation. However, if the degree of ASLR's randomization is weak, it is possible for attackers to easily figure out the key via replay attacks and source code analysis. For more tight protection we may want to resort on utilizing harder to guess cryptographic keys [17].

The PC-encoding implementation presented in this paper is not able to prevent every type of attacks, because it focuses only on the return addresses among many types of the PC-bound data. Many software attacks not resorting on the return address exist, and a few techniques have been proposed to solve this problem. For example, it is possible to insert decoding instructions into PLT region to prevent GOT overwrite attacks with the encode instruction added to *dl\_resolve()* function [11]. Also, one may utilize the relocation table to find every indirect jumps as in CCFIR [25]. We are currently in the process of incorporating these into our PC-Encoding compiler for ARM® processors.

## REFERENCES

- [1] M. Abadi, M. Budiu, U. Erlingsson, and J. Ligatti, "Control-Flow Integrity: principles, implementations, and applications", In Proceedings of the 12th ACM Conference on Computer and Communications Security, 2005.
- [2] T. Bletsch, X. Jiang, V. W. Freeh, and Z. Liang, "Jump-oriented programming: a new class of code-reuse attack," in Proceedings of the 6th ASIACCS, pp. 30–40, March 2011.
- [3] E. Buchanan, R. Roemer, H. Shacham, and S. Savage, "When good Instructions go bad: generalizing return-oriented programming to RISC", In ACM Conference on Computer and Communications Security (CCS), pp. 27–38, 2008.
- [4] S. Checkoway, L. Davi, A. Dmitrienko, A.-R. Sadeghi, H. Shacham, and M. Winandy, "Return-oriented programming without returns", In ACM Conference on Computer and Communications Security (CCS), 2010.
- [5] Computer Science Department at the University of Illinois at Urbana-Champaign. "The LLVM Compiler Infrastructure", [Online]. Available:<http://llvm.org>.
- [6] C. Cowan, C. Pu, D. Maier, H. Hintongif, J. Walpole, P. Bakke, S. Beattie, A. Grier, P. Wagle, and Q. Zhang, "Stackguard: automatic adaptive detection and prevention of buffer-overflow attacks," in Proc. of the 7th USENIX Security Symposium, Jan 1998, pp. 63–78.
- [7] D. Dai Zovi, "Practical return-oriented programming," SOURCE Boston, 2010.
- [8] L. Davi, A. Dmitrienko, A.-R. Sadeghi, and M. Winandy, "Return-oriented programming without returns on ARM", Technical Report HGI-TR-2010-002, Ruhr-University Bochum, July 2010. Available:<http://www.trust.rub.de/home/publications/DaDmSaWi2010>
- [9] S. Forrest, S. Hofmeyr, Anil Somayaji, "The evolution of system-call monitoring", ACSAC, 2008.
- [10] E. Göktas, E. Athanasopoulos, H. Bos and G. Portokalidis, "Out of control: overcoming control-flow integrity", Proceeding SP '14 Proceedings of the 2014 IEEE Symposium on Security and Privacy, pp. 575-589, 2014.
- [11] G. Lee and C. Pyo, "Method and apparatus for securing indirect function calls by using program counter encoding", U.S. Patent No. US8583939 B2, Nov. 2013.
- [12] G. Lee and A. Tyagi, "Encoded program counter: self-protection from buffer overflow attacks", International Conference on Internet Computing, pp 387-394, 2000.
- [13] Microsoft, "A detailed description of the data execution prevention (DEP) feature in Windows XP Service Pack 2, Windows XP Tablet PC Edition 2005, and Windows Server 2003", [Online]. Available: <https://support.microsoft.com/en-us/kb/875352>
- [14] Nergal, "The advanced return-into-lib(c) exploits: PaX case study". Phrack Magazine, 58(4), 2001.
- [15] Pax Team, "PaX address space layout randomization (ASLR)", [Online]. Available: <http://pax.grsecurity.net/docs/aslr.txt>
- [16] J. Powny and T. Holz, "Control-flow restrictor: compiler-based CFI for iOS," in Proceedings of the 29th Annual Computer Security Applications Conference, 2013.
- [17] C. Pyo and Gyungho Lee, "Encoding function pointers and memory arrangement checking against buffer overflow attack", *Lecture Notes in Computer Science*, Vol. 2513, pp. 25 - 36, Springer Berlin / Heidelberg, Dec. 2002.
- [18] Red Hat, "New security enhancements in red hat enterprise Linux v.3, update3", [Online]. Available:<http://h10032.www1.hp.com/ctg/Manual/c00387685.pdf>
- [19] H. Shacham, "The geometry of innocent flesh on the bone: return-into-libc without function calls (on the x86)", in ACM conference on computer and communications security (CCS), pp.552–561, 2007.
- [20] H. Shacham, M. Page, B. Pfaff, E. Goh, N. Modadugu, D. Boneh, "On the effectiveness of address-space randomization." In V. Atluri, B. Pfitzmann, and P. McDaniel, editors, Proceedings of the 11th ACM

- Conference on Computer and Communications Security, CCS 2004, Washington, D.C. ACM, October 2004.
- [21] S. Sinnadurai, Q. Zhao, W. Wong, "Transparent runtime shadow stack: protection against malicious return address modifications".
  - [22] M. Tran, M. Etheridge, T. Bletsch, X. Jiang, V. Freeh and P. Ning , " On the expressiveness of return-into-libc attacks", Proceeding RAID'11 proceedings of the 14th international conference on recent advances in Intrusion Detection, pp. 121-141, 2011.
  - [23] B. Zeng, G. Tan, and G. Morrisett, "Combining control-flow integrity and static analysis for efficient and validated data sandboxing," in proceedings of the 18th ACM conference on Computer and Communications Security, pp. 29–40, 2011.
  - [24] M. Zhang, R. Sekar, "Control flow integrity for COTS binaries," in 22nd USENIX Security Symposium, 2013.
  - [25] C. Zhang, T. Wei, Z. Chen, L. Duan, L. Szekeres, S. McCamant, D. Song, and W. Zou, "Practical control flow integrity and randomization for binary executables," in Proceedings of the 1013 Security and Privacy Symposium, pp. 559–573, 2013.
  - [26] "Payload already inside: data reuse for ROP exploits", Black Hat USA Whitepaper, 2010.
  - [27] "Gem5 simulator", [Online]. Available: <http://www.gem5.org>

# Wide-Area Damping Controller Design Aiming at Robust Performance of the Controller

M. Beiraghi and A.M. Ranjbar

**Abstract**—the main bottleneck in effective design of wide-area damping controller (WADC) is the power system operating condition changes and uncertain time delay of remote signals. They can deteriorate the controller performance and the whole system stability if not properly accounted for in the design procedure. This paper focuses on design a WADC pointing at robust performance in the face of both power system model and time delay uncertainties. The proposed method is based on structured singular value (SSV) theory and the controller is designed by applying  $\mu$ -synthesis and D-K iteration approaches. The SSV theory provides the correct test for robust performance and is an effective means to handle model uncertainties and destructive effects of uncertain time delay. A comprehensive case study is conducted on a 68-bus 16-machine test system including SSV assessment, small signal analysis, and nonlinear time domain simulation. The results reveal that the proposed method provides effective damping action when the operating condition changes and the time delay varies within a certain range.

**Index Terms**—WADC, time delay, model uncertainty, inter-area oscillations

## I. INTRODUCTION

INTER-area oscillations of a bulk power systems usually suffer from poor damping [1]-[3]. The inadequacy of damping varies with the level of power transfer through tie-lines, strength of tie lines, nature of loads and other related factors. [4]. It might have dissuasive effects and prevents the operation of the power system in a desired economic operating conditions. Therefore, it is essential to employ an effective damping control strategy that provides adequate damping over pre-defined operating scenarios. The traditional approach for damping electromechanical oscillations is to install power system stabilizers (PSSs) which use local measurements such as rotor speed or active power as feedback signals. These PSSs can damp local modes effectively, while the effectiveness in damping inter-area modes is reduced because such modes are not observable/controllable from local measurements [6].

With a rapid development of wide-area measurement system (WAMS), controllers have been recommended which use wide-area signals to enhance damping of electromechanical oscillations. Wide-area damping controller (WADC) was primarily proposed in [3] as a two level hierarchical controller, in

which a local signal is used to damp local mode and remote signal feedback to damp the inter-area mode. Afterwards, the power system model uncertainties and inherent time delay resulting from wide-area signals transmission have been challenging researchers to make WADC practical.

Time delay of wide-area signals is a key factor influencing the applicability of WADC and the whole system stability. The delay value depends on the type of communication link, the distance, signal routines, protocol of transmission, communication load and several other factors [8]. In the literature, different approaches have been proposed to consider time delay effects in controller design procedure. A gain scheduling based method has been investigated to design a WADC modeling the delay by a first order approximation [9]. This method results a parameter varying high order controller subject to delay variation. Delay compensation by a predictor based controller has been suggested in [4]-[5], in which the predictor is designed in a predefined operating condition. In [10] and [11], the authors have been employed Lyapunov stability theory and linear matrix inequalities (LMIs) approach to design WADC considering delayed communication network. Stability analysis of power system with WADC embedded has been carried out applying a set of LMIs to obtain delay margin of different controllers [1]. In [12], conventional lead-lag controller has been designed conducting particle swarm optimization approach to compensate the time lags caused by remote signal delays. In [13], a fuzzy logic based WADC has been proposed in order to have continuous delay compensation.

The system model uncertainty is another key factor in WADC realization and it mainly results from the variation of apparent linearized plant parameters as the operating point changes and the neglected high order dynamics. Robust methods such as  $H_\infty$  controller [4], [5], multi agent  $H_\infty$  controller [7] and mixed  $H_2/H_\infty$  controller [6] have been employed in order to handle these uncertainties.

Although several methods have been proposed to design WADCs aiming at one of these circumstances, less attention has been devoted to both model and delay uncertainties. Moreover, these methods do not ensure the robust performance of the controller and performance judgment is based on a number of time domain simulations.

This research focuses on a WADC design aiming at the robust performance in the face of operating condition changes and uncertain time delays. The method is based structured singular value (SSV) theorem and  $\mu$ -synthesis approach and it can handle the both power system model and time delay

The authors are with the Electrical Engineering Department, Sharif University of Technology, Tehran 11365-11155, Iran (e-mail: mojtaba\_beiraghi@ee.sharif.edu; amranjbar@sharif.edu).

uncertainties. The main procedure of the proposed method includes: 1) to obtain linearized model of the power system and small signal analysis; 2) to select measurement signals and control locations which maximize the joint controllability or observability measure of the critical inter-area modes; 3) to reduce linearized system model order for the sake of controller design simplification; 4) to design the controller by applying  $\mu$ -synthesis approach; 5) to evaluate robust performance of the controller applying SSV theory; 6) to carry out small signal analysis of the closed loop system and nonlinear time domain simulations. The proposed method is prosperously tested on a 68-bus 16-machine test system and various case studies are conducted to demonstrate the potential of the controller. The results indicate that this method is highly efficient to damp inter-area modes when the time delay varies within a certain range.

## II. WIDE-AREA DAMPING CONTROLLER STRUCTURE

Several control structures such as decentralized, quasi-decentralized centralized and hierarchical control schemes are proposed to enhance the electromechanical oscillations damping in power systems [15]. A decentralized structure is a set of local controllers. This control structure utilizes local measurements and needs no additional telecommunication equipment. However, it may not be sufficient to meet the requirements of the future power systems which are highly interconnected and stressed. In contrast, other structures need to utilize communication link to transmit remote signals. Any control which requires communication links to either collect the input or to send out the control signal is entitled as a wide-area control systems [16].

It is found that applying remote signals in the controller results enhanced system dynamic performance with respect to inter-area oscillations [3]. Although additional telecommunication equipment is required to realize a wide-area control system, it is still recognized to be more cost-effective compared to installing new control devices. In a quasi-decentralized control structure, the local controller receives some information from remote location; nevertheless, most signals are collected locally. In a centralized structure, the controller receives relevant system measurements and sends control signals to control locations in order to meet the required system performance; for instance, to ensure a minimum damping for a set of inter-area modes. The hierarchical control structure, as shown in Fig. 1 comprises two control levels and comprehends decentralized local controllers (i.e., PSSs) in the first level and the WADC as a central controller in the second level. The central controller receives measured signals from system and sends control signals to the local controllers.

Transmission delays of the remote signals and robustness to the loss of communication link can be a challenge for the application of quasi-decentralized, centralized, and hierarchical control structures. However, the use of remote signals can improve the observability and controllability of inter-area modes; hence, it results in to effective damping of these modes

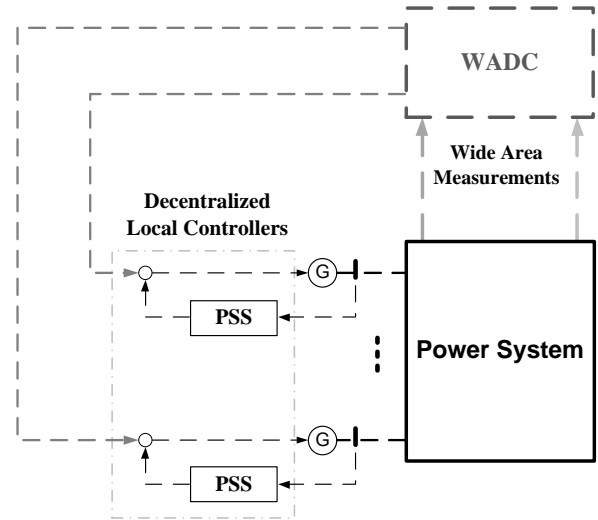


Fig.1: structure of wide-area damping controller

[15]. The hierarchical scheme allows minimum system damping performance ensured by the decentralized level in the event of loss of communication link and is preferred relative to other control structures. Hence in this study, it is applied for the WADC implementation.

## III. CONTROLLER DESIGN APPROACH

Robust controllers were introduced into wide-area damping controller design in the last decade to handle modeling errors and uncertainties. In the robust controller synthesis approach, the controller requirements are weighted and formulated to meet the control objectives. If the control objectives have been satisfied for the nominal plant, the system is stated to have nominal performance. The existence of power system model uncertainty and uncertain time delay make it necessary that the control objectives to be satisfied for the uncertain plant. In other words, the controller should have robust performance. The most frequently used approach to achieve robust performance is to reformulate the problem as a single objective controller design. For instance, the trade-off between requirements has to be made so that the single  $H_\infty$  norm of the weighted control interconnection can represent all design requirements. However, to minimize the performance index, which is obtained by small gain theorem, is not effective in this case since the resulting controller is constantly conservative and achievable closed-loop performance is limited. The controller conservatism is caused by the fact that in the reformulated robust performance problem, the uncertainty is structured and small gain theorem does not account for structure in the uncertainty. The structured singular value (SSV) theory provides the necessary and sufficient conditions for robustness to a structured uncertainty and hence is the correct test for robust performance [17]. The approach of controller design by applying the SSV theory is referred as mu-synthesis. This method is applied to design a wide-area damping controller in this research.

### A. Uncertainties and weighting functions

The power system model uncertainty mainly results from







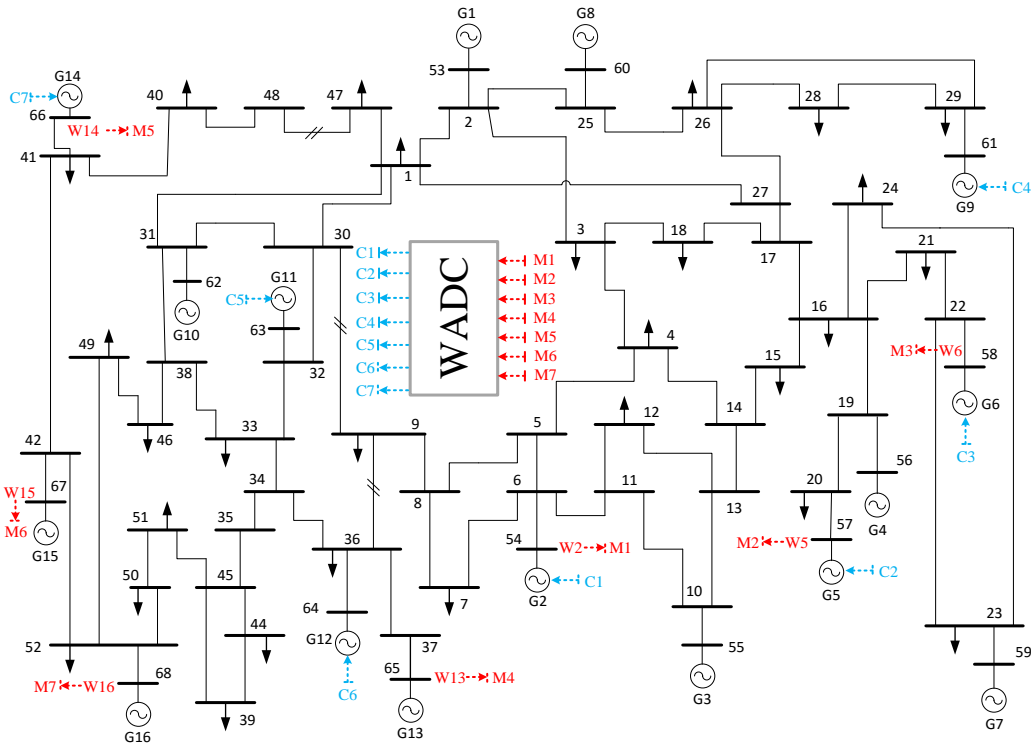


Fig.5: 68-bus 16-machine test system with WADC embedded

## V. RESULTS OF CASE STUDIES

A wide-area damping controller is designed for 68-bus 16-machine system, which is shown in Fig. 5. The detail model descriptions and parameters of this system can be found in [21]. Some modifications have been made so that the controller design procedure can be illustrated more clearly. In this system, each of the generators G13, G15 and G16 is the equivalent model for an area. However, the parameters of G14 is considered to be same as G12. The system is stressed by increasing load and generation level. Moreover, the conventional local PSSs are installed on the generators G3, G4, G9, G10, G11, and G12 and there is no PSS installed on the remaining generators. In addition, damping coefficients of the equivalent generators G13 and G16 have been reduced to some small values.

Small signal analysis reveals that the system has four inter-area modes with damping ratios less than 7%. These modes will last more than 20 sec due to their poor damping. The oscillating generators in these lightly damped modes are shown in Table I. Table II demonstrates the control locations and associated measurement signals which are selected for each mode. In this Table,  $\omega_{i-j}$  is the rotor speed difference between generators  $i$  and  $j$ . The generators with maximum GMC for the first three modes are equivalent generators; hence, it is not possible to select them as control locations. Then again, for these particular modes the maximum GMC of the generators G1-G12 and G14 are not comparable with that of equivalent generators. For instance in mode #2, G9 has the maximum GMC among G1-G12 and it is almost 22% of the G16's GMC. Therefore, the selection of G9 as the control lo-

Table I  
Critical inter-area modes of the study system

Mode index	Mode Shape	Freq. (Hz)	Damping Ratio
1	(14,15,16) vs. all others	0.476	0.0643
2	(12,13,16) vs. (1 to 9) and (14,15)	0.6804	0.0130
3	(12,13,14,15) vs. (1 to 9)	0.7119	0.0371
4	(2,3) vs. (4,5,6,7,9)	1.1141	0.0247

Table II  
Control locations and measurement signals

Mode index	Max. GMC	Control Locations based on AGMC	Max. GMO
1	G16	G9 (22%), G12 (19%), G14 (12%), G11 (10%)	$\omega_{5-15}$
2	G16	G12 (31%), G5 (21%), G6 (18%)	$\omega_{14-16}$
3	G13	G12 (45%), G14 (31%)	$\omega_{5-13}$
4	G2	G2	$\omega_{2-6}$

cation solely, results in a significant increase in the controller effort and excessive interface of local modes. To overcome this problem, the control locations are selected based on the aggregate geometric measure of controllability (AGMC) as indicated in Table II. In which for a particular mode, among G1-G12 and G14 the generators with highest values of GMC are selected in a way that the total GMC to be at least 60% of the equivalent generator's GMC. In mode #1, G9, G11, G12, and G14 are selected as control locations and the speed difference between G5 and G15 is designated as the input signal. Mode #2 suffers from a poor damping and can be controlled through the generators G5, G6, and G12. The best input signal

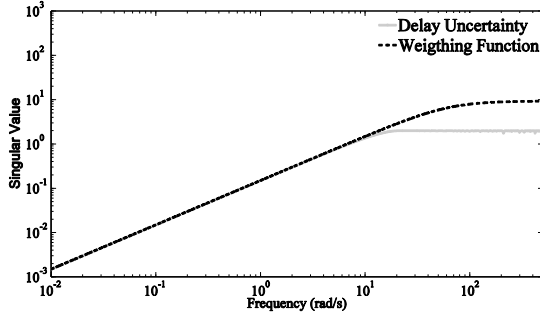


Fig.6: singular value plot of MU and weighting function

Table III

Controller order and peak  $\mu$  value through  $D$ - $K$  iteration

Iteration	peak $\mu$	Controller order
1	2.49	34
2	1.26	45
3	0.98	53
4	0.84	67
5	0.73	77
6	0.71	92

to observe this mode is the speed difference between G14 and G16. Mode #3 that demonstrates the oscillation of G12-G15 vs. G1-G9 can be damped through the control action on the generators G12 and G14. Moreover the measurement signal with maximum GMO for this mode is the speed difference between G5 and G13. Finally, mode #4 in which G2 and G3 oscillates vs. G4-G7 and G9 can be effectively damped out through selecting speed difference between G2 and G6 as the input signal and controlling the generator G2. Therefore, to enhance damping of inter-area oscillations the control locations are G2, G5, G6, G9, G11, G12 and G14.

The original linear model order is 178. To simplify the WADC design, the order of the linear model was reduced to 12 applying *Schur* model reduction technique. In this approach the order of the reduced model is determined based on Hankel singular values of the original model. This reduced model is sufficiently accurate to represent system dynamics in low frequency oscillation studies.

The communication channel which transmits measurement signals to and control signals from centralized WADC is considered to be digital fiber optic link. This type of communication is recognized as the best option due to its high speed data rate, low propagation delay, security, reliability as well as noise immunity. Hence, time delay tolerance of the control loop is expected to be between 50 to 150 ms [11]. The weighting functions in the weighted control design interconnection are given by:

$$W_p = \frac{10}{s+10}, W_n = \frac{2s+2.2}{s+220}, W_s = \frac{5s+2}{s+100}, W_d = \frac{10s}{s+66.67}$$

The MSV plot of the time delay uncertainty which is obtained by (2) and  $W_d$  are shown in Fig. 6. As illustrated in this figure, the selected weighting function satisfies the condition in (3). The *dkit* function available in MATLAB Robust Con-

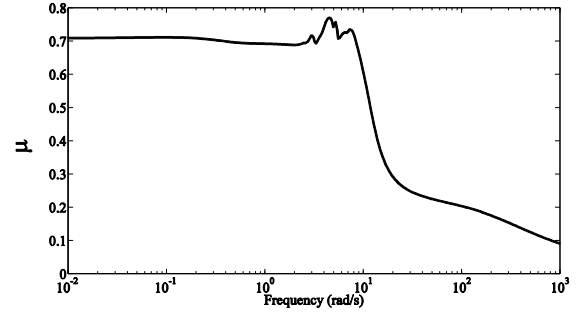


Fig.7: SSV plot of the closed-loop system with reduced order controller

Table IV

Damping ratios of inter-area modes for different load types and time delays

Load Type	Delay (ms)	Damping Ratio			
		Mode 1	Mode 2	Mode 3	Mode 4
CI	50	0.2103	0.1548	0.1452	0.1214
	100	0.1970	0.1334	0.1274	0.1192
	150	0.1810	0.1121	0.1099	0.1171
CC+CI	50	0.1780	0.1529	0.1326	0.1166
	100	0.1759	0.1319	0.1186	0.1158
	150	0.1683	0.1110	0.1037	0.1135
CP+CI	50	0.1719	0.1525	0.1296	0.1156
	100	0.1702	0.1315	0.1165	0.1152
	150	0.1646	0.1108	0.1021	0.1128

troll Toolbox is used to design the controller. The peak  $\mu$  value and controller order for various iterations are represented in Table III. In the first iteration, the peak  $\mu$  value is 2.49 and robust performance condition is not satisfied. The peak value of SSV reduces as the number of iteration increases. The best value of the SSV belongs to iteration 6 and it is 0.71 demonstrating that the designed controller fully satisfies the robust performance requirements. Moreover, this value of SSV indicates that the system model and time delay uncertainties, which can diminish the control performance, to be more than the specified AU and MU in the design procedure. For instance, it is expected that time delays, which can deteriorate the controller performance, to be more than 150 ms. For our case, time domain simulation shows that constant time delays above 230 ms weaken the damping action of the controller.

As shown in Table III, the controller order increases through the  $D$ - $K$  iteration approach. It is 34 in the first iteration and increases to 92 in the iteration 6. Therefore, in order to ease the implementation, it is necessary to reduce the designed controller order. Applying *Schur* technique again the controller order reduces to 12. The  $\mu$  plot of closed-loop system with reduced controller is depicted in Fig. 7. It can be seen from Fig. 7 that the peak value of  $\mu$  is 0.76. Therefore, the reduced controller satisfies the robust performance constraint as well as the full order controller.

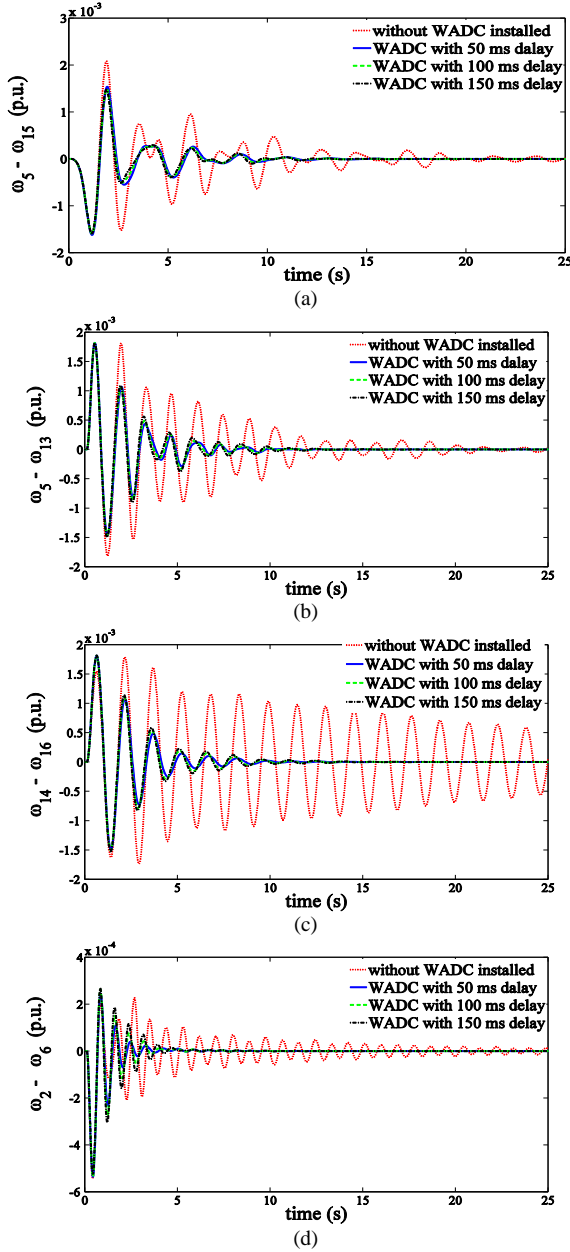


Fig. 8: dynamic response of the system to step changes in AVR set points. (a) & (b) Step change in G13 set point. (c) Step change in G16 set point. (d) Step change in G2 set point.

The damping action of the designed controller is evaluated applying eigenvalue analysis considering various load types and time delays as shown in Table IV. It is noteworthy that in this table, CI, CC, and CP mean constant impedance, constant current and constant power loads, respectively. It can be seen from Table IV that the designed WADC provides effective stabilization action under different load types when the time delay varies within the certain range.

Nonlinear time domain simulations have been carried out for 25 sec to verify the performance of the designed controller in the presence of system nonlinearities and changes of the operating condition. First 0.1 p.u. step changes in the AVR inputs of the generators G2, G13, and G16 were applied as small disturbances. The dynamic response of the system mod-

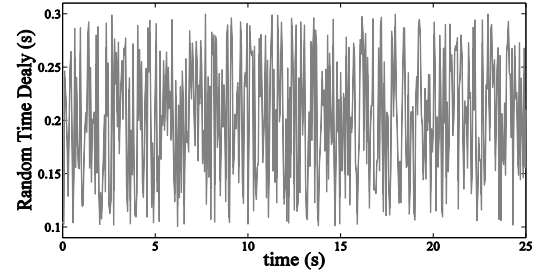


Fig. 9: random time delay of the control loop

el following these changes for different amount of time delays is shown in Fig. 8. The displays in Fig. 8 demonstrate the input signals to WADC. As aforementioned, the effects of inter-area oscillations are manifested most obviously in these particular machine speed differences. It can be seen from Fig. 8 that all inter-area oscillations last more than 25 sec when the WADC is not installed and they are damped out in 6-10 sec through WADC action in the face of different time delays.

To test the performance of the controller in the presence of large disturbances a 3-phase short circuit faults are applied to transmission lines for four cycles. The fault is cleared by taking out the faulted line. It should be noted that the system operating condition changes during post-fault period because of the outage of the faulted line. Most of the transmission lines are tested. Moreover, in this case varying time delay is applied randomly in a way that the total delay is  $200 \pm \text{rand}(100)$  ms, as shown in Fig. 9. The designed controller achieves satisfactory damping outcome for all these fault scenarios considering different time delays. The simulation results to a 3-phase fault applied to the lines 4-14, 49-52, 15-16, and 6-11 are depicted in Fig. 10. The displays in this figure show the controller's input signals. It is observed that inter-area oscillations are poorly damped without the WADC. However, they do not last more than 10 sec through the action of the wide-area controller. Moreover, it is observed that the performance of the controller is nearly the same for different fault scenarios and time delays. This similarity is due to the robust performance of the WADC which has been deliberated in the proposed approach.

## VI. CONCLUSION

A method for robust wide-area damping controller design based on  $\mu$ -synthesis approach was presented. The proposed method aimed at robust performance of the controller in the face of uncertainties. The both power system model and time delay uncertainties were accounted for in design procedure. A comprehensive case study was conducted on a 68-bus 16-machine benchmark system. Applying SSV theorem it has been demonstrated that the designed controller attains robust performance. Moreover, the response of the closed loop system was verified through small signal analysis and nonlinear time domain simulations. The results have confirmed that the designed controller can effectively damp multiple inter-area oscillations triggered by either small or large disturbances very well. Even though different constant and random time delays

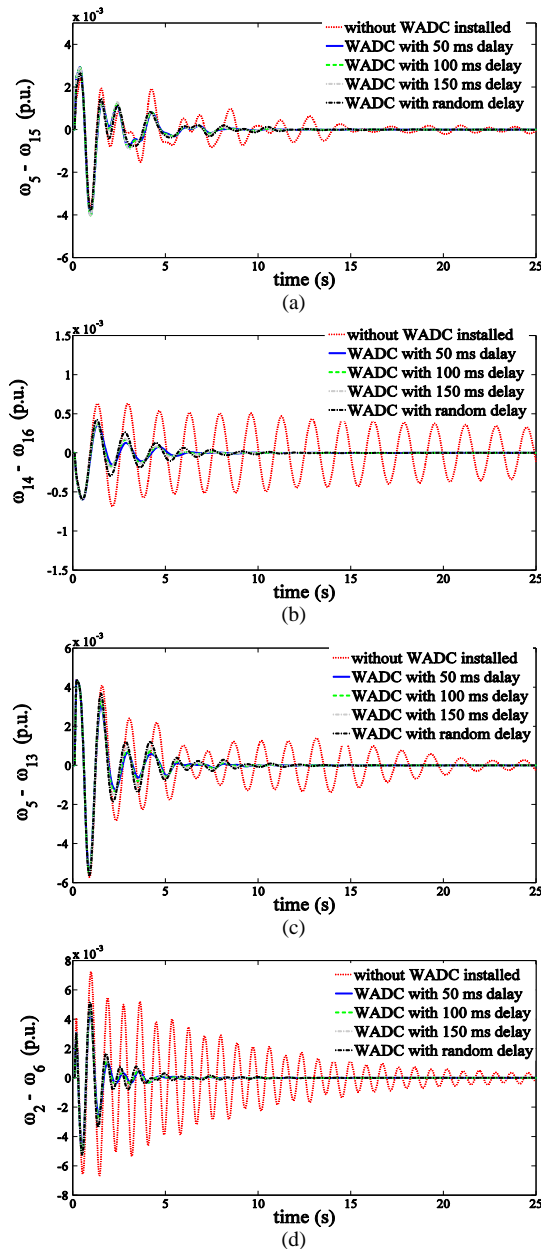


Fig. 10: dynamic response of the system to three-phase faults. (a) Fault applied to line 4-14. (b) Fault applied to line 49-52. (c) Fault applied to line 15-16. (d) Fault applied to line 6-11.

affect the control loop, it has preserved the sought stabilizing action.

#### REFERENCES

- [1] Yao, W., L. Jiang, Q. H. Wu, J. Y. Wen, and S. J. Cheng. "Delay-dependent stability analysis of the power system with a wide-area damping controller embedded." *Power Systems, IEEE Transactions on* 26, no. 1 (2011): 233-240.
- [2] P. Kundur, *Power System Stability and Control*. New York: McGraw-Hill, 1994.
- [3] Aboul-Ela, Magdy E., A. A. Sallam, James D. McCalley, and A. A. Fouad. "Damping controller design for power system oscillations using global signals." *Power Systems, IEEE Transactions on* 11, no. 2 (1996): 767-773.
- [4] Majumder, Rajat, Balarko Chaudhuri, Bikash C. Pal, and Qing-Chang Zhong. "A unified Smith predictor approach for power system damping control design using remote signals." *Control Systems Technology, IEEE Transactions on* 13, no. 6 (2005): 1063-1068.
- [5] Chaudhuri, Balarko, Rajat Majumder, and Bikash C. Pal. "Wide-area measurement-based stabilizing control of power system considering signal transmission delay." *Power Systems, IEEE Transactions on* 19, no. 4 (2004): 1971-1979.
- [6] Zhang, Yang, and Anjan Bose. "Design of wide-area damping controllers for inter-area oscillations." *Power Systems, IEEE Transactions on* 23, no. 3 (2008): 1136-1143.
- [7] Ni, Hui, Gerald Thomas Heydt, and Lamine Mili. "Power system stability agents using robust wide area control." *Power Systems, IEEE Transactions on* 17, no. 4 (2002): 1123-1131.
- [8] Naduvathuparambil, Biju, M. Valenti, and Ali Feliachi. "Communication delays in wide-area measurement systems." In *Southeastern Symposium on System Theory*, vol. 34, pp. 118-122. 2002.
- [9] Wu, Hongxia, Konstantinos S. Tsakalis, and Gerald Thomas Heydt. "Evaluation of time delay effects to wide-area power system stabilizer design." *Power Systems, IEEE Transactions on* 19, no. 4 (2004): 1935-1941.
- [10] Li, Y., C. Rehtanz, D. Yang, S. Rüberg, and U. Häger. "Robust high-voltage direct current stabilising control using wide-area measurement and taking transmission time delay into consideration." *IET generation, transmission & distribution* 5, no. 3 (2011): 289-297.
- [11] Wang, Shaobu, Xiangyu Meng, and Tongwen Chen. "Wide-area control of power systems through delayed network communication." *Control Systems Technology, IEEE Transactions on* 20, no. 2 (2012): 495-503.
- [12] Mokhtari, Maghsoud, and Farrokh Aminifar. "Toward Wide-Area Oscillation Control Through Doubly-Fed Induction Generator Wind Farms." *IEEE Trans. on Power Sys.*, to be published.
- [13] Mokhtari, Maghsoud, Farrokh Aminifar, Daryoosh Nazarpour, and Sajjad Golshannavaz. "Wide-area power oscillation damping with a fuzzy controller compensating the continuous communication delays." *Power Systems, IEEE Transactions on* 28, no. 2 (2013): 1997-2005.
- [14] Stahlhut, Jonathan W., Timothy J. Browne, Gerald Thomas Heydt, and Vijay Vittal. "Latency viewed as a stochastic process and its impact on wide area power system control signals." *Power Systems, IEEE Transactions on* 23, no. 1 (2008): 84-91.
- [15] Dotta, Daniel, and I. C. Decker. "Wide-area measurements-based two-level control design considering signal transmission delay." *Power Systems, IEEE Transactions on* 24, no. 1 (2009): 208-216.
- [16] Tomovic, Kevin, David E. Bakken, Vaithianathan Venkatasubramanian, and Anjan Bose. "Designing the next generation of real-time control, communication, and computations for large power systems." *Proceedings of the IEEE* 93, no. 5 (2005): 965-979.
- [17] Zhou, Kemin, and John Comstock Doyle. *Essentials of robust control*. Vol. 104. Upper Saddle River, NJ: Prentice hall, 1998.
- [18] Heniche, Anissa, and Innocent Kamwa. "Assessment of two methods to select wide-area signals for power system damping control." *Power Systems, IEEE Transactions on* 23, no. 2 (2008): 572-581.
- [19] Safonov, M. G., and R. Y. Chiang. "A Schur method for balanced-truncation model reduction." *Automatic Control, IEEE Transactions on* 34, no. 7 (1989): 729-733.
- [20] Chaudhuri, Balarko, and Bikash C. Pal. "Robust damping of multiple swing modes employing global stabilizing signals with a TCSC." *Power Systems, IEEE Transactions on* 19, no. 1 (2004): 499-506.
- [21] Rogers, Graham. *Power system oscillations*. Boston: Kluwer Academic, 2000.
- [22] Taylor, Carson W., Mani V. Venkatasubramanian, and Yonghong Chen. "Wide-area stability and voltage control." In *Proc. VII Symp. Specialties Electr. Oper. Expansion Planning*, pp. 21-26. 2000.

# DEVELOPING A WEB-BASED ONTOLOGY FOR EXPERT DECISION SUPPORT SYSTEMS

Saud Ben Mahmoud Mandourah\*

\*Business Administration dept.  
Faculty of Economics and administration,  
King Abed-Alziz University, Jeddah, KSA  
[dr\\_saud2002@yahoo.com](mailto:dr_saud2002@yahoo.com)

Abdel-Badeeh M. Salem\*\*

\*\*Faculty of Computer & Information Sciences, Ain  
Shams University, Abbassia, Cairo, Egypt  
[absalem@cis.asu.edu.eg](mailto:absalem@cis.asu.edu.eg)

**Abstract**—Expert systems were originally developed to solve ill-defined problems and well-defined problems that are not efficiently solved with algorithmic approaches. This technology provides an innovative and robust techniques to capture and package knowledge. Its strength lies in its ability to be put to practical use when an expert is not available. This technology has proven to be especially effective when the task is in a rapidly changing environment. On the other side, ontology is the foundation of describing a domain of interest and it consists in a collection of terms organized in a hierarchical structure that shape the reality. The main objective of using ontologies is to share knowledge between computers or computers and human. Most of the usages of ontologies in the field of artificial intelligence are related to knowledge based systems and intelligent systems. These types of ontologies include a small number of concepts and their main objective is to facilitate reasoning tasks. This paper presents the developing of web-based ontology for expert systems technology. The developed ontology was encoded in OWL-DL format using the Protégé-OWL editing environment.

**Keywords**-Ontological Engineering, Expert Systems, Artificial Intelligence, Knowledge Engineering, Web Technology

## I. INTRODUCTION

Ontologies provide a common vocabulary of an area and define, with different levels of formality, the meaning of the terms and the relationships between them. During the last decade, increasing attention has been focused on ontologies [11, 22]. The main benefits of using ontological engineering approach are: (a) to share common understanding of the structure of information among people or software agents (b) to enable reuse of domain knowledge (c) to make domain assumptions explicit (d) to separate domain knowledge from operational knowledge (e) to analyze domain knowledge. At present, there are applications of ontologies with commercial, industrial, biology, medical, education and research focuses [9,13,14].

On the other side, artificial intelligence (AI) technology includes the following sub-technologies (or applications): general problem-solving, expert systems, natural language processing, computer vision, robotics, education and games. The recent fruits of artificial life include the following AI technologies: (a) Software Agents: may be able to act autonomously learning how to solve problems, (b) Software Code: automatically evolves using genetic algorithms, (c)

Complex Computer Simulation: predict environmental, social and biological trends and (d) Biological Robots: programmed to mimic the reasoning of insects may learn to find their way. In this paper we focused our discussion on the expert systems technology.

In this paper our goal is exploiting the ontological engineering approach to develop a web-based ontology for expert systems technology. Section 2 presents an overview of the general features of the expert systems. Section 3 discusses the ontological engineering approach from the perspective of computer science. Section 4 introduces of the research issues for building ontologies. Section 5 presents the developed expert systems ontology. Finally section 6 concludes the work.

## II. EXPERT SYSTEMS

An expert system (ES) is an intelligent system incorporating a knowledge base and inference engine [20]. It is a highly specialized piece of software that attempts to duplicate the function of an expert in some field of expertise. The ES acts as an intelligent consultant or advisor in the domain of interest, capturing the heuristic knowledge of one or more experts. Non-experts can then tap the ES to answer questions, solve problems, and make decisions in that domain [3].

Expert systems will make knowledge more widely available and will help overcome the age-old problem of translating knowledge into practical, useful results. And perhaps best of all, it is a new more way that technology is helping us get a handle on the information glut. All AI software is knowledge based as it contains useful facts, data, and relationships that are applied to a problem. From the AI point of view, expert systems include the following topics; (a) Knowledge-representation techniques, (b) knowledge engineering tools and shells, (c) intelligent programming languages, (d) inference techniques, (e) reasoning methodologies, (f) machine learning and (g) user interface technologies.

### A. Knowledge representation for expert systems

The first step in constructing AI software is to build a knowledge base. In order to act intelligently, a computer must have knowledge about the domain of interest. The knowledge of the domain must be collected and codified. It



must be organized, outlined, or otherwise arranged in a systematic order. This process of collecting and organizing the knowledge is called knowledge engineering. It is the most difficult and time-consuming stage of any AI software development process. Although a variety of knowledge representation schemes have been developed over the years, these representation schemes share two common characteristics. First, they can be programmed with computer languages and stored in memory. Second, they are designed so that the facts and other knowledge contained within them can be manipulated by an inference system, the other major part of an AI program. The inference system uses search and pattern matching techniques on the knowledge base to answer questions, draw conclusions, or otherwise perform an intelligent function. A brief overview of each of these schemes is presented in the following subsections.

#### - **Logic**

Logic is the oldest form of knowledge representation technique. For a computer to perform reasoning using logic, some method must be used to convert the deductive or inductive reasoning process into a form suitable for manipulation by a computer. The result is what is known as symbolic logic or mathematical logic. It is a system of rules and procedures that permit the drawing of inferences from various premises using a variety of logical techniques. These methods are generally known as computational logic. There are two basic forms of computational logic, propositional logic and predicate logic. Since propositional logic deals primarily with complete statements and whether they are true or false, its ability to represent real world knowledge is limited. Consequently, intelligent tutoring training technology uses predicate logic instead. Predicate logic gives added ability to represent knowledge in finer detail.

#### - **Lists and Trees**

Lists and trees are simple structures used for representing hierarchical knowledge. A list is a series of related items. Objects are divided into groups or classes of similar items. Their relationships are shown by linking them together. The simplest form is one list, but a hierarchy is created when two or more related lists are combined. On the other hand, a tree is a graphical structure of hierarchy, it is simply a way of illustrating lists and other hierarchical knowledge.

#### - **Semantic Networks**

Semantic networks are basically graphical depictions of knowledge that show hierarchical relationships between objects. A semantic network is made up of a number of nodes, which represent objects and descriptive information about those objects. Objects can be any physical items such as a book, car, desk, or even a person. Nodes can also be concepts, events, or actions. The nodes in a semantic network are also interconnected by link or arcs. The arcs show the relationships between the various objects and descriptive factors. Some of the most common arcs are of the is-a or has-a type.

#### - **Frames**

A frame is a relatively large block or chunk of knowledge about a particular object, event, location, situation, or other element. The frame describes that object in great detail. The detail is given in the form of slots which

describe the various attributes and characteristics of the object or situation. Frames are normally used to represent stereotyped or knowledge based on well-known characteristics and experiences. With frames, it is easy to make inferences about new objects, events, or situations because they provide a base of knowledge drawn from previous experience. For example, the items of the components of any automobile or the animal kingdom can be represented in frames format.

#### - **Scripts**

A script is a knowledge representation scheme similar to a frame, but instead of describing an object, the script describes a sequence of events. Like the frame, the script portrays a stereotyped situation. Unlike the frame, it is usually presented in a particular context. To describe a sequence of events, the script uses a series of slots containing information about the people, objects, and actions that are involved in the events. Some of the elements of a typical script include entry conditions, props, roles, tracks and senes.

### *B. Expert systems types based on the reasoning methodology*

The field of reasoning is very important for the development of expert systems and all intelligent systems. The research area in this field covers a variety of topics, e.g.; automated reasoning, case-based reasoning, commonsense reasoning, fuzzy reasoning, geometric reasoning, non-monotonic reasoning, model-based reasoning, probabilistic reasoning, causal reasoning, qualitative reasoning, spatial reasoning and temporal reasoning. This subsection is dealing with rule-based and case-based reasoning systems.

In rule-based expert system, the inference engine contains a set of formal logic relationships which may or may not resemble the way that real human expert reach conclusions. The knowledge base is structured in a if-then organization. The rules have to be defined in a limited number of formal ways. Typically they may be a set of some hundreds of if-then (or if A and B but not C then D) types of relationships that describe all the domain specific knowledge used by the human expert. The most difficult and time consuming part of the developing a rule-based expert system is the extraction of knowledge form the head of an acknowledged expert (or a group of experts) and then transforming it into a form acceptable to the expert system knowledge based structure.

Case-based expert system (CES) uses the case-based reasoning (CBR) methodology as an efficient method for inference instead of the production rules in the traditional rule-based systems . CBR is an analogical reasoning method provides both a methodology for problem solving and a cognitive model of people. It is consistent with much that psychologist have observed in the natural problem solving that people do. Decision makers tend to be comfortable using CBR methodology in dynamically changing situations and other situations were much is unknown and when solutions are not clear.

The methodology of CES can be summarized in the following two main processes:



- **Case-search process:** In this process the system will search its Case-Memory for an existing case that matches the input problem specification. If we are lucky (our luck increases as we add new cases to the system), we will find a case that exactly matches the input problem and goes directly to a solution. If we are not lucky, we will retrieve a case that is similar to our input situation but not entirely appropriate to provide a complete solution.
- **Case-adaptation process:** In this process the system must find and modify small portions of the retrieved case that do not meet the input specification. The result of case adaptation process is (a) completed solution, and (b) generates a new case that can be automatically added to the system's case-memory for future use.

The technology of CBR directly addresses the following problems found in rule-based technology.

- **Knowledge acquisition:** The unit of knowledge is the case, not the rule. It is easier to articulate, examine, and evaluate cases than rules.
- **Performance:** A CBR system can remember its own performance, and can modify its behavior to avoid repeating prior mistakes.
- **Adaptive Solutions:** By reasoning from analogy with past cases, a CBR system should be able to construct solutions to novel problems.
- **Maintaining:** Maintaining CBR system is easier than rule-based system since adding new knowledge can be as simple as adding a new case.

### III. ONTOLOGICAL ENGINEERING FROM THE COMPUTER SCIENCE PERSPECTIVE

According to Sowa [17], the components of ontology are: (a) concepts, terms; (b) relations between concepts, terms; (c) properties, attributes of the concepts; and (d) rules, axioms, predicates, constraints. The main objective of using ontologies is to share knowledge between computers or computers and human. Computers are capable to transmit and present the information stored in files with different formats, but they are not yet compatible to interpret them. To facilitate communication and intelligent processing of information, it is necessary that all actors of the digital space (computers and humans) have the same vocabulary.

Most of the usages of ontologies in the field of artificial intelligence are related to knowledge based systems and intelligent systems. These types of ontologies include a small number of concepts and their main objective is to facilitate reasoning. For example, in a multi-agent systems, the knowledge representation is accomplished through a basic ontology, private ontologies and a knowledge base. Private ontologies of the agents are derived from the basic ontology. The names of the concepts used in private ontologies of the agents are unknown, but their definitions use terms from the basic ontology.

### IV. ONTOLOGICAL ENGINEERING RESEARCH ISSUES FOR BUILDING ONTOLOGIES

Ontological engineering refers to the set of activities that concern the ontology development process, the ontology life cycle, the methods and methodologies for building ontologies, and the tool suites and languages that support them[6,16].

#### A. Methodologies

Ontological engineering is still relatively immature discipline; each research group employs its own methodology. Ontology methodologies differ according to the strategy of identifying concepts. The well known three possible strategies for identifying concepts are: (a) bottom-up from the most concrete to the most abstract; (b) top-down from the most abstract to the most concrete; and (c) middle-out from the most relevant to the most abstract and most concrete. The last one is the most common strategy.

#### B. Ontological Languages and Tools

A great range of languages have been used for implementing ontologies during the last decade; e.g. LOOM, OCML, FLogic, CARIN, OKBC, Telos, Cycl [1,2,10]. These languages are in a stable phase of development, and their syntax consists of plain text where ontologies are specified (many of them have a Lisp-like syntax). Recently, Web-based ontology specification languages have been developed in the context of the World Wide Web. These languages have had great impact in the development of the Semantic Web (e.g. XOL, OML, OIL, and OWL. [1]). The syntax of these languages is based on XML, which has been widely adopted as a standard language for exchanging information on the web, except for SHOE, whose syntax is based on HTML.

On the other side, ontological tools have emerged for creating, editing and managing ontologies written in the various languages. These tools usually provide a graphical user interface for building ontologies, which allows the ontologist to create ontologies without using directly a specific ontology specification language (e.g. WebODE, Ontolingua, Ontosaurus, and OntoEdit, OilEd).

#### C. Ontology Interoperability

The domain of ontologies is extremely vast. A lot of ontologies were developed, even different ontologies for the same domain. In order to assure the interoperability between software applications, it is necessary to guarantee the interoperability between their ontologies. In the literature, there are different technologies related to the ontologies' interoperability, namely; ontology alignment, ontology mapping matching, ontology translation, ontology integration, ontology refinement and ontology unification [18,19].

#### D. Ontology Validation

Validation is the process to determine whether a work product satisfies its requirements. Validation can be performed after the ontology has been developed, but it is

usually better to validate while the ontology is being built. There are several techniques that can be used to validate ontology: (a) Verify the fulfillment of the purpose, (b) Check that all usage examples are expressible, (c) Create examples that are consistent with the ontology, and determine whether they are meaningful, and (d) Check that the ontology is formally consistent. Quality of the ontology is validated based on the following criteria: (a) consistency; (b) completeness; (c) conciseness; (d) clarity; (e) generality; and (f) robustness.

#### E. Ontology Evaluation

From the ontological engineering perspective, ontology evaluation is based on the following criteria: (a) completeness; (b) correctness; (c) decidability; (d) maintainability; (e) minimal redundancy; (f) rich axiomatization; and (g) efficiency. A more formal ontology evaluation method, proposed by Obrst et al.[12], includes: (a) development of an ontology and ontology tool competition; (b) principled certification of ontologies by a reviewing organization or community; and (c) the development of an ontology maturity model.

### V. DEVELOPING A WEB-BASED “EXPERT SYSTEMS” ONTOLOGY

Our methodology for developing the web-based “Expert Systems” ontology are summarized in the following steps:

- Step 1 Organizing and scoping: Determining the objectives and defining the boundaries of the ontology.
- Step 2 Data collection: the raw data needed for ontology development is acquired. In our study data was collected from AI and expert books [ 5, 8, 15].
- Step 3 Data analysis: the ontology is extracted from the results of data collection. This step includes the following tasks: (a) define the classes and class hierarchy, (b) define the properties of classes (slots) , (c) define the facets of the slots (e.g. domain and range of a slot, cardinality ,slot-value type ), and (d) create individual instances of classes.
- Step 4 Development of initial ontology: a preliminary ontology is developed (i.e. classes, relations and properties). This process was done by using OWL-DL language using the Protégé-OWL editing environment.
- Step 5 Ontology refinement: the initial development is iteratively refined.
- Step 6 Ontology validation.

Figure 1 shows the semantic net of expert systems (identifications of main object of interest and relationships between objects). Figure 2 shows the developed expert systems ontology encoded in OWL-DL format using Protégé OWL editing environment. In this ontology four main super-classes namely (a) expert system tools; (b) knowledge base; (c) inference mechanism; and (d) user interface. Expert system tools have four subclasses: (a) programming languages (b) knowledge-engineering language; (c) system-building aids; and (d) support-environment tools.

### Characteristics and Features of Expert Systems

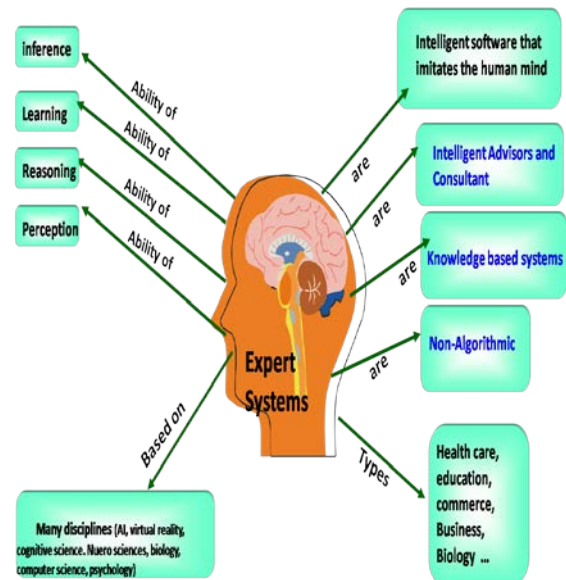


Figure 1.a the General Characteristics of Expert Systems

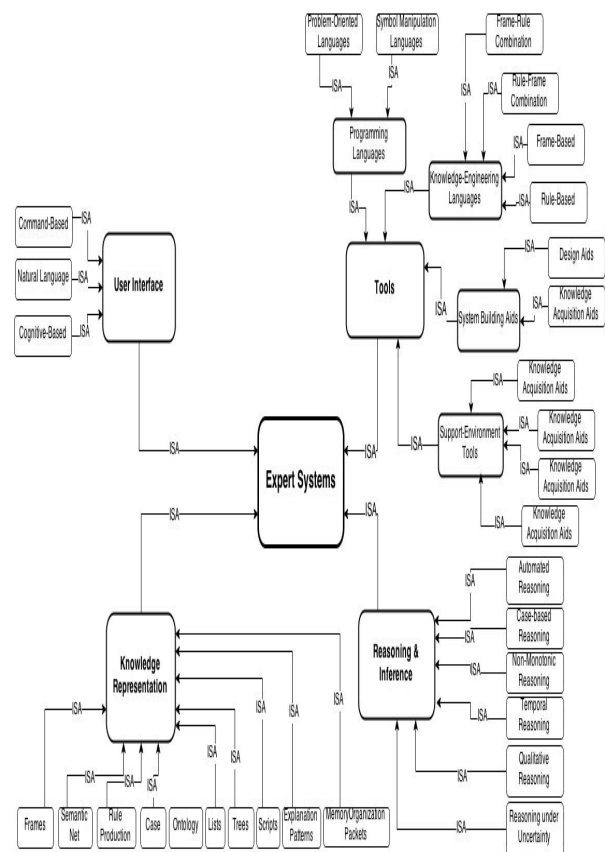


Figure 1.b the developed Semantic Net of Expert Systems



Figure 2. a Expert systems applications ontology

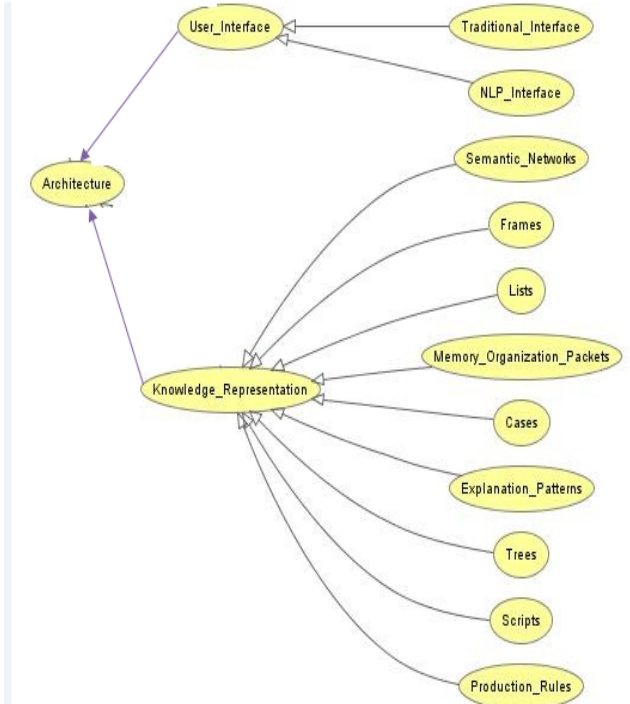


Figure 2.c Expert System architecture ontology



Fig 2 ( b) Expert systems tools ontology

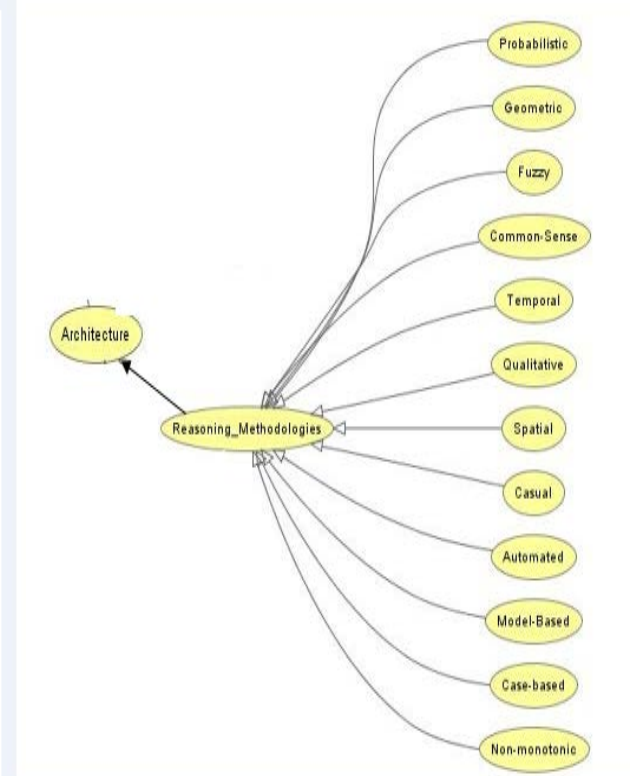


Figure 2.d Expert System architecture ontology  
Figure 2 Expert Systems Ontology: encoded in OWL-DL format using Protégé OWL Editing Environment

## VI. DISCUSSION

Ontologies are now ubiquitous in many information-systems enterprises. They constitute the backbone for the Semantic Web as well as they are used in all applications of e-activities technologies (e.g. e-Government, e-Learning, e-Health, e-Business). Ontological engineering approach is an effective methodology to manage and represent knowledge. Ontologies were developed in intelligent learning systems to facilitate knowledge sharing, refine, search and reuse. These ontologies may be used as an assessment procedure. For example and from the educational point of view, candidates show their knowledge and understanding while creating ontologies. Knowledge entities that represent static knowledge of the domain are stored in the hierarchical order in the knowledge repository and can be reused by other candidates. At the same time those knowledge entities can be also reused in description of the properties or arguments of methods of another knowledge entity. Moreover, ontology approach enables to solve the complexity and the incertitude of the instructional systems. An intelligent learning system based on a multi-agent approach consists in a set of intelligent agents, which have to communicate. They collaborate through messages. Software agents can understand and interpret the messages due to a common ontology or the interoperability of the private ontologies.

## VII. CONCLUSION

Our contribution in this paper are summarized as follows:

1. Building Semantic Net of the Expert Systems Technology: we built the semantic net of expert systems technology which defines the main objects of interest and relationships between objects.
2. Determination the Main Components of the Semantic Net: we determine the classes and its hierarchy, the properties of classes (slots), the slot's domain and range, and the individual instances of classes.
3. Developing the Web-based Ontology for Expert Systems Technology: To achieve this aim, we proposed a new methodology for building this ontology using the ontology web-based language (OWL) and the Protégé-OWL editing environment.

## REFERENCES

- [1] A. Duineveld, R. Studer, M. Weiden, B. Kenepa, R. Benjamins, WonderTools? A comparative study of ontological engineering tools, in: Proc. 12th Knowledge Acquisition Workshop (KAW99), Banff, 1999.
- [2] A Salem, Marco Alphonse, Medical Informatics Workshop, proceedings of 3rd international conference on Intelligent Computing and Information Systems, Cairo, Egypt, pp 59- 74, 2007
- [3] Cornelius T. Leondes , Fuzzy Logic and Expert Systems Applications (Neural Network Systems Techniques and Applications),1998.
- [4] Cyc ontology, www.cyc.com, accessed on August 12, 2009
- [5] George F Luger, Artificial Intelligence Structure and Strategies for Complex Problem Saving, Addison Wesley, 2005
- [6] Gurber T.- Ontology, Encyclopedia of Database Systems, Ling Liu and M. Tamer Özsu (Eds.), Springer-Verlag, in press, 2009
- [7] JOSEPH C. GIARRATANO, GARY D. RILEY, EXPERT SYSTEMS: PRINCIPLES AND PROGRAMMING, THIRD EDITION: PRINCIPLES AND PROGRAMMING , 1998.
- [8] Louis E. Frenzel, Jr. , Crash Course in Artificial Intelligence and Expert Systems, Howard W. Sams & Co. A Division of Macmillan, Inc. First Edition , 1987.
- [9] L. Tankeleviciene, R. Damasevicius, Characteristics for Domain Ontologies for Web Based Learning and their Application for Quality Evaluation, Informatics in Education, vol. 8 No 1, pp. 131-152,2009.
- [10] M. Dean, D. Connolly, F. van Harmelen, J. Hendler, I. Horrocks, D.L. McGuinness, P.F. Patel-Schneider, L.A. Stein, OWL Web Ontology Language, W3C Recommendation, 2004
- [11] M. Fernández-López, Asunción Gómez-Pérez. "Deliverable 1.4: A survey on methodologies for developing, maintaining, evaluating and reengineering ontologies". Part of a research project funded by the IST Programme of the Commission of the European Communities as project number IST-2000-29243, 2002.
- [12] Obrst, L., Hughes, T. and Ray, S. (2006). Prospects and possibilities for ontology evaluation: The view from NCOR. In Proc. of 15th Int. World Wide Web Conf., WWW 2006, Edinburgh, UK.
- [13] Olivier Bodenreider1, Anita Burgun, "Biomedical Ontologies", Medical informatics: Advances in knowledge management and data mining in biomedicine, Springer-Verlag; 2005. p. 211-236.
- [14] Ontology for Education, <http://o4e.iiscs.wssu.edu/xwiki/bin/view/Blog/About>
- [15] PETER JACKSON , INTRODUCTION TO EXPERT SYSTEMS ,1998.
- [16] Rajapbothr hack and Mooter F. A Generic Task Ontology for Scheduling
- [17] Applications, <http://projects.kmi.open.ac.uk/akt/publication-pdf/generictaskontology.pdf> accessed on March 22, 2009
- [18] Sowa web site, <http://www.jfsowa.com/>, accessed on May 10, 2010
- [19] Sowa J . F . - Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks Cole Publishing Co., Pacific Grove, CA., 1999
- [20] State of the Art Report 3. Ontology Interoperability, Draft version 0.3.2, [http://twiki.di.uniroma1.it/pub/Estrinfo/Materiale/3\\_Ontology\\_merging.pdf](http://twiki.di.uniroma1.it/pub/Estrinfo/Materiale/3_Ontology_merging.pdf), accessed on July 18, 2009
- [21] STUART J. RUSSELL AND PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH , PRENTICE HALL 2ED EDITION, 2002.
- [22] T r a u s a n - M a t u , S . - Programming in Lisp, Artificial Intelligence and Semantic Web, Polirom, Iasi, 2004
- [23] X. Su and L. Ilebrikke A Comparative Study of Ontology Languages and Tools, In Proceeding of the 14th Conference on Advanced Information Systems Engineering (CAiSE'02), Toronto, Canada, May 2002.

# Design of Smart Bridge based on WSN for efficient measuring of Temperature, Strain and Humidity

Mohamed. Bouyahi, Houria. Rezig, and Tahar. Ezzedine

**Abstract**—This paper presents the development of an automatic wireless sensor monitoring system for civil engineering structures. The objective is to provide a solution to measure temperature, humidity and strain inside a concrete structure. The research has been focused in the early age and curing phase period. Two sensors have been used. The first sensor fiber optic Brillouin Strain and temperature, the second is the sensor fiber optic Bragg of Temperature and Humidity. The two sensors used with Zolertia mote MSP430F2617 microcontroller allowing for the creation of an IEEE 802.15.4 network.

**Keywords**— Structural Health Monitoring (SHM); Optic Fiber Sensor (OFS); Wireless Sensor Network (WSN).

## I. INTRODUCTION

SHM is a new concept of civil engineering, it is able to monitor the structure on real time and evaluate the performance under many loads and determine the damage of structure. The goal of the monitoring is understand the structure and prevent the damage and prepares to repair it before the catastrophe failure. Many systems based maintenance are implemented in many structure and many sensors used. [2] Explain the SHM for Yongjong Grand Bridge (a long span Bridge such as Yongjong Grand Bridge). It contains a technology called MBM (Monitoring Based Maintenance), the sensors installed on cables vibration of stiffening stress, inclination and vibration of pylon, seismic acceleration, main cable temperature and wind direction/velocity. [9] implements 800 sensors in the super-tall, the 400 sensors vibrating strain gauges installed in different heights, the displacements and tilts sensors installed on the structure top under normal and typhoon conditions. [3] Use in the railway structure these sensors; the strain gages sensor, a distributed fiber optic strain sensor based on Rayleigh backscatter and embedded the long gage sensor in the concrete slabs, these sensors are used to measure the strain of structure.

This work was supported in part by the SysCom Laboratory. Communication System Laboratory Sys'Com, National Engineering School of Tunis, University Tunis El Manar, BP 37, belvedere 1002 Tunis, Tunisia

M. B. Author is with the National Engineering School of Tunis, University Tunis El Manar, BP 37, belvedere 1002 Tunis, Tunisia; (e-mail: mohamed.bouyahi@enit.rnu.tn).

H. R. Author is with the National Engineering School of Tunis, University Tunis El Manar, BP 37, belvedere 1002 Tunis, Tunisia; (e-mail: houria.rezig@enit.rnu.tn).

T. E. Author is with the National Engineering School of Tunis, University Tunis El Manar, BP 37, belvedere 1002 Tunis, Tunisia; (e-mail: tahar.ezzedine@enit.rnu.tn).

SHM enables the measurement of structural safety and integrity and ensure integration with bridge maintenance and management to contribute to bridge management plan. The long term loads will suffer damage, the loads result the variation of structure. Therefore, regularly testing the bridge during construction and operation to ensure its safety.

To evaluate the health of bridge structure and transmit the data with wired solution or optical cable, this method will costs lots of manpower and material resources, and also cause measurement inefficiency, maintenance difficulty, and even causes unreliability of data transmission. The advancement in wireless technology has provided motives to the authors to develop the wireless network-based bridge health monitoring system.

Wireless Sensor Networks (WSNs) integrated with sensor technology, embedded computing technology, modern networking and wireless communications technologies, and distributed information processing technology can monitor, perceive and acquire various environmental or object information [7].

This paper presents a prototype of wireless sensor network system for a structure health monitoring. Sensor devices such as fiber optic Bragg of Humidity and Temperature and fiber optic Brillouin of Temperature and Strain and ZigBee modules are combined to implement the System Hardware. The Collect Tree Protocol (CTP) is tested to send the data to gateway. The collected data sent to database for SHM and the application software view the result [2].

This article opens with motivation and objectives carried out on objective to used the Wireless Sensor Network which shall be introduced in Section 2. Next section will present the system architecture. Then, in section 4 a sensor selection along with our installation. While Section 5 will include system hardware, Section 6 presents the system software. Last section 7 concludes the paper.

## II. MOTIVATION AND OBJECTIVES

The objective of this researcher is to develop a prototype of Wireless Sensor Networks for Remotely Monitoring Bridge. WSN is formed by tiny devices; the tiny device is the mote as composed by microcontroller, sensors, memory, power unit and a communication module. They are able to sense the value of sensors and communicate with the base station as a wireless links. These sensors are reading the physical parameters of the structure at a given location, detecting of events and the signal as processing from the structure to the application. The WSNs are useful in situation where the structure is required. The research work was focused on detecting the structure parameter based on temperature, strain and humidity of the Bridge. This is accomplished by using the 802.15.4



communication enabling a significant reduction of the installation time and costs.

The WSN using by the civil engineering is able to accommodate the large number of sensor, depending of needs, the small cost, sharing the information using the network. For the context the remote agents can collect the information and storage in the database through base station Gateway. The gateway has an interface of ZigBee module and another interface module 3G that is connected to a PC. The Fig. 1. Resume the system architecture.

Temperature is an important parameter, if the temperature decreases the hydration reaction slow down, if the temperature of concrete increase the hydration reaction accelerates. Result the differential of temperature within the concrete causing the lead to the fissuring.

The life cycle of the concrete bridge and her development linked with rate of hydration. It is important to study the impact of the temperature increase caused by the occurrence of the hydration reaction. A structural health monitoring is an important field of application for WSNs. Since with traditional wired-based solutions present the many problems like installation and maintenance cost. A WSN was created to collect the data from a sensor which facilitates the monitoring.

The properties of civil structure involve an important quantity of uncertainties in several parameters, this parameter like temperature, strain and humidity. In the ages of structure the temperature, strain and humidity can caused the most factors of deterioration of structure and can plays an important role to during of the concrete, and have a long-term consequence [1].

### III. SYSTEM ARCHITECTURE

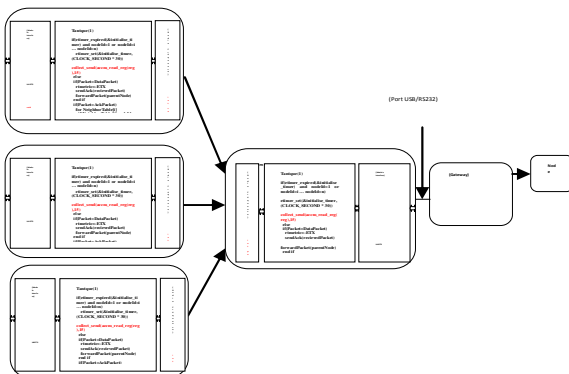


Fig. 1 System Design

The system is designed based on the wireless sensor network principles and we adopt the ZigBee communication inter-nodes. We have the central station like base station and many sensors in the Structure. The node data is transmitted in the 250 kbps radio channel. The Fig. 1. is shown the system design. The frequency to send packet from node to another is deferent when the frequency of the sensor fiber optic is deferent the frequency to send packet is deferent.

A wireless communication between the mote and the Gateway; the data request is transmitted to the Gateway after the sensor is detected the measurement to the Bridge, the

Gateway is connected to the server when to store the data in the database server via the module 3G. The Gateway transmits the data to the server. The data is sent after the moment of temp, this temp depends on the sensors and the variation of these parameters [7].

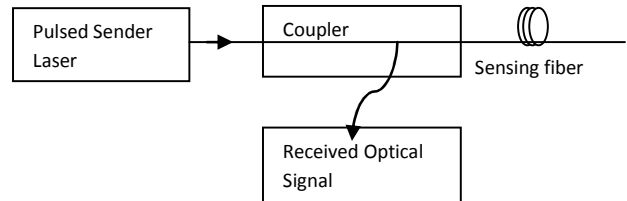


Fig. 2 Optical Time Domain Reflectometer functional schematic.

### IV. SENSOR SELECTION

In this section, we present the sensors are used in the structure. The Fiber Optic Sensor based on Stimulated Brillouin Scattering (SBS) for measuring the Temperature and Strain and the Fiber Bragg (FBG) for measuring the Temperature and Humidity. The installation model of sensor when the sensor is used for backscattering optic signal present on Fig. 2.

#### A. Sensor Fiber Brillouin

The physical parameters of optical fiber can be affected by temperature and strain. So the fiber sensors are able to detect the variation of temperature and strain over long distances.

These parameters have becoming the essence of distributed fiber optic sensing. The Rayleigh, Raman, and Brillouin scattering represent the basic scattering mechanism of the distributed sensing techniques which commonly occurred inside the fiber. The distributed fiber sensing is a really attractive technique for structural health monitoring (SHM). While it provide information of strain and temperature about a section or the complete structure with durability, robustness and measurement reliability. The distributed optical fiber sensing systems give us the opportunity to determine physical parameters, when large structures are to be monitored, such as bridges, dams and tunnels so this mechanism is suitable.

The Brillouin frequency shift has good linear relationship to the strain and temperature of the optical fiber, which can realize simultaneous temperature and strain measurement by using one optical fiber, with a lower accuracy [2] [5].

A pulse of laser is launched into the fiber, we amplify the input signal pulse in an optical fiber in a similar way as it undergoes in Erbium-Doped Fiber Amplifiers (EDFA). The directional coupler is backscattering the light to the same fiber to measure. The pulse propagates along the fiber to the receiver.

The backscattering of light is resumed of the scattering of incident photon in the fiber by the acoustic photon of the medium, this mechanism generate the frequency shift when we are measured the temperature and strain.

The backscattering signal refers to the Stokes frequency of the Stimulated Brillouin Scattering (SBS) and the occurrence of

the bathochromic shift (the Stokes component)  $w_s = w_p - w_{fa}$  with respect to the pumping beam, where  $w_{fa}$  denotes acoustic photon of the fiber. The fundamental difference from Brillouin and Raman scattering is the interaction of light photon with acoustic phonons in contrast to the Raman scattering where the interaction of light photons and molecular vibrations. The frequency of acoustic photon is lower than optical photon. The acoustic photon is determined by the following formula.

$$w_{fa} = \frac{4\pi n v_s}{\lambda} \quad (1)$$

$v_s$  indicate the speed of sound in optical waveguide

$n$  refraction index.

As a result of interaction between the matter and light photons the energy exchange occurs via acoustic photons leading to the third-order polarization of in medium.

$$p^{(3)} = \chi_{ijkl}^{(3)} E_j E_k E_l \quad (2)$$

$$\epsilon_s = \frac{F}{E * A} (\cos^2 \theta - \sin^2 \theta) \quad (3)$$

The variation is detected by sensor Fiber Brillouin is  $\epsilon_s$  focused by  $F$  is the force applied by concrete,  $A$  is the surface of optical fiber,  $E$  Young Module and  $\theta$  the direction of light in the fiber [7].

The Brillouin Frequency Shift (BFS) of back scattering is linearly sensitive to strain and temperature.

It's expressed by:

$$\Delta V_B(T, \epsilon) = C_T \Delta T + C_\epsilon \epsilon \quad (4)$$

#### B. Sensor Fiber Bragg

The sensor fiber optic is built to change their index of refraction periodically. The fiber Bragg grating is a linear superposition of temperature (T) and Humidity (RH) effects. In the presence of variations in relative temperature,  $\Delta T$ , and humidity,  $\Delta RH$ , the relative Bragg wavelength shift  $\Delta\lambda / \lambda$  is therefore given by

$$\frac{\Delta\lambda}{\lambda} = S_{RH} \Delta RH + S_T \Delta T \quad (5)$$

When the  $S_T$  and  $S_{RH}$  are the sensor sensitivities to relative temperature and humidity, respectively. To relate the sensitivities to material properties, one many express  $S_{RH}$  as the sum of a mechanical, a strain-optic, and, for  $S_T$  only, a thermo-optic contribution [6].

$$S_{RH} = \beta_{cf} - P_e (\beta_{cf} - \beta_f) [\%RH^{-1}] \quad (6)$$

$$S_T = \alpha_{cf} - P_e (\alpha_{cf} - \alpha_f) + \xi [K^{-1}] \quad (7)$$

Where  $\beta_i$  is the hygroscopic longitudinal expansion

coefficient, and  $\alpha_i$  is the thermal longitudinal expansion coefficient The subscript stands for bare fiber (i=f), coating (i=c), and coated fiber (i=cf).  $\hat{P}_e = n_{eff}^2 [P_{12} + \epsilon_{f,r} / \epsilon_{f,z} (P_{11} + P_{12})] / 2$  is the effective photoelastic coefficient of the coated fiber, where  $n_{eff}$  is the effective refractive index of the mode,  $P_{ij}$  are the coefficients of the strain-optic tensor, and  $\epsilon_{f,r}$  and  $\epsilon_{f,z}$  are the radial and axial elastic fiber strains, respectively.  $\xi$  is the thermo-optic coefficient of the fiber core [4] [8].

#### V. SYSTEM HARDWARE

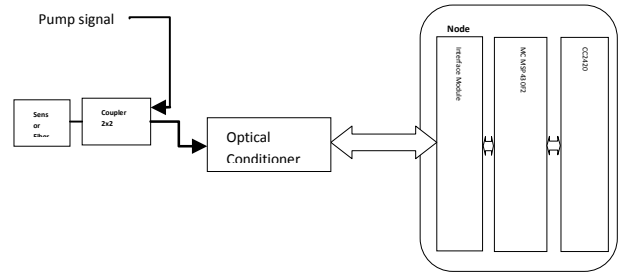


Fig. 3 Mote and Sensors.

The node defined in Fig. 3. Correspond to a Pump laser, Sensor Fiber Optic, Coupler 2x2; Optical Conditioner, Interface Module and the wireless communication module are integrated in single unit. This unit transmits the signals with the digital format by wireless module and the development of unit combine the ADC coupling with ZigBee module.

The output from the optical sensor in a form of the impulsion of optical signal. The sensory output need to be converted to the electric by Optical Conditioner. The electric signal converts to the digital signal by interface module. After reading the data from the sensor and deliver to the ZigBee module. This sends the data to other ZigBee module wirelessly. The unit is powered by the battery [2].

#### A. Zolertia Mote microcontroller

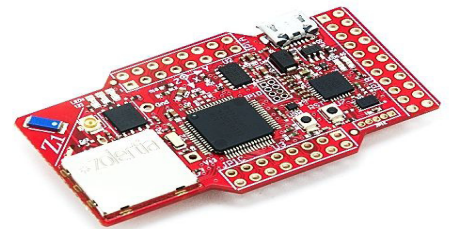


Fig. 4 Mote Zolertia.

The Z1 module Fig. 4. Developed by Zolertia, Z1 is a low power wireless module compliant with IEEE 802.15.4 and ZigBee protocols intended to help WSN developers to test and deploy their own applications and prototypes with the best tradeoff between time of development and hardware flexibility. Its core architecture is based upon the MSP430 and CC2420 family of microcontrollers and radio transceivers by Texas Instruments, which makes it compatible with motes based on this same architecture.

It has many port like UART, SPI, I2C, Phidgets, ADCs and DACs which facilitate the communication with external sensors, It is used in ultralow power consumption mode consist of several devices featuring different sets of peripherals targeted for various applications and allows its use for the high-frequency sampling required for dynamic structural monitoring. The architecture, combined with five low power modes is optimized to achieve extended battery life in portable measurement applications. It has the 16 MB Flash Memory, 8KB RAM [5] [7] [10].

#### B. Transceivers CC2430

The CC2420 is a true-chip 2.4 GHz IEEE 802.15.4 compliant RF transceiver designed for low-power and low-voltage wireless applications. CC2420 includes a digital direct sequence spread spectrum baseband modem providing a spreading gain of 9 dB and an effective data rate of 250 kbps.

The CC2420 is a low-cost, highly integrated solution for robust wireless communication in the 2.4 GHz unlicensed ISM band [5] [7] [10].

#### C. Installation

The Mote is able to connect an Optical Conditioner with Universal Asynchronous Receiver Transmission (UART) [2] [5] [10].

The UART connection is embedded in the MSP430F2617. The resolution used in UART is 13-bit. The format of UART is a start bit, seven or eight bits of data, a parity bit, an address bit, and one or two stop bits. The baud rate of UART can be detected automatically. The processor sends the value when it receives to the base station via radio [2].

### VI. SYSTEM SOFTWARE

The system software includes the two parts, the embedded software of MSP430F2617 and the one of the Upper Machine.

#### A. Software of MSP430F2617

The main program flow of the proposed system is shown in Fig. 5. The MSP430F2617 will be initialized first, include the initialization time and when the time is expired the mote detect the signal from the optical conditioner, if the signal is defer then value registered by the mote can send a packet to the gateway.

If the signal is not defer then recent value the Mote can initialize the timer for new instance [10].

#### B. Development of sensor data management software

We can use the sensor Middleware to manage the sensor and reconfigured mote, the time of signal sends it is modified from user wirelessly [10].

#### C. Flow Alarm circuit of Damage Detection

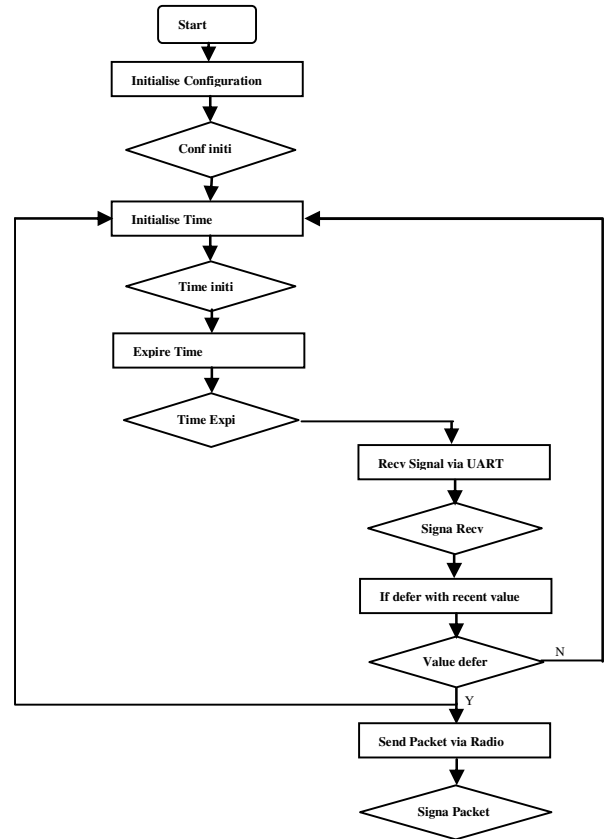


Fig. 5 Main Program Flow.

If the value of sensor measurement is equal or higher than the value of threshold memories in the mote the alarm is lunched [10].

#### D. Wireless communication protocol

In wireless sensor networks we can use the CTP protocol to achieve the destination of packet. We based on the Residual Energy and the Link Quality Indicator. The Link Quality Indicator and Energy Residual of the nodes can increase the performance of network. The link quality increases when the hop count decreases thus the node with high value of link quality will be chosen as the forwarding candidate. The threshold for the link quality can change the throughout, the lifetime of the network it makes more alive by selecting the longer routes with more energy. The optimal path will be chosen when the route has more Residual Energy or through Link Quality Indicator.

### VII. CONCLUSION

This article developed a prototype of wireless sensor bridge health monitoring system. The wireless communication,



the sensor and the node are used to make the system integration. Digital-Analogic and the electric to optical converters between module ZigBee and the optical sensors, the optical electric and the Analogic-Digital converters between sensors and module ZigBee.

This system able to detect any variation in the Bridge depends on the sensor parameters.

In addition, a protocol CTP can transmit the data between the sensors and Gateway. The sensors were installed on the Bridge.

#### REFERENCES

- [1] Barroca N. Borges LM. Velez FJ. Monteiro F. Gorski M. Castro-Gomes J. "Wireless sensor networks for temperature and humidity monitoring within concrete structures", *Construction and Building Materials*, Vol. 40, pp. 1156–66, March 2013.
- [2] Chae M.J. Yoo H.S. Kim J.Y. Cho M.Y. Development of a wireless sensor network system for suspension bridge health monitoring. *Automation in Construction*, Vol. 21, pp. 237-252, 2012.
- [3] Chapeleau X. Sedran T. Cottineau L.-M. Cailliau J. Taillade F. Gueguen I. Henault J. Study of ballastless track structure monitoring by distributed optical fiber sensors on a real-scale mockup in laboratory, *Engineering Structures*, Vol. 56, pp.1751-1757, 2013.
- [4] Chen Q. Lu P. Fiber Bragg Gratings and Their Applications as Temperature and Humidity Sensors. *Advances In Atomic, Molecular, and Optical Physics*, Vol. 56, pp. 235-260, 2008.
- [5] Jafer E. Ibala CS. Harris J. Design and development of multi-node based wireless system for efficient measuring of resistive and capacitive sensors. *Sensors and Actuators A: Physical*, Vol. 189, pp. 276-287, 2013.
- [6] Kronenberg P., Rastogi P.-K., Giaccari P., Limberger H.-G., Relative Humidity sensor with optical fiber bragg gratings. *OPTICS LETTERS*, Vol. 27, No. 16, pp. 1385-1387, 2002.
- [7] Liu Z. Yu Y. Liu G. Wang J. Mao X. Design of wireless measurement system based on WSNs for large bridges. *Measurement*, Vol. 50, pp. 324-330, 2014.
- [8] Sandra F.-H.-C., Paulo A., Edison P., Patrícia P.-L., Humberto V., Luis D.-C., Rute A.-S.-F., Paulo S.-A., Optical Fiber Relative Humidity Sensor Based on a FBG with a Di-Ureasil Coating. *Sensors* Vol. 12, pp. 8847-8860, 2012.
- [9] Xia Y. Zhang P. Ni Y. Zhu H. Deformation monitoring of a super-tall structure using real-time strain data. *Engineering Structures*, Vol. 67, pp. 29-38, 2014.
- [10] Yu Y. Zhao X. Shi Y. Ou J. Design of a real-time overload monitoring system for bridges and roads based on structural response. *Measurement* vol. 46, pp. 345-352, January 2013.

# OVA Heuristic Approach Based on Variable Neighborhood Search Method for Multiclass Support Vector Machine

B. Sidaoui, K. Sadouni.

**Abstract**—this paper describes a simple heuristic approach for solving Multiclass problems with Support Vector Machine (SVM). In this approach we start with the construction of binary tree, using One-Versus-All strategy and two criteria's of natural classification; Separation and Homogeneity. Then an optimization of the binary tree by Variable Neighborhood Search method is applied with the aim to explore the search space and to avoid the problem of local minima, the search process of optimal tree is guided by the One-Versus-All strategy. On the other hand, the proposed paradigm decomposes the multiclass problem to several bi-classes sub-problems, in order to facilitate the resolution of real and complex problems. Our aim is this paper, is to improve the temporal complexity of multiclass SVM by reducing the Support Vectors number and decrease the recognition time. This combination leads to much faster convergence comparing with state of the art methods, both in terms of running time in training and test phases and of classification accuracy. In this context, we have evaluated our approach on three corpus; ISOLET, PENDIGITS and MNIST. Encouraging results have been achieved.

**Keywords**—SVM Multiclass, One-Versus-All, Variable Neighborhood Search, Classification, Binary Tree.

## I. INTRODUCTION

Support Vector Machines (SVM) introduced in [5, 7, 9] are arguably the single most important development in supervised classification in the recent years. Support Vector Machines have been successfully used for the solution of a large class of machine learning tasks such as categorization, prediction, novelty detection, ranking and clustering [3, 4]. The success of SVM on a number of high dimensional tasks has prompted a renewed interest in kernel methods.

While SVM are a very robust and powerful technique for supervised classification, the large size and slow query time of a trained SVM is the main hindrance to their practical application especially for Multiclass problems. Knowing that,

originally SVM are binary classifiers and its extension to multiclass problems is not straightforward. The challenge is how effectively SVM extends for solving multiclass classification problem of many real-world applications such as: recognition of handwritten digits, automatic speech recognition, bioinformatics ...etc. Several multiclass methods have been proposed which successfully help to alleviate this problem. The most popular and widely used methods for applying SVM to multiclass problems usually decompose the original problem into several binary class problems, as the One-Versus-One and One-Versus-All strategies [8, 12].

In this context, a framework is proposed for applying SVM to multiclass problems. Precisely, we propose a heuristic based on two ideas: construct the binary tree using One-Versus-All strategy then find the optimal tree by an optimization algorithm called; Variable Neighborhood Search (VNS). More details are given in the next sections.

Both recognition of handwritten digits and automatic speech recognition have been the focus of machine learning techniques. The large variability and richness of voice data represent a fertile field to evaluate the performance of recognition systems. In this paper, the task of letters classification and handwritten digits are used to evaluate our work.

The remainder of the paper is organized as follows. First, we describe Support Vector Machines (SVM), their formulation and the most popular multiclass SVM. In the next section, we introduce our framework for multiclass SVM: binary tree and VNS algorithm. Implementation details are also given followed by experiments on diverse datasets. We conclude in the last section our paper with some interesting remarks and conclusions.

## II. SUPPORT VECTOR MACHINES

The main idea of binary SVM is to implicitly map data to a higher dimensional space via a kernel function and then solve an optimization problem to identify the maximum-margin hyper-plane that separates training instances [8]. The separator is based on a set of boundary training instances (training examples). Kernels can be interpreted as dissimilarity measures of pairs of objects in the training set  $X$ . In standard SVM formulations, the optimal hypothesis sought is of the form (1).

This work was by project research No.B\*01920100056of National Committee of evaluation and programming for scientific research of the Minister of Higher Education and Scientific Research, Algeria.

Bouthkil Sidaoui was with Mathematics and Computer Science Department, University center of Naama 'Salhi Ahmed', Naama, 45000, Algeria, (corresponding author : email: b.sidaoui@gmail.com).

Kaddour Sadouni, was with Computer Science Department, Faculty of Mathematics and Computer Science, University of Science and Technology of Oran 'Mohamed Boudiaf', BP1505 El-M'naouer Oran, 31100, Algeria, email: kaddour.sadouni@univ-usto.dz.

$$\Phi(\xi) = \sum \alpha_i k(x, x_i) \quad (1)$$

Where  $\alpha_i$  are the components of the unique solution of a linearly constrained quadratic programming problem, whose size is equal to the number of training patterns. The solution vector obtained is generally sparse and the non-zero  $\alpha_i$ 's are called support vectors (SV's). Clearly, the number of SV's determines the query time which is the time it takes to predict novel observations and subsequently, is critical for some real time applications such as speech recognition tasks. It is worth noting that in contrast to connectionist methods such as neural networks, the examples need not have a Euclidean or fixed-length representation when used in kernel methods.

The training process is implicitly performed in a Reproducing Kernel Hilbert Space in which  $k(x, y)$  is the inner product of the images of two examples  $x, y$ . Moreover, the optimal hypothesis can be expressed in terms of the kernel that can be defined for non-Euclidean data such biological sequences, speech utterances etc. Popular positive kernels include the Polynomial (2) and the Gaussian (3), kernels:

$$k(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \quad (2)$$

$$k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (3)$$

#### A. SVM Formulation

Given training vectors  $x_i \in \mathcal{R}^n, i = 1, \dots, m$ , in two classes, and a vector  $y \in \mathcal{R}^m$  such that  $y_i \in \{1, -1\}$ , Support Vector Classifiers [7, 8, 9] solve the following linearly constrained convex quadratic programming problem:

$$\begin{aligned} \text{maximize } W(\alpha) &= \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i, x_j) - \sum_{i=1}^m \alpha_i \\ \text{subject to } &\forall i, 0 \leq \alpha_i \leq C \\ &\sum_{i=1}^m \alpha_i y_i = 0 \end{aligned} \quad (4)$$

The optimal hypothesis is:

$$f(x) = \sum_{i=1}^m \alpha_i y_i k(x, x_i) + b \quad (5)$$

Where the bias term  $b$  can be computed separately. Clearly, the hypothesis  $f$  depends only on the non-null coefficients  $\alpha_i$  which correspond to the Support Vectors (SVs). The quadratic programming (QP) objective function involves the problem Gram matrix  $k$  whose entries are the similarities  $k(x_i, x_j)$  between the patterns  $x_i$  and  $x_j$ . It is important to

note, on one hand, that the pattern input dimension  $d$ , in the above formulation, is implicit and does not affect to some extent the complexity of training, provided that the Gram matrix  $k$  can be efficiently computed for the learning task at hand. On the other hand, the patterns representation is not needed and only pairwise similarities between objects must be specified.

This feature makes SVM very attractive for high input dimensional recognition problems and for the ones where patterns can't be represented as fixed dimensional real vectors such as text, strings etc. For large scale corpora however, the quadratic programming problem becomes quickly computationally expensive, in terms of storage and CPU time. It is well known that general-purpose QP solvers scale with the cube of the problem dimension which is, in our case, the number of training patterns  $m$ . Specialized algorithms, typically based on gradient descent methods, achieve impressive gains in efficiency, but still become impractically slow for problems whose size exceeds 100,000 examples. Several attempts have been made to overcome this shortcoming by using heuristically based decomposition techniques such as Sequential minimal optimization SMO [4] implemented in LibSVM package [20].

#### B. Multiclass SVM

Many real applications consist of multiclass classification problems. Unfortunately, Support Vector Machine is intrinsically bi-class and its efficient extension to multiclass problems is still an ongoing research issue [12, 13, 14]. Several frameworks have been introduced to extend SVM to multiclass contexts and a detailed account of the literature is out of the scope of this paper. The most common way to build a multiclass SVM is to combine several sub-problems that involve only binary classification. This idea is used by various approaches as One-Versus-All (OVA), One-Versus-One (OVO) [8, 12] and Directed Acyclic Graph (DAGSVM) [13].

In OVA strategy,  $K$  binary classifiers are constructed, where  $K$  is the number of classes. The classifier is trained by labeling all the examples in the class as positive and the remainder as negative. The final hypothesis is given by the formula:

$$f_{ova}(x) = \arg \max_{i=1, \dots, K} (f_i(x)) \quad (6)$$

In the second strategy, OVO proceeds by training  $K(K-1)/2$  binary classifiers corresponding to all the pairs of classes. The hypothesis consists of choosing either the class with most votes (voting). In Directed Acyclic Graph (DAGSVM), each node of graph represents a binary classifier (DAGSVM).

The multiclass approaches; OVO and DAGSVM, are cheaper to train in terms of memory and computer speed than OVA. The Authors in [13] investigated the performance of

several SVM multi-class paradigms and found that the One-Versus-One achieved slightly better results on some small to medium size benchmark data sets. Furthermore, other interesting implementations of multi-class SVM have been developed recently, such as, the two architectures proposed in [14, 15], in the first Cha and Tappert build the decision tree by genetic algorithms where they considered each tree as chromosome, while in the second architecture Madzarov and All propose a clustering method to construct the binary tree. Another approach in this direction has been presented by Lei and Venu in [1], they use a recursive method to build the binary tree.

In term of complexity, OVA approach needs to create  $K$  binary classifiers; the training time is estimated empirically by a power law [2] stating that  $T \approx \alpha M^2$  where  $M$  is the number of training samples and  $\alpha$  is proportionality constant. According to this law, the estimated training time for OVA is:

$$Train\_Time_{OVA} \approx k\alpha M^2 \quad (7)$$

Without loss of generality, let's suppose that each of the  $K$  classes has the same number of training samples. Therefore, each binary SVM-OVO approach requires  $2M/k$  samples. Hence, the training time for OVO (and also for DAG) is:

$$Train\_Time_{OVO} \approx \alpha \frac{k(k-1)}{2} \left( \frac{2M}{k} \right)^2 \approx 2\alpha M^2 \quad (8)$$

In the testing phase, OVA strategy requires  $K$  binary SVM evaluations and OVO strategy  $K(K-1)/2$  binary SVM evaluations, while DAGSVM performs faster than OVO and OVA, since it requires only  $K-1$  binary SVM evaluations. The approaches proposed in [1, 14 and 15] are even faster than DAGSVM because the depth of the binary tree is  $\lceil \log_2(k) \rceil$ . In addition, the total number of support vectors (SVs) in all models will be smaller than the total number of SVs in the others (OVA, OVO and DAGSVM). Therefore it allows quick convergence in test phase. Recently, approaches based on optimization and estimation of several SVM parameters, are presented in [17, 18 and 19].

The advantage of the approaches presented in [1, 14, 15] and the approach proposed in this paper lie mainly in the test phase, because it uses only the models necessary for recognition, which make the testing phase faster. However a problem of local optima appears when searching the optimal binary tree [1, 14 and 15]. To avoid this problem, we propose in this paper to find the optimal binary tree, using the similarity partitioning algorithms; and an optimization based on Variable Neighborhood Search (VNS) method. The detail of One-Versus-All Tree multiclass proposed in this paper is discussed in the next section.

### III. APPROACH MULTICLASS BASED ON BINARY TREE

Our approach uses multiple binary SVMs in a binary decision tree structure [10]. For each node of the binary tree, a binary SVM is associated. All samples in the node are assigned to the two sub-nodes derived from the current node. This step is repeated at every node until each node contains only samples from one class, e.g until reaching leaves of the tree. In this context, many works have been proposed [1, 14 and 15], but the main problem that should be considered seriously here is how to construct the optimal tree? In other words, how we can do a correct partitioning of the training samples in two groups, in each node of the tree? In this paper, we introduce a new multiclass approach called VNS-Tree Multiclass for Variable Neighborhood Search Tree for solving multiclass problems with Support Vector Machines. The main idea is to construct an optimal binary tree by VNS method. Our contribution is divided in two steps: construction of the binary tree and, implementation of the tree result.

#### A. Construction of Binary Tree

Binary tree constructing is based on two ideas: the principle of One-Versus-All approach and, tree optimization by Variable Neighborhood Search algorithm. The aim is to construct a binary tree where each node represents a partition of two classes (one class against rest), while the leaves represent the classes' labels'. For each node (we begin from the root) we exploit the main idea of the OVA approach for calculate the best hyper-plane separating one class against the rest, while using two similarity criteria; i.e. for each node the best partition of two subsets of classes, where one subset contain just one class.

Furthermore, for a fixed number of classes, the optimal partition is not necessarily unique. To avoid this problem, Variable Neighborhood Search optimization algorithm is used. The construction process is composed on two phases: the first consist to create a binary tree and then, an optimization of the binary tree is applied in the second phase.

In the first phase, we analyze several criteria specific to each family by systematically applying the same optimization strategy: a number of classes are fixed and, at each step, we construct a binary partition in each node of binary tree. At each level of the tree, we try to build a partition of fixed number of classes (equal to 2) that optimizes a certain criterion. In this work, we begin with a partition which has one group containing all classes. So we start with root (up) of tree, and we looks for the class of larger diameter (farthest compared to the other). We use the same procedure until no partitioning is possible.

There are three families of natural classification criteria; separation, homogeneity and dispersion. In this paper, we used only separation and homogeneity criteria.

In the first one, we seek to maximize the differences between classes, which are functions of distances between classes, i.e. the smallest distances inter-classes and greater distances possible intra-class (in our case: two classes in each partition). This top-down approach consists to separate the

class which is the most distant for the remaining classes. We calculate a distance between all pairs  $(x_i, x_j)$  classes prototypes and we deduce the class most distant from the remainder.

We obtain the first partition (root of the tree) which contains all classes. This process is repeated only for the remainder; until there is only one class per group (no partitioning is possible). It is well known that for this type of optimization problem, the optimal partition is not necessarily unique, particularly if there are values of equal distance, there may be several optimal trees and different partitions with the same optimal value. Finally, we calculate the total inertia of each tree (this is practically possible for a value not big of  $k$ ) defined by:

$$f_{\text{Overallinertia}} = \sum_{i=1}^L \text{Inertia}(i) \quad (9)$$

Where;  $L$  is the length of the tree. Then the global optimum is:

$$\text{Optimum\_tree} = \text{argmax}\{f_{\text{Overallinertia}}, i = 1 \dots T\} \quad (10)$$

Where  $T$  is the number of all possible trees.

In Homogeneity, the classes are as concise as possible; it is desired to minimize the diameter, that is, the maximum intra-distances classes. In this case, we generate partitions by an ascendant clustering technique. We are constructing binary partitions, which one of the two classes containing only one element, by grouping the closest classes. This bottom-up approach is similar to a hierarchical classification method.

Initially, we compute a similarity matrix of all pair's  $(x_i, x_j)$  prototypes of classes and then select the smallest similarity, which gives us the partition of the last level of the tree (the smallest partition):

$$\text{partition} = \{x_i, x_j\} \quad (11)$$

Where  $(x_i, x_j)$  have minimal distance:

$$d(x_i, x_j) = \min\{d(x_i, x_j); i, j = 1 \dots K\} \quad (12)$$

This process is repeated, updating the gravity centers at each level, till the root of the tree. To prevent the problem of local minima, we calculate the total inertia defined by the formula (12) of each tree.

Finally in Dispersion, we minimize a function of inertia, the sum of squared deviations in a center, whether real or virtual.

The OVA Tree Multiclass method that we propose is based on recursively partitioning the classes in two disjoint groups

(one contains one class only) in every node of the binary tree, and training a SVM that will decide in which of the groups the incoming unknown sample should be assigned. In the general case, the number of partitions into two parts (groups) of a set of  $k$  elements is given by the following formula [6]:

$$N_{\text{partitions}_{k,2}} = \sum_{i=0}^2 (-1)^i \frac{(2-i)^k}{i!(2-i)!} \quad (13)$$

For any partition problem, two cases can be considered: if the number  $k$  is small ( $k < 6$ ), we calculate all possible partitions and then deduce the optimal partition, else we determine the optimal partition by optimization methods, because it is impossible to cover all possible partitions. To avoid the problem of local minima and to obtain the optimal binary tree, a VNS optimization algorithm is used.

In this paper, we start by a set  $X$  of training samples labeled by  $y_i \in \{c_1, c_2, \dots, c_k\}$ , each OVA Tree Multiclass is summarized in two phases: Calculate  $k$  gravity centers and then, find the right partition of  $k$  gravity centers into two groups, by using separation and homogeneity criteria. For that, a function of overall inertia is calculated for each binary tree by summing inertias of various partitions of the tree.

### B. Optimization of Binary Tree

Metaheuristics are general framework to build heuristics for combinatorial and global optimization problems; our problem (binary tree optimization) is included in this context. Variable Neighborhood Search (VNS) [11], is a recent Metaheuristic which exploits systematically the idea of neighborhood change, both in the descent to local minima and in the escape from the valleys which contain them. It exploits the following notions:

- Systematic change of neighborhood in search.
- A local optimum with respect to one neighborhood structure is not necessarily so for another.
- A global optimum is local with respect to all possible neighborhood structures.

#### 1) Variable Neighborhood Search algorithm

Initialization: Select a set of neighborhood structures  $N_k$  with  $k = 1, \dots, N_{\text{max}}$ , find an initial solution  $x$ , set  $k = 1$ , and choose a stopping condition.

- Step 1: Generate randomly  $x' \in N_k(x)$
- Step 2: Apply a local search method starting with  $x'$  to find local optimum  $x''$ .
- Step 3: If  $x''$  is better than the current solution, then set  $x = x''$  and  $k = 1$ , otherwise set  $k = k + 1$ ; go back to step 1.

## 2) Optimization of Binary Tree with VNS

In this paper, we are interested to Variable Neighborhood Search (VNS) algorithm for optimize the binary tree, for its simplicity and efficiency. The application of VNS is resumed as follows:

- Initially, we defined a set of neighborhoods of the tree  $T(k)$ ,  $k \in \mathbb{N}$ . In our case,  $k$  is the set of labels and therefore the set  $T(k)$  is all subspaces referenced by all binary partitions one-versus-all, of all classes. That strategy allows exploring the different regions of feasible solutions space.
- Generate an initial solution (binary tree where each node is one-versus-all classes).
- For each subspace of solution (binary trees) or neighborhood, we apply a local search to find a local optimum.
- Iterates until a global optimum is reached; tree with greater total inertia.

## C. Complexity of VNS-Tree Multiclass

The aim of this work is to improve the complexity of the SVM multiclass approach based on binary tree and to solve the problem of local minima. For that latter, a VNS algorithm is used. The complexity of VNS-Tree multiclass approach is clearly better compared to other approaches.

In the training phase, our approach One-Versus-All Tree multiclass has  $(K - 1)$  binary classifiers ( $K$  is the number of classes). The random structure of the optimal tree complicates the calculation of the training time. However, an approximation can be given. Assume that each of the  $K$  classes has the same number of training samples. The training time is summed over all  $(K - 1)$  nodes in the different levels of tree. In the  $i^{th}$  level, there are one node and each node uses,

$M - \left(\frac{M}{K}\right)(i - 1)$  training samples. Hence, the total training time is:

$$Train\_Time_{VNSTree} \approx \sum_{i=1}^{K-1} \alpha \left( M - \left(\frac{M}{K}\right)(i - 1) \right)^2 \quad (14)$$

$$Train\_Time_{VNSTree} \approx \alpha M^2 \quad (15)$$

It must be noted that the training time of our approach does not include the time to build the hierarchy structure (binary tree) of the  $k$  classes. Furthermore, in test or recognition step, our work is faster than state of the art, for several reasons: the simplicity of the binary architecture, the reduced number of Support Vectors, small number of models recognition (classifiers) and the strategy used in recognition.

## D. Implementation of VNS-Tree Approach

To implement the optimal tree obtained in construction step, we take advantage of both the efficient computation of the tree architecture and the high classification accuracy of SVM utilizing this architecture.  $k-1$  binary SVM classifiers have been trained for a problem of  $k$  class, where each node is represented by a SVM classifier, i.e. we construct  $k-1$  models of binary classifiers.

We start from the root node of the tree, using the samples of the binary partition (one versus all), so we first group as positive examples and the samples of the second group as negative examples. The classes from the first group will be assigned to the left sub-tree, while the classes of the second group will be assigned to the right sub-tree. This recursive process (dividing each of the groups into two subgroups applying the procedure explained above) is repeated, until there is only one class per group, which defines the leaves of the decision tree. Finally, we obtain  $k-1$  SVM models, which are used for the recognition of new examples.

The recognition of each test example starts at the root of the tree. At each node of the binary tree, a decision is made about the assignment of the input pattern into one of the two possible groups (descent of the current node), i.e. we affect the query pattern to the left or to the right sub-tree. Each of these groups may contain multiple classes. This is repeated recursively downward the tree until the pattern arrive a leaf node that represents the class it has been assigned to.

## IV. EXPERIMENTS AND RESULTS

In this paper, we performed our experiments on three corpuses: ISOLET corpus [21] consisting of 5999 examples for training and 1559 examples for test labelled by 26 letters (classes): {A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z}. PENDIGITS corpus [22] consists of 7494 examples for training and 3590 for test labelled by {0, 1, 2, 3, 4, 5, 6, 7, 8, 9} classes. MNIST corpus [23] consists of 60,000 and 10,000 examples for training and test sets respectively of 10 digits classes: {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}. For the three datasets used in this work, any pre-processing has been required.

In all the experiments reported below, we performed cross validation for tuning SVM hyper parameters. The GNU SVM light [24] implementation is used for our OVA Tree Multiclass Machine used in this paper, and LibSVM for SVC [25], software to compare our results. The following tables: Table 1, Table 2 and Table 3, summarize the best results obtained in the experimentation phase. In this paper, supervised classification task concerns two problems: letters classification problem with ISOLET corpus and handwritten digits problem with two datasets; PENDIGITS and MNIST.

The hyper-parameters of binary SVM used in our work are:  $\gamma$ ; kernel type; takes the following values, linear (0), Polynomial (1), Gaussian (2), C; regularization parameter of SVM,  $g$ ; Gaussian kernel parameter. The C and  $g$  were calculated by cross validation. Time (s) is time of prediction or test,

Error(\%) is the error rate of prediction. The three datasets are experimented for the different classifiers: OVA (One Versus All), OVO (One Versus One), VNS-Tree-Hom (VNS-Tree with Homogeneity) and VNS-Tree-Sep (VNS-Tree with separation).

Table 1: ISOLET Results.

Classifier	T	C	D	Error (%)	Time (s)
OVA	0	10	/	01,67	532
OVO	1	100	2	01,68	489
VNS-Tree-Hom	1	0	2	02,11	360
VNS-Tree-Hom	1	1000	2	01,98	320
VNS-Tree-Sep	1	100	2	02,13	310
VNS-Tree-Sep	1	10	2	02,20	280

Table 2: MNIST results.

Classifier	T	C	D	Error (%)	Time (s)
OVA	1	10	3	03,28	15
OVO	1	100	4	03,27	17
VNS-Tree-Hom	1	100	3	03,66	14
VNS-Tree-Hom	1	100	4	03,73	13
VNS-Tree-Sep	1	100	3	04,30	12
VNS-Tree-Sep	1	100	4	04,11	14

Table 3. PENDIGITS Results

Classifier	T	C	g	Error (%)	Time (s)
OVA	0	15	/	01,20	42
OVO	2	10	0,00008	01,28	48
VNS-Tree-Hom	2	100	0,0001	01,52	40
VNS-Tree-Hom	2	100	0,00008	01,78	40
VNS-Tree-Sep	2	100	0,00015	01,69	40
VNS-Tree-Sep	2	100	0,0001	01,52	30

## V. ANALYSIS OF RESULTS

In this study, we find to improve the complexity of SVM machine, in particularity for the big datasets and the problems who's the number of labels is important. It is clear that the proposal approach improve clearly the recognition time as shown figure (figure. 2). In terms of error rates, the results are similar to those of the OVO and OVA approaches as shown figure (Fig. 1). We have not won big thing in this sense. It is very logical because we have based on the binary classifier as OVA.

This research may be considered as a contribution to improve the temporal complexity, as shown figure (figure. 2), with MNIST datasets (of 70000 examples for training and test sets). We believe that the results obtained by this study is encouraging step to improve the performance of our approach by an optimization of the hyper parameters of binary SVM and the uses of big data.

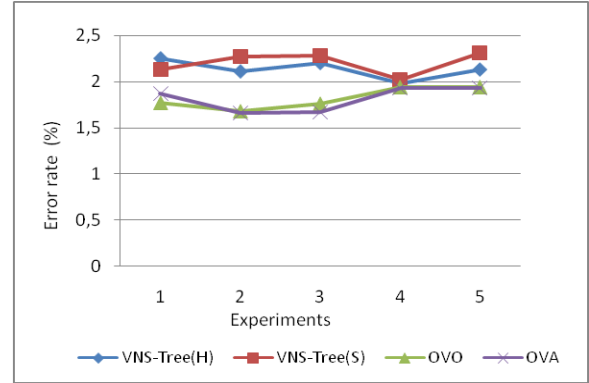


Fig. 1 Variation of error rate for different approaches

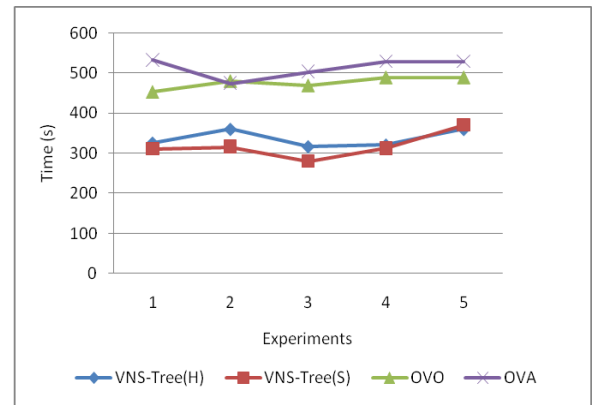


Fig. 2 Variation of recognition time for different approaches

## VI. CONCLUSION

We introduced and implemented an efficient framework for multiclass SVM using separation and homogeneity criteria combined with an optimization algorithm; Variable Neighborhood Search (VNS). The binary tree of support vector machine multiclass paradigm was shown through extensive experiments to achieve state of the art results in terms of accuracy. The ultimate objective of this study was improving the complexity and the performance of multiclass SVM machine, using a VNS method and the binary tree architecture.

The experiments carried in this paper for investigating our binary architecture named VNS-Tree Multiclass indicate that the results are comparable to those of other multiclass approaches, particularly in [16] where the same corpuses was used. In this work, we have tried to improve the complexity of multiclass SVM. Furthermore, our approach is faster, due to its logarithmic complexity, and the reduced number of support vectors (SV). Our aim was to improve the complexity of SVM multi class, in terms of test time. The structure of our approach oriented binary tree, allow in one hand to reduce the number of SV and, on the other hand, it used the necessary models for classification.

Nevertheless, there are various challenges for the improvement of SVM i.e. the optimization of SVM hyper parameters which can improve considerably this type of contribution. We believe also that, in order to improve the pattern recognition technology, more research efforts must take advantage of the solid mathematical basis, and the power of binary SVM should be invested in developing general purpose efficient machine learning paradigms capable of handling large scale multi-class problems.

## REFERENCES

- [1] Hansheng L, Govindaraju V, "Half-against-half multi-class support vector machines", In Multiple Classifier Systems, pp. 156-164, 2005.
- [2] Platt J, "Fast training of support vector machines is using sequential minimal optimization", In Advances in Kernel Methods-Support Vector Learning, pp. 185-208, 1999.
- [3] Joachims T, "Making large-scale support vector machine learning practical", In Advances in Kernel Methods- Support Vector Learning, pp. 169-184, 1999.
- [4] Osuna E, et All, "Training support vector machines, an application to face detection", In Proceedings of CVPR'97, vol 5, pp. 130-136 1997.
- [5] Boser B, Guyon I M, Vapnik V, "A training algorithm for optimal margin classifiers", In Proceedings of AWCLT, pp. 144-152, 1992.
- [6] Jain A and all, "Algorithms for clustering data", 334. Prentice Hall Advanced, New York, USA, 1988.
- [7] Vapnik V, "The nature of statistical learning theory", 314. Springer, New York, USA, 2010.
- [8] Vapnik V, "Statistical learning theory", 736. A Wiley Interscience publication, New York, USA, 1998.
- [9] Vapnik V, Chervonenkis A, "Theory of pattern recognition". USSR, Nauka, Russian, 1974.
- [10] Yann Y, Bottou L, Bengio Y, Patrick H, "Gradient-based learning applied to document recognition", Proceedings of the IEEE, vol 86, pp. 2278-2324, 1998.
- [11] Mladenovic N, Hansen P, "Variable neighborhood search, Principles and applications", European Journal of Operational Research, vol 130, pp. 449-467, 2001.
- [12] Platt J C, Taylor J S, "Large margin dags for multiclass classification", Advances in Neural Information Processing Systems, 12, pp. 547-553, 2000.
- [13] Hsu C, Lin J, "A comparison of methods for multiclass support vector machines", IEEE Transactions on Neural Networks, 13, pp. 4-15, 2002.
- [14] Cha S H, Tappert C, "A genetic algorithm for constructing compact binary decision trees", Journal of Pattern Recognition Research, 1, pp. 1-13, 2009.
- [15] Gjorgji M, Chorbev I, "A multi-class svm classifier utilizing binary decision tree", Journal of Informatica, 33, pp. 233-241, 2008.
- [16] Rifkin R, Klautau A, "In defense of one-vs-all classification", Journal of Machine Learning Research, 5, pp. 101-141, 2004.
- [17] Vretos P N, Tefas A, "Using robust dispersion estimation in support vector machines", Journal of Pattern Recognition, 46, pp. 3157-3558, 2013.
- [18] Mikel Galar E B, Fernandez A, "Dynamic classifier selection for one-vs-one strategy: Avoiding non-competent classifiers", Journal of Pattern Recognition, 46, pp. 3412-3424, 2013.
- [19] {WenmingZheng} Wenming Zheng Z L, "A new discriminant subspace analysis approach for multi-class problems", Journal of Pattern Recognition, 45, pp. 1426-1435, 2012).
- [20] A library for support vector machines, developed by Chang and Lin, available at: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [21] Isolated letter speech recognition database created by R. Cole and M. Fanty, available at: <http://archive.ics.uci.edu/ml/machine-learning-databases/isolet/>
- [22] Pen-based recognition of handwritten digits database created by Alpaydin, available at: <http://archive.ics.uci.edu/ml/machine-learning-databases/pendigits/>
- [23] database of handwritten digits database created by yann, available at: <http://yann.lecun.com/exdb/mnist/>
- [24] Software developed by Joachims, available at: <http://svmlight.joachims.org>
- [25] Libsvm toolkit of the improved core vector machine, training algorithms using core-set approximation, Software of Joachims, available at: <http://www.cs.ust.hk/~ivor/cvm.html>



# An Image Fusion Algorithm Based on Wavelet Transform and Tensor Analysis

Yongping Zhang, Dechun Zheng, Zhongkun He

School of Electronic and Information Engineering,  
Ningbo University of Technology,  
Ningbo, China

email: ypz@nbut.edu.cn, zhengdc197543@sina.com.cn, hzk@nbut.edu.cn

**Abstract**—How to select and reconstruct the different subband coefficients is a key issue in image fusion based on wavelet transform. In this paper, a novel algorithm is proposed for the selection and reconstruction of the coefficients in wavelet domain. For choosing the low frequency subband coefficients, a scheme based on the regional standard deviation and mean of the high frequency sub-images, which can combine with low frequency and high frequency subband coefficients, is provided. To reconstruct the high-frequency sub-band coefficient, a novel method based on the structure tensor analysis is employed. In the method, the high frequency subband coefficients in horizontal and vertical orientation are reconstructed by using eigenvalue analysis of the weighted structure tensor. The method can obtain more detailed information and enhance contrast. Experimental results show that the proposed algorithm can get a better fusion results in terms of both visual effects and quantitative indicators, compared to traditional wavelet based algorithm.

**Keywords**—image fusion, multi-focus images, wavelet transform, structure tensor matrix, local standard deviation

## I. INTRODUCTION

Image fusion technology is playing a crucial role in many fields such as automatic scale recognition, remote sensing, multimodal medical imaging, surveillance, machine vision, and so on. The goal of image fusion is to integrate multiple images of the same scene into a composite image so that the new image is more suitable for visualization, detection or recognition tasks.

In the study, we aim at the multi-focus image fusion algorithm based on wavelet transform. These years, the wavelet transform has become a useful tool in multi-resolution image fusion [1, 2, 7]. Those existing image fusion algorithms based on wavelet transform use Mallat algorithm [3] in combination with certain selection rules to perform image decomposition and image information fusion.

The key issue in wavelet-based image fusion is how to appropriately merge the sub-band coefficients. The

traditional wavelet-based image fusion algorithms adopt the maximal selection rule and weighted average to reconstruct the wavelet coefficients. Generally, the maximal of absolute values is employed as a criterion to choose the high-frequency coefficients, while the low-frequency coefficients are computed by using the method of weighted average [4, 5], the fused results then are obtained by using wavelet reconstruction algorithm. The main drawback of such a method is that it will lead to the Gibbs phenomenon. Hereafter, a series improved methods have been developed. In order to overcome the limitation of only considering the grayscale values of a single pixel, an area-based maximum selection rule has been provided in [7]. In addition to this, the local energy [9], and the local gradient selection [8, 9] have also been provided to improve the fusion performance.

In this paper, a new wavelet-based image fusion method is developed which does not adopt simple selection criterion, but uses structure tensor analysis method to reconstruct the horizontal and vertical sub-band coefficients, uses the regional standard deviation and mean of the high frequency subband to choose the low frequency subband coefficients. The diagonal high frequency sub-band coefficients are obtained via the regional standard deviation choosing. While the high-frequency sub-band coefficients in both horizontal and vertical orientation are computed by using structure tensor analysis method. Finally the fused image is reconstructed by inverse wavelet transform.

The rest of the paper is organized as follows. Section II gives a brief review of image fusion algorithm based on wavelet transform. Section III presents the fusion rules of wavelet coefficients. The image fusion scheme and the experimental results are presented in Section IV, and the paper is concluded in Section V.

## II. WAVELET FUSION FRAMEWORK

The block diagram of an image fusion system is shown in Fig. 1. Here we assume two input images are image A and B, and image F is fused image, the basic steps of image fusion are as follows:

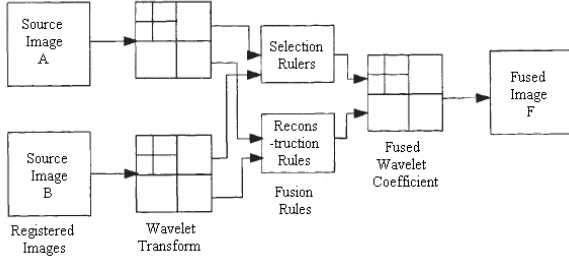


Fig. 1 Fusion Processing of Image Wavelet Transformation.

(1) Using wavelet transform to decompose source image A and B into sub-images  $\{Cn_{k_0}^L(i, j), Cn_k(i, j) \ (k \geq k_0)\}$  at multi-scale, where  $Cn_{k_0}^L(i, j)$ 、 $Cn_k(i, j)$  represents low frequency sub-image、high frequency sub-image in three directions, respectively.

(2) The wavelet coefficients of sub-images are fused using different fusion rules at different scale.

(3) The final fused image F is obtained by applying the inverse wavelet transform on fused wavelet coefficients.

In the next Section we will give a detailed discussion of selection and reconstruction of the subband coefficients.

### III. FUSION RULES

In order to obtain better quality in the fused image, we adopt the divide and rule strategy to reconstruct the different subband coefficients in wavelet domain. For the construction of low frequency subband and the diagonal high frequency subband, the maximal selection rule is used, while the construction of horizontal and vertical sub-band is performed by using tensor analysis method.

#### A. Selection of the low frequency subband coefficients

For the construction of low frequency subband coefficient, the simple method is to computing the average of coefficients corresponding to different input image. This method can effectively suppresses the noise in input images, but lose the useful information in images [10]. To overcome the drawback of simple averaging method, some improved methods have been proposed, such as the method based on local variance [11]. The limitation of those methods is without considering the role of high frequency sub-band.

A modified algorithm of selection low frequency sub-image is proposed, which based on the value of mean value and standard deviation of the corresponding high frequency coefficient. The basic steps of this method are as follows:

1) Add the absolute value of the corresponding high frequency coefficient, obtained sub-image  $CHr_{k_0}(i, j)$ , where  $Cn_{k_0}^L(i, j)$ 、 $Cn_{k_0}^H(i, j)$ 、 $Cn_{k_0}^V(i, j)$ 、 $Cn_{k_0}^D(i, j)$  represent

the low frequency、high frequency in horizontal、vertical and diagonal orientation, respectively. And  $n = a, b$  represents the image A, B

$$CHr_{k_0}(i, j) = |Cn_{k_0}^H(i, j)| + |Cn_{k_0}^V(i, j)| + |Cn_{k_0}^D(i, j)|. \quad (1)$$

2) Find the mean  $Ma(i, j)$ 、 $Mb(i, j)$  and the standard deviation  $SDa(i, j)$ 、 $SDb(i, j)$  of as activity level from image A and B, respectively.

3) Get fused low frequency sub-band coefficients  $F_{k_0}^L(i, j)$  from wavelet coefficients  $Cn_{k_0}^L(i, j)$ , and the fusion rules we used as follow:

$$F_{k_0}^L(i, j) = \begin{cases} Ca_{k_0}^L(i, j); & \text{if } SDa(i, j) > SDb(i, j) \\ & \text{and } Ma(i, j) > Mb(i, j) \\ Cb_{k_0}^L(i, j); & \text{if } SDa(i, j) < SDb(i, j) \\ & \text{and } Ma(i, j) < Mb(i, j) \\ \frac{(Ca_{k_0}^L(i, j) + Cb_{k_0}^L(i, j))}{2}; & \text{otherwise} \end{cases} \quad (2)$$

#### B. Reconstruction of high frequency subband coefficient in horizontal and vertical orientation

The tensor matrices are used to describe multi-image or vector-value images [12]. For a vector field  $\xi = (\xi_1, \xi_2)^T$  defined on domain  $\Omega$ , its tensor matrix at position  $(i, j) \in \Omega$  can be represented by

$$G(i, j) = \begin{pmatrix} (\xi_1)_{i,j}^2 & (\xi_1)_{i,j}(\xi_2)_{i,j} \\ (\xi_1)_{i,j}(\xi_2)_{i,j} & (\xi_2)_{i,j}^2 \end{pmatrix}. \quad (3)$$

This matrix is positive and semi-definite. Denoted by  $\lambda^+$  and  $\lambda^-$  the maximum and minimum eigenvalues of  $G(i, j)$ , respectively. The corresponding eigenvectors are denoted by  $\theta^+$  and  $\theta^-$ . It is easy to see that  $\lambda^+ = \|\xi\|^2$ ,  $\lambda^- = 0$ , and  $\theta^+ = \xi / \|\xi\|$ ,  $\theta^- = \xi^\perp / \|\xi\|$ . That is,

$$\xi = \sqrt{\lambda^+} \theta^+ \text{ and } G = \lambda^+ \theta^+ \theta^{+T}. \quad (4)$$

For two vector fields  $\xi^1 = (\xi_1^1, \xi_2^1)^T$  and  $\xi^2 = (\xi_1^2, \xi_2^2)^T$ , their weighted tensor matrix at position  $(i, j)$  can be represented as:

$$G_s(i, j) = \begin{pmatrix} \sum_k s_k^2(i, j) (\xi_1^k)_{i,j}^2 & \sum_k s_k^2(i, j) (\xi_1^k)_{i,j} (\xi_2^k)_{i,j} \\ \sum_k s_k^2(i, j) (\xi_1^k)_{i,j} (\xi_2^k)_{i,j} & \sum_k s_k^2(i, j) (\xi_2^k)_{i,j}^2 \end{pmatrix}. \quad (5)$$

where  $s_k(i, j)$  is weight function defined as:

$$s_k(i, j) = \frac{\|(\xi^k)_{i,j}\|}{(\|(\xi^1)_{i,j}\|^2 + \|(\xi^2)_{i,j}\|^2)^{1/2}}. \quad (6)$$

Since  $G_s(i, j)$  is also a positive semi-definite matrix, its eigenvalues are both real and non-negative. Denoted by  $\lambda^+$  and  $\lambda^-$  the maximum and minimum eigenvalues of  $G_s(i, j)$ , respectively. The corresponding eigenvectors are denoted by  $\theta^+$  and  $\theta^-$ . Then, we have

$$G_s = \lambda^+ \theta^+ \theta^{+T} + \lambda^- \theta^- \theta^{-T}. \quad (7)$$

We define the reconstructed vector  $\xi = (\xi_1, \xi_2)^T$  from the vector fields  $\xi^1 = (\xi_1^1, \xi_2^1)^T$  and  $\xi^2 = (\xi_1^2, \xi_2^2)^T$  as the follows:

$$\xi = (\xi_1, \xi_2)^T = \sqrt{\lambda^+} \theta^+ \text{sign}(\theta^+ \cdot \sum_k s_k \xi^k). \quad (8)$$

That is, the magnitude of the vector  $\xi_{i,j} = ((\xi_1)_{i,j}, (\xi_2)_{i,j})^T$  is the square root of the maximum eigenvalues of  $G_s(i, j)$  and the direction agrees with the average of  $\xi_{i,j}^1$  and  $\xi_{i,j}^2$ .

Based on the above notions, a new coefficient vector filed can be reconstructed from the vector fields  $((Ca_k^H)_{i,j}, (Ca_k^V)_{i,j})^T$  and  $((Cb_k^H)_{i,j}, (Cb_k^V)_{i,j})^T$ , where  $(Ca_k^H)_{i,j}$ 、 $(Ca_k^V)_{i,j}$  express the corresponding horizontal and vertical high frequency sub-images with wavelet decomposition, and  $n = a, b$  present the image A and B. In the new coefficient vector filed, the first component is the reconstructed horizontal high frequency sub-images  $F_k^H(i, j)$  and another component is the reconstructed sub-images  $F_k^V(i, j)$  in vertical direction.

#### C. Selection of diagonal high frequency sub-band coefficient

Variance is a measure of contrast, which reflects the detail of the images [7]. Here, in order to obtain the better results on the whole, the local regional standard deviation is used to select the diagonal high frequency sub-images. The fusion rule is defined as the following:

$$F_k^D(i, j) = \begin{cases} Ca_k^D(i, j) & SDa_k^D(i, j) > SDb_k^D(i, j) \\ Cb_k^D(i, j) & SDa_k^D(i, j) < SDb_k^D(i, j) \\ (Ca_k^D(i, j) + Cb_k^D(i, j)) / 2 & \text{otherwise} \end{cases} \quad (9)$$

Where,  $SDa_k^D(i, j)$ 、 $SDb_k^D(i, j)$  represent the standard deviation of diagonal high-frequency sub-band coefficient in image A、B, respectively.

The rule of fusion based on the improved regional statistical characteristics can correspond to low-frequency and high-frequency coefficient selective highlighting the contrast of the source image. While the algorithm based

reconstruction of the structure tensor matrix, can achieve the purpose of enhanced clarity. The rule is proposed in this paper, not only reflected local contrast information, but can get more regional features and details.

#### IV. FUSION RESULTS AND DISCUSSION

In this section, we use two groups of multifocus images analyses experimental result of the algorithm is present in this paper. The basis wavelet fusion is bior2.4, and the images are decomposed at three layers, the size of local windows is  $3 \times 3$ . The experimental images are provided by [www.imagefusion.org](http://www.imagefusion.org).

##### A. Subjective evaluation of qualitative

The experiments are performed on two groups of images: clock images and Pepsi images. The multifocus images used for fusion are left focused and right focused as shown in Fig.2 and Fig.3.

The results of tradition algorithm (high frequency coefficients select via bigger absolute value and low frequency coefficients select via weighted average), the local gradient method, local variance method and proposed method are shown by (c)-(f) in Fig.2 and Fig.3, respectively.

From the results of four methods, the definition of results obtained by using traditional method is lower than others; the local methods get the lighter results; the result using the proposed method can improve the definition and contrast of images, and remove shadows around the edges of images.

##### B. Objective quantitative assessment

In order to objectively evaluate the performance of our algorithm, this article also adopted the mutual information (MI) [13] and edge information Q [14].

MI value reflects the fused image contains information content size in source image A and B. The higher MI value,

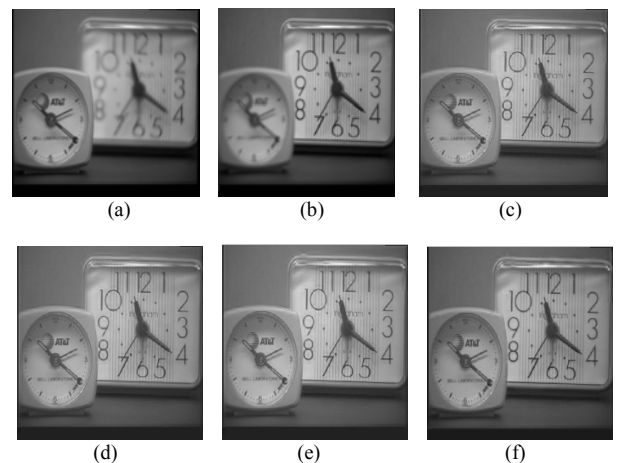


Fig. 2. Fusion results of the multi-focus Clock Image. (a) Source image A. (b) Source image B. (c) Traditional fusion. (d) Local gradient fusion. (e) Local variance fusion. (f) Fused image using the proposed method.

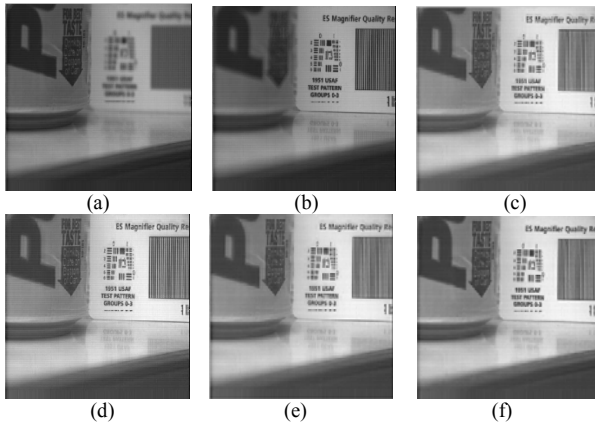


Fig. 3. Fusion results of the multi-focus Pepsi Image. (a) Source image A. (b) Source image B. (c) Traditional fusion. (d) Local gradient fusion. (e) Local variance fusion. (f) Fused image using the proposed method.

the fused image received the more information from source images, and get the better fusion results.

Q shows the fused image relative to edge information retention capacity of source image A and B. The higher value of Q indicates the edge information of source images to retain well in the fused images.

Table 1 gives the values of the measures MI and Q. We

TABLE I. MI AND Q RESULTS FOR TWO SETS OF IMAGE FUSION

Methods	Fig.2.clock images		Fig.3.pepsi images	
	MI	Q	MI	Q
Traditional method	6.3291	0.6161	6.6745	0.7184
Local Gradient	6.4176	0.6301	6.7684	0.7326
Local Variance	6.4121	0.6213	6.7332	0.7264
Paper	7.0555	0.6572	7.3324	0.7484

can see that the proposed algorithm has a better fusion performance, compared to the traditional fusion methods.

## V. CONCLUSION

A novel method based on the selection and reconstruction of the wavelet coefficients for fusing multifocus image is proposed. The method uses the regional standard deviation and mean of the high frequency subband to choose the low frequency subband coefficients, uses the regional standard deviation to select the diagonal high frequency sub-band coefficients, and adopts the structure tensor analysis to reconstruct the high-frequency sub-band coefficients in horizontal and vertical orientation. The experiments indicate that the quality of the presented algorithm is better than that of the traditional method.

## REFERENCES

- [1] K. Kazemi, H. A. Moghaddam. Fusion of Multifocus Images Using Discrete Multi-wavelet Transform. *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI2003)*. Tokyo, Japan. 2003
- [2] L. J. Chipman, T. M. Orr, L. N. Graham. Wavelets and image fusion. *Proc Int Conf on Image Processing*. LOS Alamitos USA: IEEE Computer Society, 1995, 248-251
- [3] S. G. Mallat. A Theory for Multiresolution Approximation and Wavelet Representation. *IEEE Trans. Pattern Anal. Machine Intell.*, 1989, 11(7): 674-693
- [4] P. J. Burt. The pyramid as a structure for efficient computation. In : *Multiresolution Image Processing and Analysis*, Rosenfield A.,Ed., Springer-Verlag, new york,1984
- [5] G. A. María, G. C. Raquel, et al. Fusion of Multispectral and Panchromatic Images Using Wavelet Transform. *Evaluation of crop classification accuracy*. Data fusion sessions within the 22<sup>nd</sup> EARSel Symposium, 2002
- [6] G. X. Liu, W. H. Yang. A wavelet—decomposition—based image fusion scheme and its performance evaluation. *Acta Automatica Sinica*, 2002, 28(6): 927-934
- [7] H. Li, B. S. Manjunath, S. K. Mitra. Multisensor Image Fusion Using Wavelet Transform. *Graphical models and image processing*, 1993, 57(3): 235-245
- [8] M. Santos, G. Pajares, M. Portlela, et al. A New Wavelets Image Fusion Strategy. *Pattern Recognition and Image Analysis*, 2003, 2652 :919-926
- [9] X. Y. Jiang, L. W. Zhou, Z. Y. Gao. Multispectral Image Fusion Using Wavelet Transform. *SPIE*, 1996, 2898: 35-42
- [10] R. Chao, K. Zhang, Y. Li. An image fusion algorithm using wavelet transform. *Acta Electronic Sinica*, 2004, 32(5): 750-753
- [11] Z. Guo, J. Yang. Wavelet Transform Image Fusion Based On Regional Variance. In: *MIPPR 2007 Remote sensing and GIS data processing and applications, and Innovative multi-spectral technology and applications*, 2007
- [12] T. Brox, et al. Nonlinear Structure Tensors. In: *Image and Vision Computing*. 2006, 24: 41-55
- [13] F. Maes, et al. Multimodality Image Registration By Maximization of Mutual Information. *IEEE Transactions on Medical Imaging*, 1997, 16 (2): 187-198
- [14] C. S. Xydeas, V. Petrovic. Objective Image Fusion Performance Measure. *Electronics Letters*, 2000, 36(4): 308-309

# Discrete Sliding Mode Control with Predictive Sliding function for Multivariable Systems

Houda Ben Mansour, khadija Dehri and Ahmed Said Nouri

Research Unit: Numerical Control of Industrial Processes

National Engineering School of Gabes, Street of Medenine, 6029 Gabes, Tunisia

University of Gabes

Email: houda.b.mansour@gmail.com

Email: Khadija.Dehri@gmail.com

Email: ahmedsaid.nouri@enig.rnu.tn

**Abstract**—This paper shows the development of a Discrete Sliding Mode Control with Predictive Sliding Function (SMC-PSF) for multivariable systems. This is an extension of our previous works synthesized in the case of single input single output. The multivariable (SMC-PSF) consists of two loops, the primary loop is a Model Predictive Control (MPC) and the secondary loop is a Sliding Mode Control (SMC). The MPC allows to the SMC an optimal sliding function. This type of scheme improves the performances of the SMC and the MPC controllers. Simulation results demonstrate that the (SMC-PSF) gives better performances, for multivariable systems, in terms of strong robustness to external disturbance and parameters variation, chattering elimination and fast convergence, in comparison with the SMC.

**Keywords**—Multivariable systems, Sliding Mode Control, Model Predictive Control, Predictive Sliding function, Chattering phenomenon.

## I. INTRODUCTION

The growth in the complexity of modern industrial systems make difficult the design of an exact mathematical model and the development of a suitable control. In fact, these systems are non linear, multivariable, also with external disturbances, parameter uncertainties and time delays. Therefore Sliding Mode Control (SMC) and Model based predictive control(MPC) are excellent candidates to utilize as a control law for these systems.[1].

For a large class of systems, the SMC is particularly interesting due to its ability to deal with non linearities, uncertainties, modeling errors and disturbances [4]. The main idea behind SMC is to synthesize a discontinuous control input to force the states trajectories to reach a specific surface called the sliding surface in finite time and to stay on it. However, in spite of the robustness of the sliding mode control, the chattering phenomenon, caused by the discontinuous term of the control law, is still the main problem of the SMC which consists in a sudden and rapid variation of the control signal leading to undesirable results [1].

Many approaches have been proposed to solve this problem such as high order sliding mode control [2], [3].

On the other hand, in recent years, model based predictive control(MPC) has received a lot of attention in the control theory and applications. It has been successfully implemented in many industrial applications, showing good performances. The basic idea of MPC is to calculate a sequence of future

control signals in such a way that it minimizes a multistage cost function defined over a prediction horizon. The index to be optimized is a difference between the predictive system output and predictive reference sequence over the prediction horizon plus a quadratic function measuring control effort [4], [5], [6].

Nevertheless the control law is model dependent, so a perfect model is required to guarantee the success of MPC control strategies. Because of the finite horizon, the stability and the robustness of the process is difficult to analyze and guarantee, especially when constraints are present [7], [8].

As a solution, we have proposed in [9], [10], [11], [12], [13], [14] a controller which combine the design of SMC and MPC for single input single output systems. This combination improves the performances of the two control laws and overcome most of their specific drawbacks.

In other works, we have proposed another combination, which, it consists on a sliding mode controller, where the optimal sliding function is allowed by a model predictive control block based on a specific objective [10], [15]. The main idea of this work is to extend our previous works, concerning the Sliding Mode Control with predictive sliding function (SMC-PSF), to multivariable systems.

The paper is organized as follows: Section II gives the synthesis of the classical discrete multivariable sliding mode control for multivariable systems. The synthesis of the multivariable sliding mode controller with predictive sliding function is presented in section III. In the following section the proposed controller is tested on a simulation example, and compared to SMC control. Finally, section V draws conclusions of the paper.

## II. THE CLASSICAL DISCRETE MULTIVARIABLE SLIDING MODE CONTROL

Consider a discrete multivariable system subjected to external disturbances and parameters variation, defined by [16]:

$$\begin{cases} x(k+1) = (A + \Delta A)x(k) + (B + \Delta B)(u(k) + v(k)) \\ y(k) = Hx(k) + Du(k) \end{cases} \quad (1)$$

where:

$x(k) \in \mathbb{R}^n$  is the state vector at the instant  $k$ ,  $u(k) \in \mathbb{R}^m$  is the input vector at the instant  $k$ ,  $y(k) \in \mathbb{R}^p$  is the output vector

at the instant  $k$ ,  $v(k) \in \mathbb{R}^m$  is the disturbance input vector at the instant  $k$ .

The matrices  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ ,  $H \in \mathbb{R}^{p \times n}$  and  $D \in \mathbb{R}^{p \times m}$  are the nominal model matrices.  $\Delta A \in \mathbb{R}^{n \times n}$  and  $\Delta B \in \mathbb{R}^{n \times m}$  are the parameter uncertainties matrices.

The system (1) can be presented by the following form:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + w(k) \\ y(k) = Hx(k) + Du(k) \end{cases} \quad (2)$$

with:

$$w(k) = \Delta Ax(k) + \Delta Bu(k) + (B + \Delta B)v(k) \quad (3)$$

where  $w(k) \in \mathbb{R}^n$ .

The sliding function is defined as [17]:

$$S(k) = Cx(k) = [s_1(k) \cdots s_m(k)]^T \quad (4)$$

where the dimension of the matrix  $C$  are  $(m, n)$ .

The sliding function vector is chosen in order to verify the following reaching law [18], [19]:

$$S(k+1) = \Phi S(k) - \begin{bmatrix} m_1 \text{sign}(s_1(k)) \\ m_2 \text{sign}(s_2(k)) \\ \vdots \\ m_m \text{sign}(s_m(k)) \end{bmatrix} \quad (5)$$

where  $\Phi$  is a diagonal matrix with  $(m, m)$  dimension and verifying  $0 \leq \Phi_{i,i} < 1$  and  $m_i > 0$  for  $i \in [1 \ m]$ . and sign is the signum function defined as :

$$\text{sign}(s_i(k)) = \begin{cases} -1 & \text{if } s_i(k) < 0 \\ +1 & \text{if } s_i(k) > 0 \end{cases} ; \quad i \in [1 \ m]$$

Thus, using equation (5), the control law ensuring the quasi-sliding mode is calculated as follows [20]:

$$u(k) = (CB)^{-1} \left( -CAx(k) + \Phi S(k) - \begin{bmatrix} m_1 \text{sign}(s_1(k)) \\ m_2 \text{sign}(s_2(k)) \\ \vdots \\ m_m \text{sign}(s_m(k)) \end{bmatrix} \right) \quad (6)$$

$(CB)$  is inversible.

### III. SYNTHESIS OF DISCRETE MULTIVARIABLE SLIDING MODE CONTROL WITH PREDICTIVE SLIDING FUNCTION

A block diagram of the SMC-PSF is shown in Figure1, where the primary loop is a Model Predictive Control (MPC) and the secondary loop is a Sliding Mode Control (SMC)[10], [15].

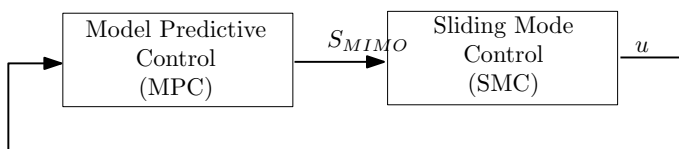


Fig. 1. SMC-PSF Controller bloc diagram.

The main purpose is to apply the discrete sliding mode control for multivariable systems which the sliding function is given optimally by the Model predictive control, based on a specific objective

We consider, now, the sliding mode control problem for multivariable system (1). The objective is to design sliding mode controller with predictive sliding function taking the reaching law (5). Define the vector  $\Delta U_{eq}(k+1)$  as:

$$\Delta U_{eq}(k+1) = \begin{bmatrix} \partial u_{eq}(k+1) \\ \partial u_{eq}(k+2) \\ \vdots \\ \partial u_{eq}(k+M) \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} u_{eq}(k+1) - u_{eq}(k) \\ u_{eq}(k+2) - u_{eq}(k+1) \\ \vdots \\ u_{eq}(k+M) - u_{eq}(k+M-1) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (7)$$

Or, the equivalent control vector is given by:

$$u_{eq}(k) = (CB)^{-1} [\Phi S(k) - CAx(k)]$$

So the vector  $\Delta U_{eq}(k+1)$  can be written as:

$$\Delta U_{eq}(k+1) = \begin{bmatrix} (CB)^{-1} \Phi [S(k+1) - S(k)] \\ (CB)^{-1} \Phi [S(k+2) - S(k+1)] \\ \vdots \\ (CB)^{-1} \Phi [S(k+N) - S(k+N-1)] \end{bmatrix} - \begin{bmatrix} (CB)^{-1} CA [x(k+1) - x(k)] \\ (CB)^{-1} CA [x(k+2) - x(k+1)] \\ \vdots \\ (CB)^{-1} CA [x(k+N) - x(k+N-1)] \end{bmatrix} \quad (8)$$

or, we have:

$$\left\{ \begin{array}{l} x(k+2) - x(k+1) = A(x(k+1) - x(k)) \\ \quad + B(u_{eq}(k+1) - u_{eq}(k)) \\ x(k+3) - x(k+2) = A^2(x(k+1) - x(k)) \\ \quad + AB(u_{eq}(k+1) - u_{eq}(k)) \\ \quad + B(u_{eq}(k+2) - u_{eq}(k+1)) \\ \vdots \\ x(k+M) - x(k+M-1) = A^{M-1}(x(k+1) - x(k)) \\ \quad + A^{M-2}B(u_{eq}(k+1) - u_{eq}(k)) \\ \quad + A^{M-3}B(u_{eq}(k+2) - u_{eq}(k+1)) \\ \quad + \dots \\ \quad + B(u_{eq}(k+M-1) - u_{eq}(k+M-2)) \\ \vdots \\ x(k+N) - x(k+N-1) = A^{N-1}(x(k+1) - x(k)) \\ \quad + A^{N-2}B(u_{eq}(k+1) - u_{eq}(k)) \\ \quad + A^{N-3}B(u_{eq}(k+2) - u_{eq}(k+1)) \\ \quad + \dots \\ \quad + A^{N-M-1}B(u_{eq}(k+M) - u_{eq}(k+M-1)) \end{array} \right. \quad \Psi_{MIMO} = \begin{pmatrix} 0 & \dots & \dots & \dots & \dots & 0 \\ TAB & 0 & \dots & \dots & \dots & 0 \\ TA^2B & TAB & 0 & \dots & \dots & 0 \\ TA^3B & TA^2B & TAB & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ TA^NB & TA^{N-1}B & \dots & TA^{N-M}B & \dots & 0 \end{pmatrix}$$

with  $T = (CB)^{-1}C$ .

Equation (10) can be written as:

$$(I + \Psi_{MIMO})\Delta U_{eq}(k+1) = \Delta S_{\Phi_{MIMO}}(k+1) - \Pi_{MIMO}[x(k+1) - x(k)] \quad (11)$$

$\Delta U_{eq}(k+1)$  can be given by:

$$\Delta U_{eq}(k+1) = (I + \Psi_{MIMO})^{-1}\Delta S_{\Phi}(k+1) - \Pi_{MIMO}[x(k+1) - x(k)] \quad (12)$$

Then the vector  $\Delta U_{eq}(k+1)$  can be presented by:

$$\Delta U_{eq}(k+1) = \begin{bmatrix} (CB)^{-1}\Phi[S(k+1) - S(k)] \\ (CB)^{-1}\Phi[S(k+2) - S(k+1)] \\ \vdots \\ (CB)^{-1}\Phi[S(k+N) - S(k+N-1)] \end{bmatrix} - \begin{bmatrix} (CB)^{-1}CA \\ (CB)^{-1}CA^2 \\ \vdots \\ (CB)^{-1}CA^N \end{bmatrix} (x(k+1) - x(k)) - \Psi_{MIMO}\Delta U_{eq}(k+1) \quad (9)$$

So, the equation (9) can be written as:

$$\Delta U_{eq}(k+1) = \Delta S_{\Phi_{MIMO}}(k+1) - \Pi_{MIMO}[x(k+1) - x(k)] - \Psi_{MIMO}\Delta U_{eq}(k+1) \quad (10)$$

with:

$$\Delta S_{\Phi_{MIMO}} = (k+1) \begin{bmatrix} (CB)^{-1}\Phi[S(k+1) - S(k)] \\ (CB)^{-1}\Phi[S(k+2) - S(k+1)] \\ \vdots \\ (CB)^{-1}\Phi[S(k+N) - S(k+N-1)] \end{bmatrix} - \Pi_{MIMO} = \begin{bmatrix} (CB)^{-1}CA \\ (CB)^{-1}CA^2 \\ \vdots \\ (CB)^{-1}CA^N \end{bmatrix}$$

and

So:

$$\Delta U_{eq}(k+1) = K_{MIMO}\Delta S_{\Phi}(k+1) + L_{MIMO}[x(k+1) - x(k)] \quad (13)$$

with:

$$\begin{cases} K_{MIMO} = (I + \Psi_{MIMO})^{-1} \\ L_{MIMO} = -(I + \Psi_{MIMO})^{-1}\Pi_{MIMO} \end{cases} \quad (14)$$

To find the predictive function vector, the following corresponding optimization cost function is defined:

$$j_{SMC-PSF}(k) = \sum_{j=1}^N q_j [\delta S_{\Phi}(k+j) - \delta S_r(k+j)]^2 + \sum_{l=1}^M g_l [\delta u_{eq}(k+l-1)]^2 \quad (15)$$

where  $\delta S_r(k+j)$  the increment of the sliding mode references trajectories vector,  $\delta S_{\Phi}(k+j)$  is the increment of the predictive sliding function vector, multiplied by the term  $(CB)^{-1}\Phi$ ,  $q_j$  and  $g_l$  are weight coefficients.

In order to simplify the synthesis of the controller, we consider  $q_j = q$  and  $g_l = g$ . So, the following corresponding optimization cost function (15) is written by:

$$j_{SMC-PSF}(k) = \sum_{j=1}^N q [\delta S_{\Phi}(k+j) - \delta S_r(k+j)]^2 + \sum_{l=1}^M g [\delta u_{eq}(k+l-1)]^2 \quad (16)$$

Rewrite equation (16) in vector form:

$$J_{SMC-PSF}(k) = \|\Delta S_{\Phi_{MIMO}}(k) - \Delta S_{r\_MIMO}(k)\|_Q^2 + \|\Delta U_{eq}(k)\|_G^2 \quad (17)$$

where

$$\Delta S_{r\_MIMO}(k+1) = [\delta S_r(k+1), \delta S_r(k+2), \dots, \delta S_r(k+N)]^T$$

$$G = [gI_m, gI_m, \dots, gI_m]$$

$$Q = [qI_m, qI_m, \dots, qI_m]$$

Since the control objective is to keep states on the sliding surface, the desired sliding mode reference trajectory vector is approximated by the predictive sliding function vector and should verify:

$$S_r(k) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(m \times 1)}, \text{ so } \Delta S_{r\_MIMO} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(mN \times 1)}$$

So, the optimization cost function can be written as:

$$J_{SMC-PSF}(k) = \|\Delta S_{\Phi_{MIMO}}(k)\|_Q^2 + \|\Delta U_{eq}(k)\|_G^2 \quad (18)$$

Rewrite equation(13),  $J_{SMC-PSF}(k)$  can be written as:

$$J_{SMC-PSF}(k) = \Delta S_{\Phi_{MIMO}}(k)^T Q \Delta S_{\Phi_{MIMO}}(k) + [K_{MIMO} \Delta S_{\Phi_{MIMO}}(k) + L_{MIMO}(x(k) - x(k-1))]^T G [K_{MIMO} \Delta S_{\Phi_{MIMO}}(k) + L_{MIMO}(x(k) - x(k-1))] \quad (19)$$

The optimal sequence of the increment of the predictive sliding function vector is obtained by minimizing the cost function  $J_{SMC-PSF}$ :

$$\frac{\partial J_{SMC-PSF}(k)}{\partial \Delta S_{\Phi_{MIMO}}(k)} = 0 \quad (20)$$

So, the increment of the predictive sliding function vector can be calculated as:

$$\Delta S_{\Phi_{MIMO}}(k) = -(GK_{MIMO}^T K_{MIMO} + Q)^{-1} GK_{MIMO}^T L_{MIMO}(x(k) - x(k-1)) \quad (21)$$

We suppose that  $\delta S_{\Phi}(k)$  is the vector of the  $m$  first elements of the vector  $\Delta S_{\Phi_{MIMO}}$ , so, the predictive sliding function vector  $S_{\Phi}(k)$  is given as:

$$S_{\Phi}(k) = S_{\Phi}(k-1) + \delta S_{\Phi}(k) \quad (22)$$

with:  $S_{\Phi}(k) = (CB)^{-1} \Phi S(k)$

So, the control law  $u_{eq}(k)$  is given by the equation:

$$u_{eq}(k) = (CB)^{-1} [\Phi S(k) - CAx(k)] \quad (23)$$

Then:

$$u_{eq}(k) = S_{\Phi}(k) - (CB)^{-1} CAx(k) \quad (24)$$

Or, we have:

$$u_{dis}(k) = -(CB)^{-1} \begin{bmatrix} m_1 \text{sign}(s_1(k)) \\ m_2 \text{sign}(s_2(k)) \\ \vdots \\ m_m \text{sign}(s_m(k)) \end{bmatrix} \quad (25)$$

So:

$$u(k) = u_{eq}(k) + u_{dis}(k) \quad (26)$$

#### IV. SIMULATION RESULTS

To evaluate the robustness of the control laws (equations 6 and 26) in presence of constant or periodic disturbances and parameters uncertainties, we consider a discrete multivariable process described by the following equation:

$$x(k+1) = (A + \Delta A)x(k) + (B + \Delta B)(u(k) + v(k))$$

where:

$$A = \begin{bmatrix} 0 & 1 \\ 0.24 & 0.2 \end{bmatrix}; \quad B = \begin{bmatrix} 1.5 & 0 \\ 0 & 1 \end{bmatrix}$$

The retained synthesis parameters are:

$$C = \begin{bmatrix} 0.6667 & 0 \\ 0 & 1 \end{bmatrix}$$

and  $m_1 = 0.01$ ,  $m_2 = 0.01$ ,  $\Phi = [0.01 \quad 0; 0 \quad 0.01]$ ,  $N = 10$ ,  $M = 5$ ,  $H_p = 0.001I(N, N)$ , and  $G = 0.001I(N, N)$

The sliding functions vector is given by:

$$S(k) = Cx(k) = C = \begin{bmatrix} 0.6667 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} = \begin{bmatrix} s_1(k) \\ s_2(k) \end{bmatrix}$$

##### A. Case of constant disturbances

The results presented in this section are obtained with the presence of constant disturbances which are given by:

$$v(k) = \begin{bmatrix} 0.15 \\ 0.2 \end{bmatrix}, \quad \forall k \geq 100$$

The parameters variation are given by:

$$\Delta A = 0.1 \begin{bmatrix} 5 \sin(-\frac{2k\pi}{10}) & 6 \sin(-\frac{2k\pi}{10}) \\ 3 \sin(-\frac{2k\pi}{10}) & 3 \sin(-\frac{2k\pi}{10}) \end{bmatrix}$$

$$\Delta B = 0.1 \begin{bmatrix} 2 \sin(-\frac{2k\pi}{10}) & 5 \sin(-\frac{2k\pi}{10}) \\ 3 \sin(-\frac{2k\pi}{10}) & 5 \sin(-\frac{2k\pi}{10}) \end{bmatrix}, \forall k \geq 300$$

The evolution of the states  $x_1(k)$  and  $x_2(k)$ , the control inputs  $u_1(k)$  and  $u_2(k)$  and the sliding mode functions  $s_1(k)$  and  $s_2(k)$  with SMC-PSF and SMC are given, respectively, in figures 2 to 7.

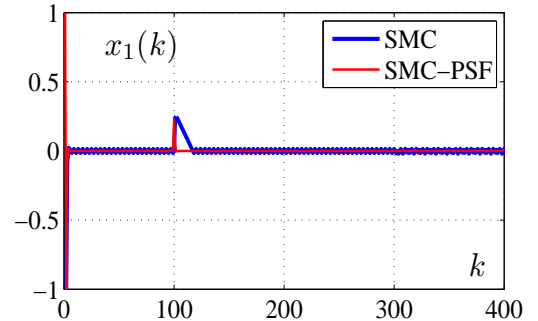
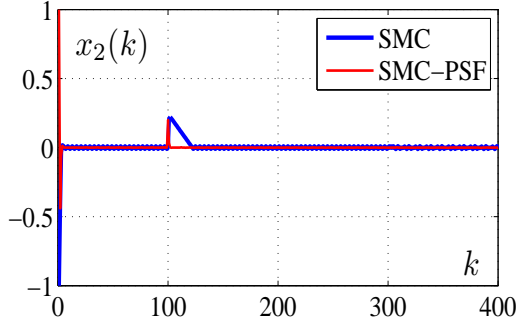
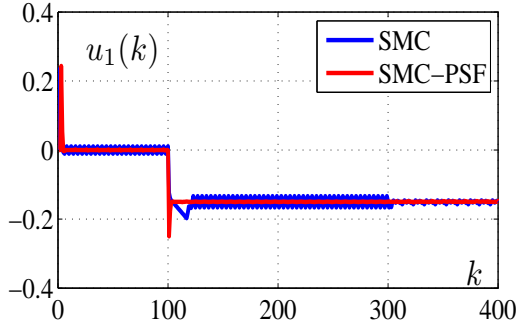
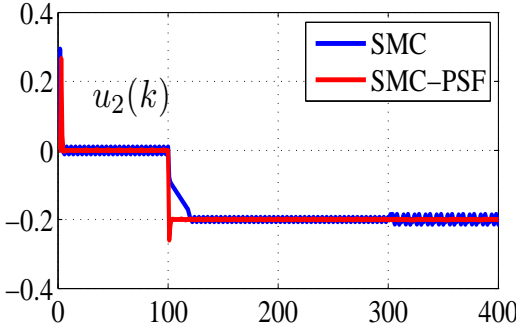
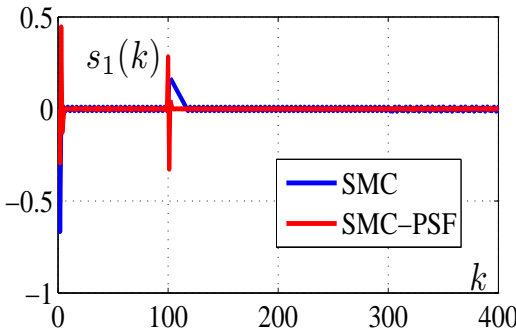
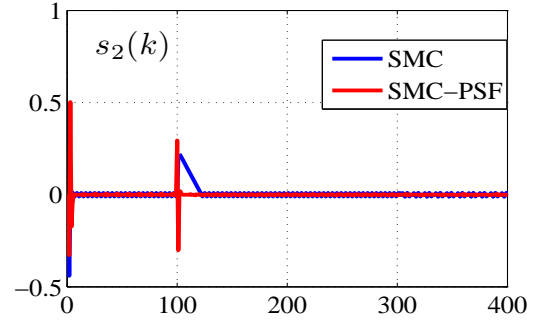


Fig. 2. Evolution of the state  $x_1(k)$ .



Fig. 3. Evolution of the state  $x_2(k)$ .Fig. 4. Evolution of the control signal  $u_1(k)$ .Fig. 5. Evolution of the control signal  $u_2(k)$ .Fig. 6. Evolution of the sliding function  $s_1(k)$ .Fig. 7. Evolution of the sliding function  $s_2(k)$ .

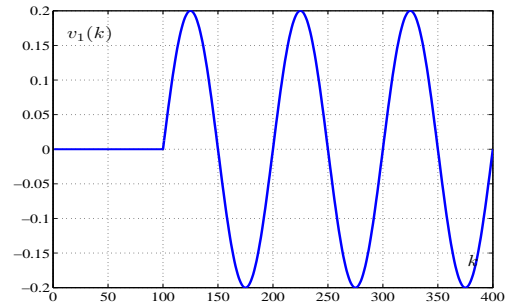
It can be seen that the performances of the SMC-PSF are better than the other studied controller, not only, for rejecting constant disturbances, but also, for eliminating chattering. In fact, without disturbances and parameters uncertainties, the results of SMC and SMC-PSF are comparable. But, in presence of constant disturbances ( $k \geq 100$ ), we find that the proposed control law ensure good performances in term of rejection of external disturbances and fast convergence. When we add parameters uncertainties, at the instant ( $k \geq 300$ ), the oscillation encountered, in the case of classical SMC, are reduced.

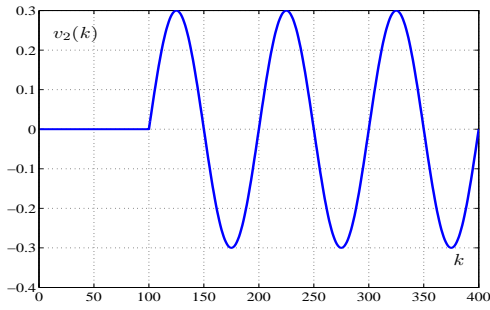
#### B. Case of periodic disturbances

The results presented in this section are obtained with the presence of disturbances, whose evolutions are given in figures 8 and 9, and with the following parameters variation:

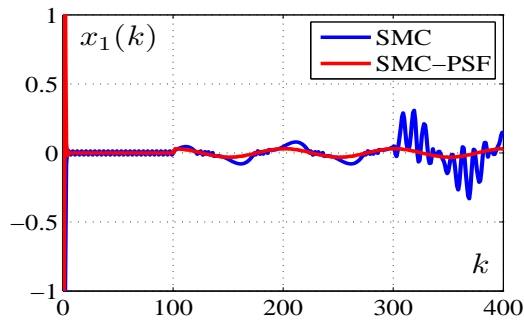
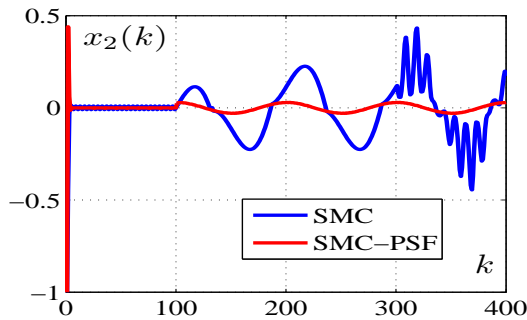
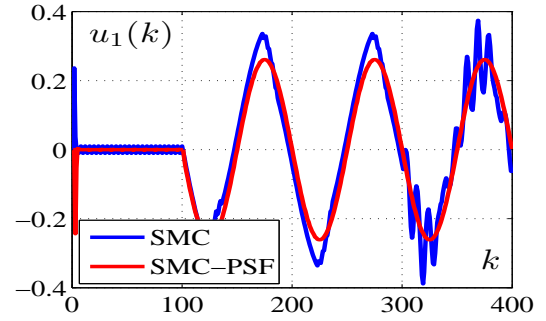
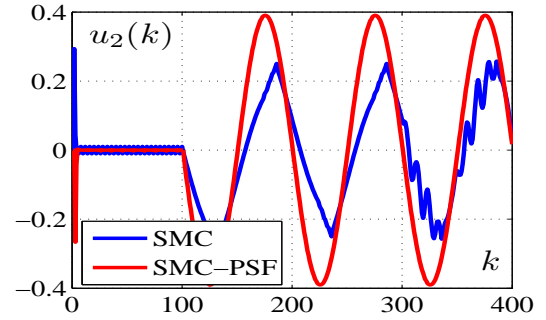
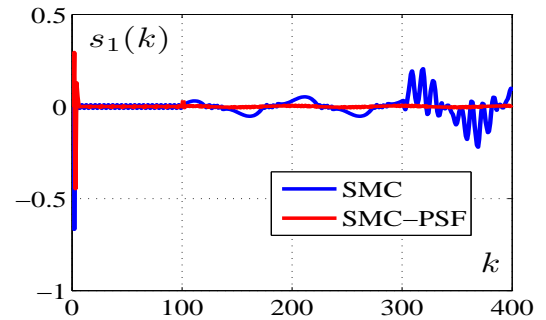
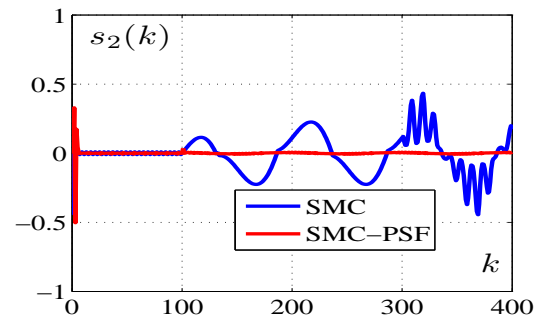
$$\Delta A = 0.1 \begin{bmatrix} 5 \sin(-\frac{2k\pi}{10}) & 6 \sin(-\frac{2k\pi}{10}) \\ 3 \sin(-\frac{2k\pi}{10}) & 3 \sin(-\frac{2k\pi}{10}) \end{bmatrix}$$

$$\Delta B = 0.1 \begin{bmatrix} 2 \sin(-\frac{2k\pi}{10}) & 5 \sin(-\frac{2k\pi}{10}) \\ 3 \sin(-\frac{2k\pi}{10}) & 5 \sin(-\frac{2k\pi}{10}) \end{bmatrix}, \forall k \geq 300$$

Fig. 8. Evolution of the disturbance  $v_1(k)$ .

Fig. 9. Evolution of the disturbance  $v_2(k)$ .

The evolution of the states  $x_1(k)$  and  $x_2(k)$ , the sliding mode functions  $s_1(k)$  and  $s_2(k)$  and the control inputs  $u_1(k)$  and  $u_2(k)$ , with SMC-PSF and SMC are given, respectively, in figures 10 to 15.

Fig. 10. Evolution of the state  $x_1(k)$ .Fig. 11. Evolution of the state  $x_2(k)$ .Fig. 12. Evolution of the control signal  $u_1(k)$ .Fig. 13. Evolution of the control signal  $u_2(k)$ .Fig. 14. Evolution of the sliding function  $s_1(k)$ .Fig. 15. Evolution of the sliding function  $s_2(k)$ .

A comparison between the SMC-PSF and SMC, in the case of multivariable systems, reveals that the use of the new control strategy SMC-PSF reduces the chattering problem effectively ( $k \geq 300$ ).

Furthermore, the results obtained prove the capability of the proposed control law to reduce periodic disturbances  $k \geq 100$  and parameter uncertainties  $k \geq 300$ .

## V. CONCLUSION

In this paper, a sliding mode controller with predictive sliding function for multivariable systems was proposed. This controller combines the design technique of the SMC and MPC.

The method was tested on a multivariable system, and compared to the results given by the SMC controller. It is shown that mixing both control techniques, for multivariable systems, gives new controller with better robustness properties in rejecting disturbances, hard parameter variations and in eliminating the chattering problem.

## ACKNOWLEDGMENT

"This work was supported by the Ministry of the Higher Education and Scientific Research in Tunisia"

## REFERENCES

- [1] P. Lopez and A.S. Nouri. *Thorie lmentaire et pratique de la commande par les rgimes glissants*. Mathmatiques et application, 2006.
- [2] M. Mihoub, A.S. Nouri, and R. Ben Abdenmour. Real-time application of discrete second order sliding mode control to a chemical reactor. *Control Engeneering Practice*, 17:1089–1095, 2009.
- [3] A. Cavallo and C. Natale. High-order sliding control of mechanical systems: theory and experiments. *Automatica*, 12:1139–1149, 2004.
- [4] D. Clarke and R. Scattolini. Constrained receding horizon predictive control. *IEEE Transaction automatic control*, 23 N2:347–354, 1991.
- [5] E. F. Camacho and C. Bordón. *Model Predictive Control*. Springer, 2004.
- [6] R. J. Culi and C. Bordón. Iterative nonlinear model predictive control, stability robustness and applications. *Chemical Engineering Practice*, 16:1023–1034, 2008.
- [7] J. M. Maciejowski. Predictive control with constraints. In *Pretice Hall, Harlow*, 2011.
- [8] D.W. Clarke, C.Mohtadi, and P.S.Tuffs. Generalised predictive control i: The basic algorithm. *Automatica*, 23:137–148, 1987.
- [9] H. BenMansour, K. Dehri, and A. S. Nouri. New predictive sliding mode control for non minimum phase systems. *International Journal of Computer Applications IJCA*, 70 - N. 11:1–8, 2013.
- [10] H. BenMansour, K. Dehri, and A. S. Nouri. New discrete sliding mode controller with predictive sliding function. *International Review of Automatic Control*, 6 N 4:1–10, 2013.
- [11] H. BenMansour and A.S. Nouri. Discrete predictive sliding mode control of uncertain systems. In *Proccedings of the 9th International Multi-Conference on System, Signals and Devices Germany*, 2012.
- [12] H. BenMansour and A. S. Nouri. Predictive sliding mode control for perturbed discrete delay time systems: Robustness analysis. In *International Conference on Electrical Engineering and Software Application ICEESA 2013*, 2013.
- [13] H. BenMansour and A. S. Nouri. Discrete predictive sliding mode control of uncertain systems. In *Proceeding of the 9th International Multi-Conference on Systems, Signals and devices*, 2012.
- [14] N. Abdennabi, H. Ben Mansour, and A. S. Nouri. A new sliding function for discrete predictive sliding mode control of time delay systems. *International Journal of Computer Applications (IJCA)*, 10(4):288–295, 2013.
- [15] H. BenMansour, K. Dehri, and A. S. Nouri. Comparison between predictive sliding mode control and sliding mode control with predictive sliding function. In *Second International Conference on Electrical Engineering and Control Applications ICEECA, Constantine, Algeria.*, November, 18-20, 2014.
- [16] R. BenAbdenmour, P. Borne, M. Ksouri, and F. Msahli. *Identification et commande numrique des procds industriels*. Technip, Paris, France, 2001.
- [17] K. Dehri. Sur le rejet des perturbations harmoniques par les rgimes glissant. Master's thesis, Ecole Nationale d'Ingnieurs de Gabs (ENIG), 2013.
- [18] J. L. Chang. Discrete sliding-mode control of mimo linear systems. *Asian Journal of Control*, 2:10–15, 2002.
- [19] G. Monsees. Discrete-time sliding mode control. Master's thesis, PhD thesis, Delft University of Technology, 2002.
- [20] K. Dehri, Ltaief M., Nouri A. S., and Ben Abdenmour R. Rejection of periodic disturbances with unknown frequency for multivariable systems. *International Journal on Sciences and Techniques of Automatic control and computer engineering*, 5, n1:1458–1471, 2011.

# Benign or malignant lesion classification in mammography images using the adaptive orthogonal transformation and the coefficients of the correlation matrix

Khalid El Fahssi, Abdelali Elmoufidi, Abdenbi Abenaou, Said Jai-Andaloussi, Abderrahim Sekkaki

**Abstract**—Segmentation and classification of breast masses in mammography play a crucial role in Computer Aided Diagnosis system (CAD). In this paper we propose an approach based on the theory of adaptive orthogonal transformation that will calculate the informative characteristics of regions of interest of mammography images and classification by comparison of the similarity between the vectors of the characteristics of regions of interest by use of the coefficient of matrix of correlation. The result obtained by this calculation method allows the increase the efficiency of diagnosis. To illustrate the effectiveness of the method we present the results of experiments carried out on the basis of images MIAS mammograms.

**Keywords**—segmentation; mammography; active contours; classification; orthogonal; correlation

## I. INTRODUCTION

Breast cancer is the most common evil in women worldwide. It is a leading cause of female mortality [1, 2,3] since every woman has a one in eight chance of developing breast cancer during her lifetime. [4]. The Prevention of the disease is very difficult because the risk factors are either poorly known or not suggestible. Scientific studies have provided a better understanding of development of cancer but it is still not possible to know why one person develops breast cancer. It should be noted that only 5 to 10% of breast cancers are hereditary related to the transmission of deleterious genes which are the most frequently incriminated are BRCA1 and BRCA2 (breast Cancer acronyms) associated with a predisposition of the disease [5]. Mammography is the most effective imaging technique to detect tumors at an early stage [2,6] and currently is the principal investigation in breast cancer screening. However, radiologists are always obliged to examine very finely the image to a better diagnosis hence the importance of Computer Aided Diagnosis system (CAD) developed over the last twenty years[7]. Generally systems (CAD) based on a series of approaches including pretreatment steps, segmentation, extraction of parameters of classification and finally the interpretation and classification of suspected abnormalities [7,8,9,10]. Generally the first indicators of malignancy parameters are connected to the mass density, size,

shape and edges, malignant masses often infiltrate the fabric of linear son purposes that extend outward from the center of the mass. From the pathology of malignant masses, the shape can be used to discriminate between malignant mass and benign masses. Approximately 80 to 85% of breast cancers are diagnosed localized tumor's appearance on mammography[4]

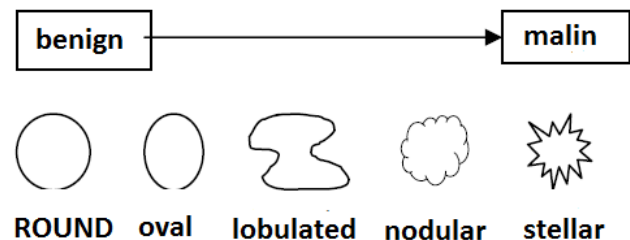


Fig. 1. The morphological spectrum of mammographic masses.

For automatic classification of tumors several techniques are used such as that based on artificial neural networks [11,12], one based on fuzzy logic [14], and other based on support vector machines [13]. In this paper we propose a new technique based on adaptive orthogonal transformation and the coefficient of correlation matrix.

This paper is organized as follows: Section II.A describes the database used for evaluation, section II.B we give our method of pretreatment, section II.C we give a result of our method of segmentation, section II.D present the method for classification and detection of the lesion. The result is presented in section III and we end with a conclusion and perspective in section IV.

## II. MATERIALS AND METHOD

### A. Database

The mini-MIAS [20] database contains a total of data 322 MLO view mammography images. This database is divided into categories such margin: speculated, circumscribed or poorly defined. The images have a resolution of 1024 1024 pixels. From this data set, a total of 111 lesions was selected. These include 60 benign and 51 malignant masses. An example of a series of the image is given by Figure 1.

Khalid El Fahssi, Abdelali Elmoufidi, Abdenbi Abenaou, Said Jai-Andaloussi, Abderrahim Sekkaki are with LIAD Labs, Casablanca, Kingdom of Morocco.; elifahssi@etude.univcasa.ma

Khalid El Fahssi, El Moufidi, Said Jai-Andaloussi and Abderrahim Sekkaki are with Faculty of science Ain-chok, Casablanca, Kingdom of Morocco.

Abdenbi Abenaou are with National School of Applied Sciences, Agadir, Kingdom of Morocco.

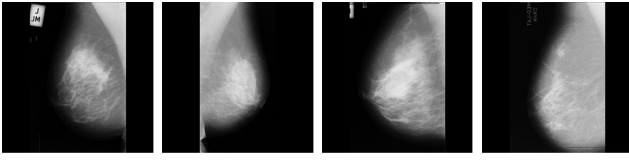


Fig. 2. Example of mammography study.

### B. preprocessing of mammography Image

This step is crucial in our system for the reduction of false positives. Indeed, in these images, the breast covers only 30% of the entire image [15]. Moreover, the digitization of mammography images may result in noise in the resulting mammograms. So in mammography images several types of noise and artifacts are present [16]. Thus for improving the quality of mammography we used an image preprocessing method based on shrinkwrap function [17], this function amplifies areas of high intensity and segments them using a front. The front is initialized on the convex hull (for speed) and erodes the map until it has converged on the edge of the areas to keep, maintaining edge geometry.

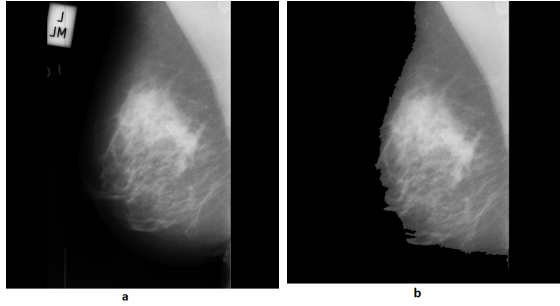


Fig. 3. Mammogram preprocessing step : (a) Original mammogram, (b) Artifact suppressed

### C. our method of segmentation

For segmentation the mammography images we have used a combined solution of the two approaches, one based on levels set theory and the other based on the principle of the minimization of the energy of active contours[18]. The elaborated algorithm which is an interactive approach permits to segment effectively the images from its coarse resolution to its refinement. The results of the segmentation are illustrated from the database of MIAS. The figure represents the result of our method of segmentation.

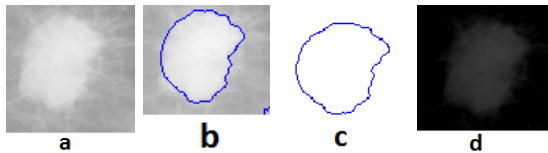


Fig. 4. Segmentation of mammograms mdb184 : (a) Original mammogram, (b)ROI detected ,(c) The contour of ROI, (d) ROI extracted

### D. Method proposed for classification of lesion

The first part of the approach proposed in this article has allowed the segmentation and selection of regions of

interest of mammography images of various types of disease. The classification of mammograms depending on the type of disease images requires the extraction of the most relevant features in these regions of interest [19],[21].

the calculation of characteristics vectors informative regions is ensured by the use of an adaptable orthogonal transformation [22]for each class regions (type of disease). The spectrum of informative characteristics of a given area (class) obtained decreases rapidly. In other words, the analysis of the spectra obtained allows a considerable difference between the classes of regions. This property will subsequently develop a simple decision rule which ensured a high certainty of classification.

The principle of the proposed method is to synthesize an adaptable operator orthogonal transformation to generate functions of bases parameterizable. Using these transformations [22],[23] is favored by the ability to adapt to the shape of their basic functions depending on the nature of the standard vector formed by mammograms each class images. In other words, for each type of disease a system of basis functions are associated parameterizable for showing his images. In addition, these functions of base parameterizable meet the criteria for completeness of the system ensuring the transformation of vectors without loss of information content. The system of basis functions formed is expressed as a factorisable orthonormal matrix operator, which therefore allows a transformation with a fast calculation algorithm:

$$Y = \frac{1}{N} H X \quad (15)$$

where:

-  $X = [x_1, x_2, \dots, x_N]^T$  is the vector representing the region of interest in the segmented image mammography given initial vector transform (of size  $N = 2^n$ ).

-  $Y = [y_1, y_2, \dots, y_N]^T$  is the vector of informational characteristics calculated by the spectral operator orthogonal  $H$  of dimension  $N \times N$ .

Factorization of Good [24] showed a possibility of representing the matrix operator  $H$  as product  $G_i$  (16) Sparse matrix with a higher proportion of zero which has allowed the construction the quick transformation algorithms of Fourier , Haar and Walsh. The matrices  $G_i (i = 1, \dots, n)$  are constructed by blocks of matrices  $V_{i,j}$  of minimum dimension that is called spectral nuclei [22]:

$$G_i = \begin{bmatrix} \alpha_{i1} & 0 & \dots & 0 & \gamma_{i1} & 0 & \dots & 0 \\ \beta_{i1} & 0 & \dots & 0 & \delta_{i1} & 0 & \dots & 0 \\ 0 & \alpha_{i2} & 0 & \dots & 0 & \gamma_{i2} & \dots & 0 \\ 0 & \beta_{i2} & 0 & \dots & 0 & \delta_{i2} & \dots & 0 \\ & & \ddots & & & & \ddots & \\ 0 & \dots & 0 & \alpha_{iN/2} & 0 & \dots & 0 & \gamma_{iN/2} \\ 0 & \dots & 0 & \beta_{iN/2} & 0 & \dots & 0 & \delta_{iN/2} \end{bmatrix} \quad (16)$$

With

$$V_{i,j} = \begin{bmatrix} \alpha_{ij} & \dots & \gamma_{ij} \\ \beta_{ij} & \dots & \delta_{ij} \end{bmatrix} = \begin{bmatrix} \cos(\varphi_{i,j}) & \dots & w_{i,j} \sin(\varphi_{i,j}) \\ \sin(\varphi_{i,j}) & \dots & -w_{i,j} \cos(\varphi_{i,j}) \end{bmatrix},$$

$$w_{i,j} = \exp(j\theta_{i,j}), \quad \varphi \in [0, 2\pi], \quad \theta \in [0, 2\pi].$$

Hence equation (15) can be written as follows:

$$Y = \frac{1}{N} HX = \frac{1}{N} G_1 G_2 \dots G_N X = \frac{1}{N} \Pi_{i=1}^N G_i \quad (17)$$

By defining the parameters  $\varphi_{i,j}$  and  $\theta_{i,j}$  can train operators orthogonal of transformations with basic functions complex, and  $\theta_{i,j} = 0$  operators with real functions. Adapting Operator  $H$  in (15) is provided by the condition:

$$\frac{1}{N} H_a Z_{sd} = Y_c = [Y_{c,1}, 0, 0, \dots, 0], Y_{c,1} \neq 0 \quad (18)$$

where:

- $Y_c$  is the target that builds the basis for adjusting the vector operator  $H_a$ .
  - $Z_{sd}$  represents the calculated by means of the estimates of the statistical characteristics of images of a given class standard vector.
  - $H_a$  is adaptable to synthesize operator.
- Synthesis adaptable operator  $H_a$  based standard  $Z_{sd}$  (for a given class) is to calculate the angular parameters  $\varphi_{i,j}$  matrices  $G_i$  according to condition (18). The procedure for calculation of the parameters is based on an iterative algorithm to calculate the target vector  $Y_c$  by step according to the equation:

$$Y_i = G_i Y_{(i-1)} \quad (19)$$

The first iteration ( $Y_0 = Z_{sd}$ )

$$\begin{bmatrix} [V_{1,1}] \\ [V_{1,2}] \\ \vdots \\ [V_{1,N/2}] \end{bmatrix} G_1 \begin{bmatrix} Y_0 = Y_1 \end{bmatrix} = \begin{bmatrix} Y_1 = Y_1 \end{bmatrix} \quad \frac{N}{2^1}$$

**number of non-zero coefficients**

The second iteration

$$\begin{bmatrix} [V_{2,1}] \\ [V_{2,2}] \\ \vdots \\ [V_{2,N/2}] \end{bmatrix} G_2 \begin{bmatrix} Y_1 = Y_2 \end{bmatrix} = \begin{bmatrix} Y_2 = Y_2 \end{bmatrix} \quad \frac{N}{2^2}$$

The nth iteration

$$\begin{bmatrix} [V_{n,1}] \\ [V_{n,2}] \\ \vdots \\ [V_{n,N/2}] \end{bmatrix} G_n \begin{bmatrix} Y_{n-1} = Y_n = Y_c \end{bmatrix} = \begin{bmatrix} Y_n = Y_n \end{bmatrix} \quad \frac{N}{2^n}$$

The final calculation of the vector  $Y_c$  allows obtaining adaptable operator  $H_a$ .

For the classification of lesions, we dispose two sets of feature vectors for each type of tumor (benign or malignant). The first was used for the calculation of the standard  $Z_{sd,i}$  (tumor i) for the synthesis of the operator, while the second set used to form the spectral standard  $Y_{sd,i}$ . The latter is obtained by projecting the feature vector of the second set in  $H_a$  adaptable bases.

To make the decision and classify each tumor, we calculate each  $Y_i$  spectrum in each base  $H_{a,i}$  of the region of interest, then we will of a rule of decision based on the calculation of the coefficient of matrix correlation between the standard spectrum of each class and the projection of  $Y_i$  spectrum of the region of interest in the base  $H_{a,i}$  of the same class. Correlation coefficient is calculated by the following equation:



$$r_p = \frac{\hat{\sigma}_{XY}}{\hat{\sigma}_X \hat{\sigma}_Y}$$

with

$$\hat{\sigma}_{XY} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})$$

$$\hat{\sigma}_X = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}, \hat{\sigma}_Y = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$

which are respectively estimators of covariance, standard deviations and expectations of the variables X and Y. Then two spectra are not linearly correlated if  $r_p$  is zero. The two spectrum are better correlated if  $r_p$  is near to 1 or -1. then according to the condition of linear correlation, the specter of the region of interest belongs to a class of which  $\max(|r_p|)$  is near to 1.

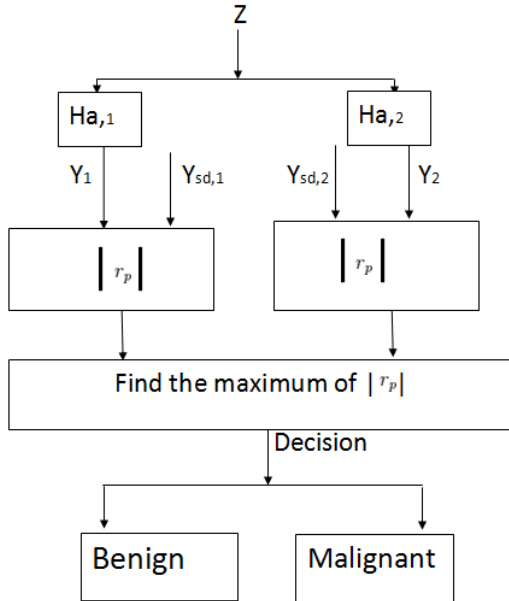


Figure 5. Classification procedure

### III. RESULTS AND DISCUSSION

This section presents the results of experiments performed on 111 selected mammography database mini-MIAS images. To evaluate the effectiveness of the proposed method, we used the information provided by the Mammographic Image Analysis Society database (MIAS) including the class of anomalous images and coordinates of their centers of regions of interest. To test our method, the algorithm is applied to each image containing a mass lesion. In the classification of ROI to Benign or Malignant, a positive case means correct classification of ROI to benign or malignant while a negative

case means incorrect classification of ROI as such a type. The definitions of the fractions are as below:

- True Positive (TP) means breast classified as Malignant that proved to be Malignant.
- False Positive (FP) means breast classified as Malignant that proved to be Benign.
- False Negative (FN) means breast classified as Benign that proved to be Malignant.
- True Negative (TN) means breast classified as Benign that proved to be Benign.

We have tested the performance of our method by calculating and analysis of accuracy, sensitivity and specificity for malignant and benign classification. These are defined and calculated as follows[25]:

Sensitivity: number of correct classified Benign mass/number of total Benign mass :

$$\text{Sensitivity} = (\text{TP} / (\text{TP} + \text{FN})) * 100$$

Specificity: number of correct classified Malignant mass/number of total Malignant mass:

$$\text{Specificity} = (\text{TN} / (\text{TN} + \text{FP})) * 100$$

Our method has been proven effective for the classification masses to benign or malignant on a mammogram with a sensitivity equal 93.78% and Specificity equal 94.54%.

### IV. CONCLUSION AND PERSPECTIVE

In this work we presented an approach for classification of lesion to benign or malignant in digital mammograms. The experimentation gives a percentage of 93.78% of a sensitivity and a percentage of 94.54% of Specificity for all cases studied. The results of the algorithm can contribute to solving the main problem in mammography image processing such as diagnostic and classification. The Efficiency of the proposed method confirms the possibility of its use in improving the computer-aided diagnosis.

### REFERENCES

- [1] Al Mutaz M. Abdalla, Safaai Dress, Nazar Zaki , Detection of Masses in Digital Mammogram Using Second Order Statistics and Artificial Neural Network, International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 3, June 2011, pp. 176-186.
- [2] Liyakathunisa & C.N. Ravi Kumar, A Novel and Efficient Lifting Scheme based Super Resolution Reconstruction for Early Detection of Cancer in Low Resolution Mammogram Images ,International Journal of Biometrics and Bioinformatics (IJBB), Volume (5) : Issue (2) : 2011 , pp 53-75.
- [3] N.Ben Hamad, N. Benromdhane, K. Taouil, M.S. Bouhlel, Reduction the False Positives in Assistance Systems Diagnosis in the Breast Cancer, SETIT 2007 4th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications, Tunisia ,2007.
- [4] L. M. Bruce & al, 1999 "Classifying Mammographic Mass Shapes Using the Wavelet Transform Modulus Maxima Method" IEEE Transactions On Medical Imaging Vol 18 No. 12.
- [5] Gulsrud, Thor. O & al, 1996 "Optimal filter for detection of stellate lesions and circumscribed masses in mammograms", Proc. SPIE vol 2727, pp 430-440, Visual communications and Image Processing

- [6] Songyang Yu and Ling Guan, A Cad System for the automatic detection of clustered microcalcifications in digitized mammogram films, *iee on medical imaging*, vol. 19, n. 2, february 2000, pp. 115-126
- [7] Rangaraj M.Rangayyan, Fabio J.Ayres, J.E.Leo Desautels, A review of computer-aided diagnosis of breast cancer: toward the detection of subtle signs, *sciencesdirect , journal of the franklin institute* 344 (2007) 312348
- [8] J. B. Jona, N. Nagaveni A Hybrid Swarm Optimization approach for Feature set reduction in Digital Mammograms, *WSEAS TRANSACTIONS on INFORMATION SCIENCE and APPLICATIONS*, Issue 11, Volume 9, November 2012.
- [9] T.Balakumaran, Dr.Ila.Vennila, C.Gowri Shankar, Detection of microcalcification in mammograms using wavelet transform and fuzzy shell clustering, (*ijcsis*) *international journal of computer science and information security*, vol. 7, no. 1, 2010
- [10] Arianna Mencattini, Marcello Salmeri, Giulia Rabottino, Simona Salicone, Metrological characterization of a cadx system for the classification of breast masses in mammograms , *iee transactions on instrumentation and measurement*, vol. 59, no. 11, november 2010.
- [11] Nguyen, H.; Hung, W.T.; Thornton, B.S.; Thornton, E.; Lee, W."Classification of microcalcifications in mammograms using artificial neural networks", *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1998.
- [12] Jasmine, J.S.L.; Govardhan, A.; Baskaran, S."Microcalcification detection in digital mammograms based on wavelet analysis and neural networks", *International Conference on Control, Automation, Communication and Energy Conservation*, 2009. INCACEC 2009.
- [13] Aarthi, R.; Divya, K.; Komala, N.; Kavitha, S."Application of Feature Extraction and clustering in mammogram classification using Support Vector Machine", 2011 Third International Conference on Advanced Computing (ICoAC),
- [14] Chumklin, S.; Auephanwiriyakul, S.; Theera-Umpon, N."Microcalcification detection in mammograms using interval type-2 fuzzy logic system with automatic membership function generation", 2010 IEEE International Conference on Fuzzy Systems (FUZZ).
- [15] Songyang Yu and Ling Guan, A CAD System for the Automatic Detection of Clustered Microcalcifications in Digitized Mammogram Films, *IEEE Tran. on Medical Imaging*, vol. 19, n. 2, February 2000, pp. 115-126.
- [16] Stylianos.D et al., A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry, *computer methods and programs in biomedicine* 2011, 47-63.
- [17] C.W.A.M. van Overveld (Department of Mathematics and Computing Science) and B.Wyvil (Department of Computer Science University of Calgary),"Shrinkwrap: An efficient adaptive algorithm for triangulating an iso-surface".
- [18] K.El Fahssi, A.Eloufidi et al., Mass segmentation in mammograms based on the minimisation of energy and active contour model, *IEEE International Symposium on Medical Measurements and Applications (MeMeA14)*, june 11-12, 2014, LISBON, PORTUGAL.
- [19] Miheala lascu, Dan Lascu Feature Extraction in Digital Mammography Using LabVIEW, 2005 WSEAS Int. Conf. on DYNAMICAL SYSTEMS and CONTROL, Venice, Italy, November 2-4, 2005 (pp427-432).
- [20] <http://peipa.essex.ac.uk/info/mias.html>
- [21] Osslán Osiris Vergara Villegas, Humberto de Jesus Ochoa Dominguez, Vianey Goadalupe Cruz Sanchez, Efrén David Gutierrez Casas, Gerardo Reyes Salgado Rules and Feature Extraction for Microcalcifications Detection in Digital Mammograms Using Neuro-Symbolic Hybrid Systems and Undecimated Filter Banks , *WSEAS TRANSACTIONS on SIGNAL PROCESSING*, issue 8, Volume 4, August 2008.
- [22] Abenaou A., Sadik M. Elaboration d'une methode de compression des signaux aleatoires base d'une transformation orthogonale parametrisable avec algorithme rapide WOTIC2011, ENSEM, Casablanca.
- [23] Abenaou A. Sadik M. Methode et algorithme de formation d'un systeme de fonctions de base adaptables pour le diagnostic des signaux biologiques, *Colloque International des Tlcommunications*, Tanger 2011.
- [24] Good, I.J., The interaction algorithm and practical Fourier analysis, *J. Roy. Statist. Soc. Ser. B*, B-20, 361-372, 1958, B-22, 372-375, 1960.
- [25] R.Mousa et al., Breast cancer diagnosis system based on wavelet analysis and fuzzy-neural, *Expert Systems with Applications* 28 (2005) 713723.



# Modeling of Solar-Diesel Hybrid Central Heating System

Musa Abdalla<sup>1</sup>

<sup>1</sup> Mechanical Engineering Department, gi, Jordan, E-mail: [admin@mechatronix.us](mailto:admin@mechatronix.us)

## ABSTRACT

A complete technical study on a Hybrid heating system is carried out with the aid of computer simulation. The main objective of this work was to provide more insight into combining Solar and Diesel energies to be utilized in domestic central heating. The motivation for such a system was basically the ever increasing Diesel prices in the Kingdom. This work revealed that the cost of Diesel in heating houses could be lowered by at least fifteen percent if the hybrid system is implemented. Complete mathematical model of a representative house was derived for the sake of computer simulation. The model was validated and verified through computer simulations (Matlab Simulink based) with real collected weather data of Jordan (complete year record). Finally a controller strategy was devised and tested using the derived mathematical model.

## 1. INTRODUCTION

Currently the whole world is on an advent of a new era of energy shortage. The energy shortage is evident by the ever-increasing prices of fuels. The impacts on average citizens take different forms: increase in food prices, hijacked gas prices, unbearable heating costs, etc. The main concern of this work is finding engineering practical solutions to lower the central heating costs by tapping into the solar free energy.

Alternative fuels and new sources of energies have attracted many scientists and engineers. One major free source of energy is the Solar Energy. Fortunately, Jordan's weather supports solar based systems, and it provides promising results in that direction. These ideas and more were the motivation and driving forces of this work complement the central heating industry.

The main goal of this project was to make use of solar energy as a complement heating system in conjunction with the classic Diesel energy during winter. Even though this idea is not novel; but the implementation and the scientific approach that we took are new. In addition, the novel control strategy that we have adopted for the hybrid system is more efficient and cost effective than commercial existing systems.

Mechanical and electrical engineers have been working to their extremes to utilize solar energy in conjunction with existing sources of heat (hybrid). A quick glance at the literature indicates two main streams: Photovoltaic and Solar collectors. Unfortunately, the efficiencies currently available is below 10% making it a cost substitute. Nevertheless, as always technology has surprised us with its fast development these photovoltaic arrays will soon be competing in the market. The other stream is to utilize solar collectors to heat a fluid then pass this heat to another fluid via a heat exchanger. This type of system is well established and is widely used even in Jordan.

Despite the lack of encouraging legislation, Jordanian houses have been using solar collectors to supply the domestic hot water demands for the past thirty years. However, the importance of alternative energy is evident in Jordan through the scattered well-established energy

centers. As a matter of fact the University of Jordan is one of the leaders in that direction.

Jordan depends almost exclusively on imported fossil fuels to meet the energy needs; about 99% of its consumption is imported. Jordan possesses one of the most moderate weather around the world. The amount of sun appearance around the year makes it a favorite place to implement solar-based systems. Figure 1 depicts the average monthly solar insolation throughout one year. These data were collected in 1984 [1].

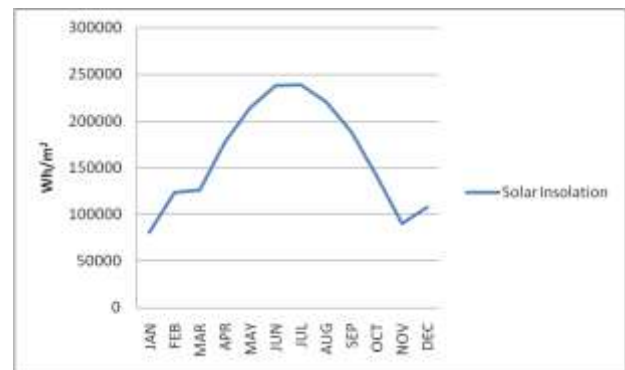


Figure 1: Solar radiation in Jordan 1984

Even for an old data record like this, it shows clearly the potential for alternative sources of power or a complement one (Hybrid). Due to the global warming effect, the amount of sunshine Jordan owns went up and it spreads out to the winter season.

The feasibility of solar energy has been studied for many years, and it started in 1973 with the fuel crisis. A pilot plant was implemented in Coolidge Arizona in 1979 [2]. After the success of this project many plant were constructed in other regions of the world.

In 1981, the International Energy Agency (IEA) initiated a project to build a 500kW plant in Spain [3].

There have been many successful attempts to generate electricity by heating fluid such as oil or water and then produce steam at 200oC, which is then fed to a turbine. The fluid heating was done using a field of parabolic collectors [4, 5].

Many recent studies on the economics of using a Solar/Diesel Hybrid systems are available such as the one provided by reference [6].

## 2. HYBRID SOLAR-DIESEL SYSTEM

### 2.1. Hybrid System Design

Basically, the main idea is to use solar collectors to heat a heat transfer fluid, which in turn will heat the returned circulating water from the central heating unit if appropriate. Figure 1 illustrates the conceptual design of the Hybrid System.

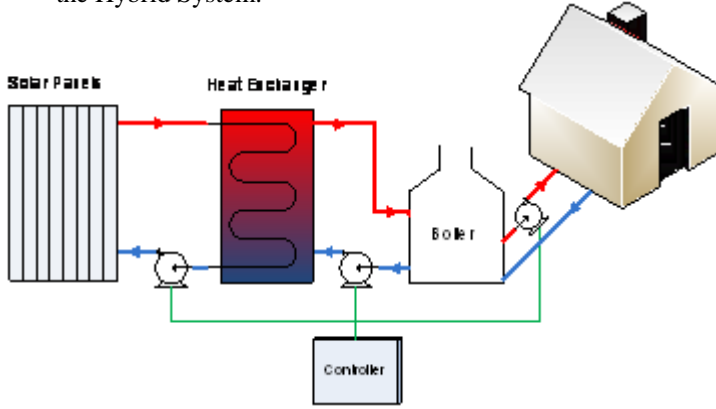


Figure 1: Hybrid System Conceptual Design

The concept of the solar system is very simple, cold water from a storage tank is circulated through the solar collectors if feasible, and if not the circulating pump is shut off. Different implementations of the same idea are possible; the design used was favored in terms of cost and simplicity.

Even though the system is easy to implement, a lot of time was spent in modeling the physical components and to generate a complete mathematical model. The cost savings of the basic system when incorporating a good control strategy exceeds 20%.

### 2.2. Multi-Rooms Modeling

The model that we will adopt assumes heat losses through side walls, doors, windows, ceiling...etc., and all these factors were combined in the value of overall heat transfer coefficient 'U'. Then it is easy to show that the following differential equations will govern the heat distribution in the in a two room system. Note that the walls are well insulated in the sides. Modeling the first room will result in the following DE

$$Q_1 = mC_p \frac{dT_1}{dt} + k_1[T_1(t) - T_a] + k_0[T_1(t) - T_2(t)]$$

Modelling the second room will generate the second coupled differential equation.

$$Q_2 = mC_p \frac{dT_2}{dt} + k_2[T_2(t) - T_a] + k_0[T_2(t) - T_1(t)]$$

Where,  $Q_1$  is the input power for the first room,  $Q_2$  is the input power for the second room,  $m$  is the mass of air inside the room,  $C_p$  is the specific heat of air,  $T_1$  is the first room inside temperature, and  $T_2$  is the second room inside temperature,  $U$  is the overall heat transfer coefficient,  $A$  is total area through which heat loss occurs, and  $T_a$  ambient temperature.

Also, for losses calculations use  $K = UA$  and  $R = C_p m$ .

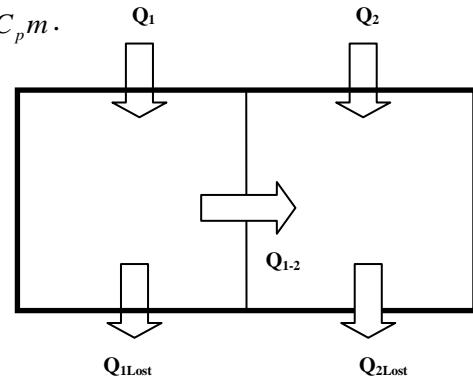


Figure 2: Energy flow for double room space

At this level we are ready to derive a complete mathematical model for a full house model. A representative house is depicted in Figure 3; this house consists of twelve rooms. This mathematical model is similar to the two room model and it is not included here to save space, however interested readers are referred to the FFF technical report [15].

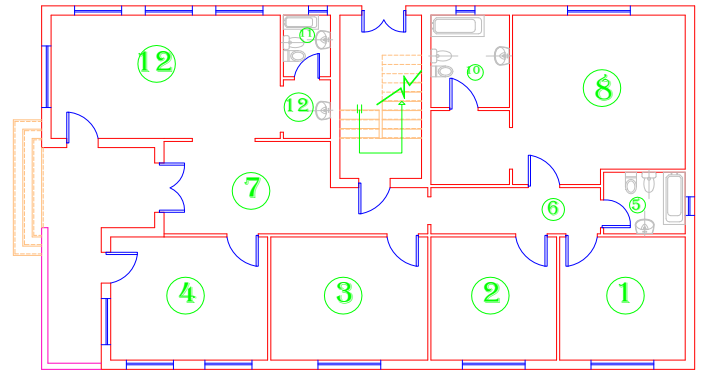


Figure 3: Full house rooms' layout plan view

The temperature distribution over time (system response) for the individual rooms is fully depicted in Figure 4. Step power input is used to perturb the temperature in the rooms.

### 2.3. Heating Equipment Modeling

In this part we shall shed more light on the central heating equipment modeling. Basically, we will look into modeling of radiators, boiler and burner, and solar system. The interaction of these components with the house model that we have derived in the previous section will create a simulated response of the system.

**Rontal Radiator:** We'll use data of commercially available radiators (Metalco, Jordan) which provided the power supplied by each radiator type with a correction factor relative to the temperature difference. The formula is approximated as a linear relationship as follows: between 50 -70 C°: Correction Factor = 0.025 \* ΔT - 0.5 Also, for values above 70 use correction Factor=1.1 and below 50 use correction Factor=0.5.

Temperature Difference	54	56	58	60	62
Correction Factor	0.85	0.9	0.95	1	1.05

Table 1: Radiator correct factor

It turns out such a simple model does not capture the dynamics changes of temperature properly. Instead the following model is used:

$$-U_r A_r (T_r - T) + \dot{m}_r C_p (T_b - T_{ref}) - \dot{m}_r C_p (T_r - T_{ref}) = m_r C_p \frac{dT}{dt}$$

Where,  $U_r$  is overall heat transfer coefficient between the radiator and the room,  $A_r$  is the effective area for convection in between the room and the radiator,  $T_r$  is the temperature of the outlet of the radiator,  $T$  is the room temperature,  $T_{ref}$  is the reference temperature,  $m_r$  is the mass of water inside the radiator,  $\dot{m}_r$  is the mass flow rate of water inside the radiator, and  $C_p$  is the specific heat of water.

**Boiler and the Burner:** To derive the model of the boiler one may take a control volume around the boiler. The governing equation of the boiler becomes

$$U_b A_b (T_g - T_b) + \dot{m}_b C_p (T_r - T_{ref}) - \dot{m}_b C_p (T_b - T_{ref}) = m_b C_p \frac{dT}{dt}$$

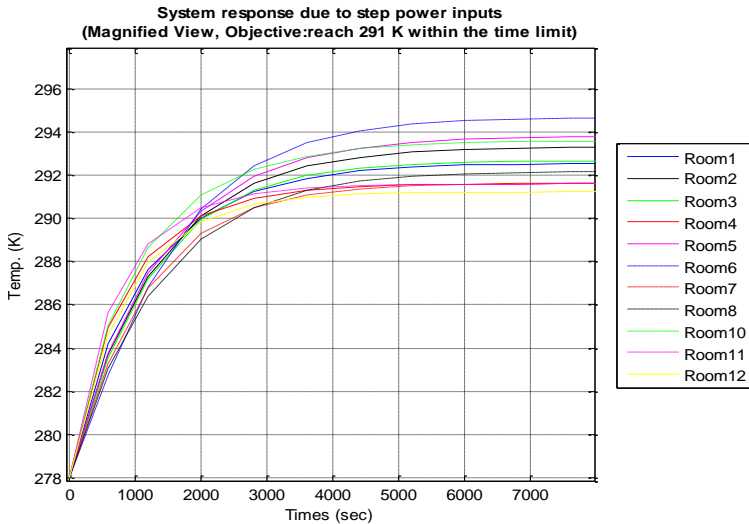


Figure 4: Full house response due to step input

Where,  $U_r$  is the overall heat transfer coefficient between the hot gases and the water inside the boiler,  $A_r$  is the effective area for convection between the hot gases and water inside the boiler,  $T_r$  is temperature of the outlet of the radiator,  $T_b$  is the temperature of the outlet of the boiler,  $T_g$  is the temperature of the hot gases,  $T$  is the room Temperature,  $T_{ref}$  is the reference Temperature,

$m_b$  is the mass of water inside the boiler,  $\dot{m}_b$  is the mass flow rate of water inside the boiler, and  $C_p$  is the specific heat of water.

It must be kept in mind that the temperature of water returned from each radiator is proportional to the mass flow rate in each radiator, and in order to find the temperature of the return line we should make a weighted sum of all temperatures, this can be done as follows:

$$T_{(r)} = \frac{\sum T_{(r)n} \dot{m}_n}{\dot{m}_b}$$

Where,  $\dot{m}_b$  is the total mass flow rate,  $\dot{m}_n$  is the mass flow rate in radiator (n),  $T_{(r)n}$  is the temperature of returned water from the radiator (n), and  $T_{(r)}$  is the temperature of return water.

**Solar System:** The solar system consists of a heat exchanger and the solar collectors. The governing equations for these components are as follows, respectively:

$$U_{ex} A_{ex} (T_h - T_m) = \dot{m}_c C_p (T_h - T_c)$$

$$U_{ex} A_{ex} (T_h - T_m) + \dot{m}_b C_p (T_r - T_m) = m_{b-ex} C_p \frac{dT_m}{dt}$$

Where,  $\dot{m}_c$  is the mass flow rate for the collector,

$\dot{m}_b$  is the mass flow rate for the boiler,  $m_c$  is the mass of water in the collector,  $T_c$  is the temperature of water at the inlet of collector,  $T_h$  is the temperature of water at the outlet of collector,  $C_p$  is the specific heat of water,  $T_{ref}$  is reference Temperature,  $q_{solar}$  is the heat gain by the collector,  $T_a$  is the ambient temperature,  $G_T$  is the irradiance for tilted surface,  $\eta$  is the collector's efficiency,  $U_c$  is the overall heat transfer coefficient between collector and atmosphere,  $A_c$  is the collector area,  $U_{ex}$  is the overall heat transfer coefficient between heat exchanger and accumulator,  $A_{ex}$  is the area for heat exchanger, and  $T_m$  is the temperature water leaving the solar system and entering the boiler.

$$\dot{m}_c C_p (T_c - T_{ref}) - \dot{m}_c C_p (T_h - T_{ref}) + q_{solar} - U_c A_c (T_h - T_a) = m_c C_p \frac{dT_h}{dt}$$

Where,  $\dot{Q}_{solar} = G_T A_c \eta$

In this work, we have used the Gobi Solar Collectors because all needed parameters were readily available from the manufacturer. Total power captured from the sun is given as  $q_{solar} = \dot{\eta} G A$

Where,  $G$  is the absolute radiation from the sun per unit area,  $\eta$  is the collector's efficiency, and  $A$  is the total area of the solar collectors.

The efficiency equation is used as supplied from the manufacturer

$$\eta = 0.7255 - 0.5633 (\Delta T/G) - 0.0022 (\Delta T^2/G)$$

$$\Delta T = T_c - T_a$$

Where,  $T_c$  is temperature of the water entering the solar collectors array and  $T_a$  is the ambient Temperature.

**Domestic Hot Water (DHW):** The governing equations of this system are derived below, it should be borne in mind that it is assumed that the circulation of the solar system water occurs through either the Solar to DHW heat exchanger or Solar to boiler heat exchanger. Please note that we have included an electric heater in the derivations for future work.

$$\begin{aligned} Q_{s-DHW} &= \dot{m}_w C_p (T_h - T_c) = U_{s-DHW} A_{s-DHW} (T_h - T_{DHW}) \\ Q_{b-DHW} &= \dot{m}_{r(13)} C_p (T_b - T_r) = U_{b-DHW} A_{b-DHW} (T_b - T_{DHW}) \\ Q_{s-DHW} + Q_{b-DHW} + Q_{e-DHW} + \dot{m}_{DHW} C_p (T_{city} - T_{DHW}) &= \\ m_{DHW} C_p \frac{dT_{DHW}}{dt} \end{aligned}$$

Where,  $Q_{s-DHW}$  is the power from the solar system to the DHW cylinder,  $Q_{b-DHW}$  is the power from the boiler system to the DHW cylinder,  $Q_{e-DHW}$  is the power from the electric heater system to the DHW cylinder,  $\dot{m}_w$  is the mass flow rate through the solar-DHW circuit,  $\dot{m}_{r(13)}$  is the a portion of the boiler's total mass flow rate which is the mass flow rate through boiler-DHW circuit,  $\dot{m}_{DHW}$  is the mass flow rate from DHW cylinder to the house which represent hot water usage,  $m_{DHW}$  is the mass of water inside DHW cylinder,  $U_{s-DHW}$  is the overall heat transfer coefficient of the solar-DHW heat exchanger,  $U_{b-DHW}$  is the overall heat transfer coefficient of the boiler-DHW heat exchanger,  $A_{s-DHW}$  is the area of the solar-DHW heat exchanger through which convection occurs,  $A_{b-DHW}$  is the area of the boiler-DHW heat exchanger through which convection occurs,  $T_h$  is the temperature of water at the inlet of collector,  $T_c$  is the temperature of water at the outlet of collector,  $T_{DHW}$  is the temperature of water at the outlet of DHW cylinder,  $T_{city}$  is the temperature of water at the inlet of DHW cylinder,  $T_r$  is the temperature of the outlet of the radiator, and  $T_b$  is the temperature of the outlet of the boiler.

Figure 5 illustrates the main components of the simulated system, while Figure 6 depicts the actual implementation of the hybrid system.

## 2.4. Hybrid System On/Off Control

The control strategy adopted in this work is a basic ON/OFF control algorithm that depends on measuring the difference between certain temperatures and actuating certain devices in accordance. The idea was to tested and verify the developed system mathematical model using simulation. The full verification of the model is deferred in another paper, which utilizes hardware to validate the model. The flow chart depicted in Figure 7 summarizes this simple control strategy.

Figure 8 illustrates the controller actions in heating a single room over a day period. At start, the controller turns ON the burner and the radiators' circulation pump to start heating the room; since room's temperature is less than the set point by the specified switch ON differential. The solar collector's pump is ON as well; since the difference between the inlet and the outlet temperatures of the collector is greater than the switching differential. A while later the controller turns OFF the burner and the radiators' pump but the collector's pump is still ON.

## 2.5. Conclusions

The model was tested on real radiation data and provided information about the use of solar energy. The savings were estimated to be approximately 25% of the energy used for heating (using Diesel alone). This is totally dependent not only on the weather conditions but also on the operating hours; operating the system during day time will save a lot of energy; on the other hand operating the system during night only is very costly. Thus for people who don't stay at home during day time and with the absence of powerful energy storage equipment this system will not exploit the use of solar energy fully.

We have tested a lot of operation schemes during different weather conditions, if the system was operating all day long the savings will reach about 5-10%, however if it was used during day time the savings will reach about 25-30%, obviously the savings will be trivial if the system operated during night only.

The model provided a good test-bed for controllers design. However, the proposed model needs to be validated using complete real measurements not just real radiation data. This stepped is deferred for future work.

## Acknowledgments

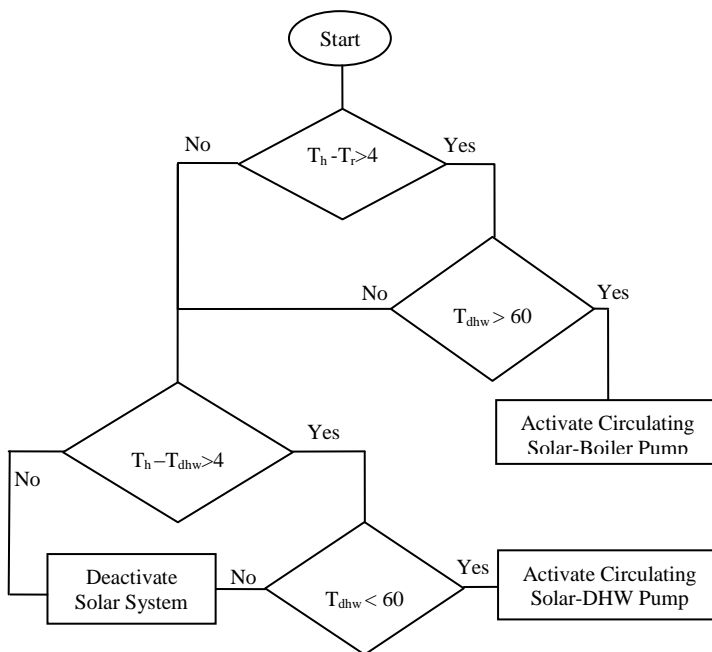
This work was not possible without the research grant (1192, Hybrid Solar-Diesel Central Heating) provided by the Scientific Research Center, University of Jordan.

## References

- [1] Royal Scientific Society, data record of solar energy. A 1984 collected data was granted by private conversation with Dr.Musa Abdalla.
- [2] Larsen, M. Performance of the Coolidge Solar Thermal Electric Power Plant. s.l. : J. Solar Energy Eng. 109. pp. 2-8. Vol. 109.
- [3] Schraub, F and Dehne, H. Electric Generation System Design: Management,Startup and Operation of IEA Distributed Collector System in America,Spain. 31(4). s.l. : Solar Energy J., 1983. pp. 351-354.
- [4]. Price, H, Keamey, D and Replogle, I. Update on the Performance and Operation of SEGS. Miami : ASME International Solar Energy Conference. pp. III-VII.
- [5] Wettermark, G. Performance of the SSPS Solar Power Plant at Almera. s.l. : Solar Energy Eng. J., 1988. pp. 235-246. Vol. 110.

- [6] Linde, Van Der and Sayigh, M. The Economics of Solar/Diesel Hybrid -Case Study. s.l. : Elsevier, 1996. pp. 884-886.
- [7] Schlager, Neil and Weisblatt, Jayne. Alternative Energy. Detroit : Thomson Gale, 2006.
- [8] Galloway, Terry. Solar House. UK : Elsevier, 2004. ISBN 0 7506 58312.
- [9] Sørensen, Bent. Renewable Energy. Denmark : Elsevier Science, 2004.
- [10] Franklin Research Center, ... Installation Guidelines for solar DHW systems. 1979.
- [11] Al-Sharif, Lutfi R. The Transistor and SCR (Article). Amman : s.n., 2007.
- [12] American Society of Heating, Refrigerating and Air Conditioning. ANSI/ASHRAE Standard 135-1995 (BACnet Standard).
- [13]. Stanford, Herber. Analysis and Design of HVAC systems. New Jersey : Prentice Hall, 1988.
- [14] Cengel, Yunus and Turner, Robert. Fundamentals of Thermal Fluid Sciences. s.l. : MC Graw Hill, 2004.
- [15] Musa Abdalla, Hybrid Heating System, FFF Technical report, Faculty of Engineering and Technology, University of Jordan, July 2007.

Figure 7: Simple control strategy for the hybrid system



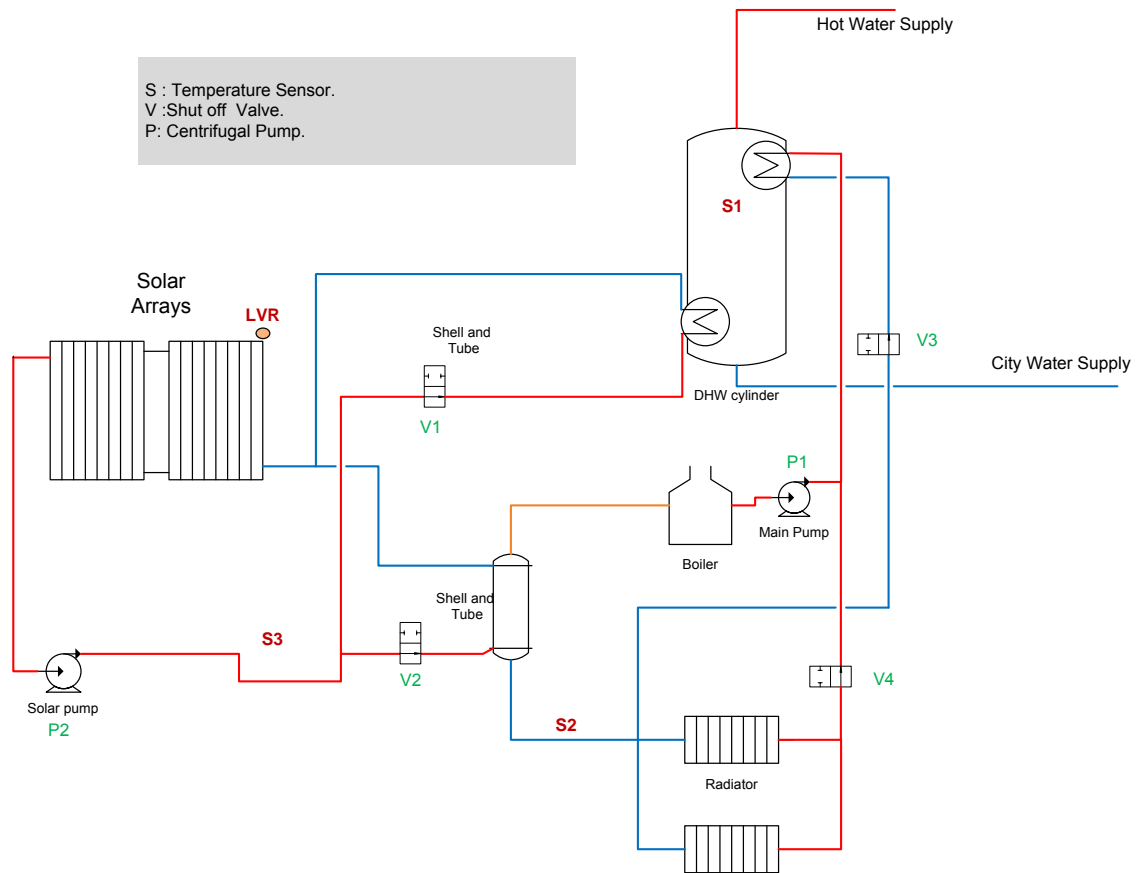


Figure 5: Hybrid system proposed system schematics



Figure 6: Hybrid system implementation at University of Jordan

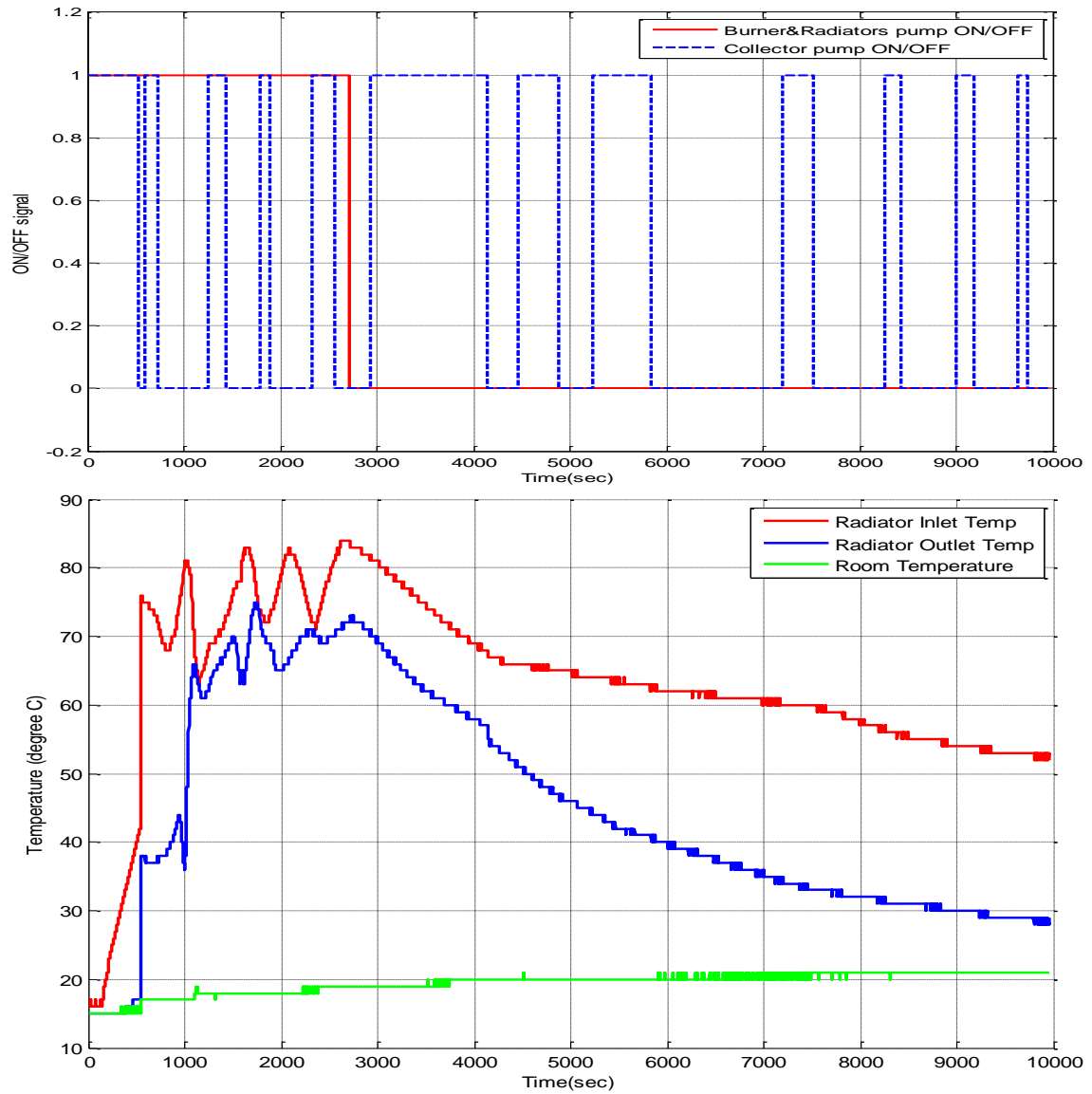


Figure 8: Single room system's response using real data for one day



# Psychological Aspects of Life Values Preferences in the Context of Various University Environment

VĚRA STRNADOVÁ, PETR VOBORNÍK, KATEŘINA PROVAZNÍKOVÁ

Department of Management, Department of Informatics, Department of Applied Linguistics

University of Hradec Králové

Rokitanského 62, Hradec Králové, 500 03

CZECH REPUBLIC

vera.strnadova@uhk.cz, petr.vobornik@uhk.cz, katerina.provaznikova@uhk.cz

**Abstract:** - The article deals with the current population's quality of life standard. It compares the relation between the beginning and the end of university studies and between university students of two regions – the country capital Prague and the regional capital Hradec Králové. It follows up Vital Mind project which was going on at FIM UHK. It presents the results of the research carried out in 2010-2015 as a part of students' projects. A total of 156 university students participated in the study. It was conducted in four groups of respondents. To assess the quality of life standard the method of subjective questioning SEIQoL based on a guided interview was used. The study provides an original newly created application for the electronic completion and evaluation of the questionnaire.

**Key-Words:** - Database, life satisfaction, methods of mental health diagnostics, online application, quality of life, university students.

## 1 Introduction

Life satisfaction, an important indicator of positive psychological well-being, has been defined as a cognitive evaluation of one's overall life or important domains [1]. This study focuses on the life satisfaction of university students of two regions – the country capital Prague and the regional capital Hradec Králové.

Although the few studies of young people have revealed similar findings, most studies have used small-scale samples, limiting their generalization [2].

The studies have suggested that most youth experience positive, overall life satisfaction. However, some studies revealed significant, although small differences related to gender, ethnicity, and socioeconomic status [3]. Most studies employed global measures, failing to differentiate among life satisfaction domains [4].

One exception involved a study of 5545 high school students in South Carolina, U.S.A. Most of these high students reported positive life satisfaction, with respect to global and domain-specific life satisfaction. Although small, race and gender effects emerged for specific domains. Further, a significant number of students reported considerable dissatisfaction with school experiences, suggesting particular concerns for this major life domain. Given differences in cognitive,

social and emotional functioning in early (versus later) adolescence, it is possible that the levels and correlates of life satisfaction differ across the two age groups. Thus, large-scale studies of life satisfaction of the adolescents are needed. [4]

The Vital Mind project at FIM UHK [5] and the issues that are studied in it can be included in the area of Mental Hygiene which is currently heavily involved with *methods of mental health diagnostics*.

These methods focus on the area of overall life satisfaction (in particular the satisfaction in a relationship, family and personal life satisfaction) and then on the state of vitality, the overall feeling of health. In contrast there are presented the states of anxiety, serious problems in life crises or emotional problems and states of depression. Other monitored areas include the ability of self-control and the state of social health, characterized primarily by the number of closed friends, frequency of social contacts and the possibility to discuss one's personal problems with close friends. Considering the mentioned overview of the subjects and the mental hygiene focus we can conclude that most of the matters studied by this discipline belong to the issues of Psychology of health. [6]

Mental hygiene and psychology of health thus target empirically detectable and experimentally observable phenomena which relate to strengthening and maintaining *mental health*. So we approach the



selected subject of our study from the position of a wider, holistic concept of a complex human, including ethical issues of human existence and quality of life.

The concept “Quality of life” can be traced back to the Greek philosopher, Aristotle. He described quality of life in terms of the concept of happiness, experienced when everything works well and the soul is satisfied. Since “life” is the thing everyone leads, pursues and is concerned about, quality of life has been studied extensively. [7]

In the last decades there has been a growing interest towards the concept of “Quality of Life” (QoL), not only in the bio-medical field, but also in other areas, such as sociology, psychology, economics, philosophy, architecture, journalism, politics, environment, sports, recreation and advertisements. Nevertheless QoL does turn out to be an ambiguous and elusive concept – a precise, clear and shared definition appears to be a long way off. [8]

The quality of life can be defined from various points of view. According to the World Health Organization WHO the quality of life is the answer to the question, *how a person perceives his/her position in life in the context of his/her culture and value system*. It also takes into account the objectives, expectations and standards of a person. [9]

WHO recognizes six basic aspects of the quality of life: *physical aspect and the level of autonomy* (the assessment of the amount of physical fatigue, pain, mobility, dependence on medical assistance, ability to work), *psychological health and intellectual realm* (self-concept and self-esteem, the ratio of negative and positive experience, the function of thinking, memory and the ability to concentrate, but also personal faith and spirituality), *social relationships* (personal relationships, intimate life, sources of wider social support) and *the environment* (financial sources, accessibility of medical and social care, conditions of home environment as well as external physical conditions – climatic conditions, amount of pollution, noise). [9]

The quality of life can be considered at different levels [10] – at *macro-level* we study the quality of life of large social units (people of a particular country), at *meso-level* we are interested in the quality of life of small social units (students at university). The third *personal level* monitors directly the life of an individual – here the subject of study of the quality of life can be a person’s way of experiencing and thinking, his/her attitudes, needs and wishes, self-evaluation or the amount of self-

realization. Also negative aspects can be included in this level (the amount of pain during an illness, emotions of sadness during a personal crisis or life trauma).

The quality of life can also be viewed from the *objective or subjective* perspective. While the objective aspect reflects the living conditions of a particular person (state of health, socio-economic status), the subjective aspect constitutes the amount of his/her personal well-being and life satisfaction. To assess a man’s quality of life it is important to know the value system of an individual and his/her different expectations. We also have to know in detail to what extent his/her expectations are in accordance with the reality of life seen from the subjective point of view, this means the way his/her life is going on in reality. In connection with that positive psychology emphasizes the importance of seeking and understanding the factors which contribute to a good, meaningful life and human happiness [11].

In the Czech Republic the issues of measuring the quality of life are dealt with by J. Křivohlavý [10], [12] in his works. He introduces three basic groups of methods used in measuring the quality of life – *objective* (e.g. verbal evaluation according to given criteria), *subjective* (an interview, a self-evaluating questionnaire) and *mixed* (a combination of previous two groups). The majority of methods are designed for the adult population, however there are also quantitative and qualitative methods for detecting children’s and adolescents’ quality of life.

The original concept of *optimal experience* was presented by M. Csikszentmihalyi [13]. He proposed the term *flow* – immersing oneself in an activity. According to this concept a good and happy life is characterized by the frequency and intensity of the *flow* experience, complete immersion in what we are currently doing, devotion to a present moment. Happiness is not an accidental state dependent on external conditions but an internal process which we can cultivate and maintain – “*being able to be happy*”. The quality of life here is closely related to the ability to control one’s consciousness and the art of establishing internal harmony. From our life we deliberately remove obstacles that prevent us from the feeling of inner fulfilment. It is about the ability to get the most pleasure and satisfaction in everything we do.

The summarising results of numerous scientific studies on the subject of human satisfaction and happiness [11], [14], [15], [16] imply that the happiest people:

- spend a lot of time with their loved ones, family members and friends and they maintain good relationships with them
- feel and express their gratitude for everything they have
- offer their help willingly and responsively when it is needed
- see their future optimistically
- can enjoy and relish the good things that the present moment brings
- are physically active, they usually do regular exercises
- have their long-term objectives and plans
- naturally also face problems and experience conflicts and discomfort but they manage to get over difficult periods in their life more easily because they are able to learn from them

Huebner, E. Scott says, that for life satisfaction are important five specific domains – family, friends, self, school and living environment. [17]

## 2 Methodology of research

The case study of this work focused on the quality of life uses the questionnaires of SEIQoL method. This method gives an insight into the individually perceived quality of life. Its principle lies in the fact that the respondent is asked about five most important areas in his/her life. The examples of the areas are health, family and education. The person selects the areas that make him/her happy, goals he/she wants to achieve. Important life goals which a person is aimed at and to achieve them he/she uses smaller goals. [18]

SEIQoL method – The Schedule for the Evaluation of Individual Quality of Life was created in 1994 [19]. It is characterized by five principles:

- defining the qualities of life by a respondent
- a respondent's evaluation of his/her life
- determining the importance of individual aspects by a respondent
- the evaluation is relevant at a given time and at some other time respondent can indicate different values and aspects
- the aspects indicated by a respondent can be medical, psychological or any other aspects

During the evaluation five most important aspects of life are taken into account and the less essential ones are not mentioned. The aspects represent respondents' life goals. That is what they

live for, what makes their life better, what gives the meaning to their life, what they hope for and aim all their energy at. At the same time we detect what importance is given to these aspects by a respondent.

The most suitable way of detecting is the analysis of subjective views and conviction that is *judgement analysis*. In the interview with a respondent the questioner helps to define the importance of the aspects.

Later the authors derived a simplified way called SEIQoL-DW (DW means *direct weighting*) where a five-colour disc is used (see Fig. 1). The respondent indicates the importance of the aspects by shifting the sectors. The bigger a sector is the bigger importance it has.

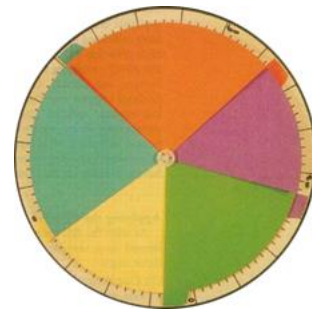


Fig. 1 colour disc of SEIQoL-DW method [20]

Compared to the work with the colour disc the division of a *hundred percent scale* into corresponding sections requires bigger amount of imagination and calculation. There is also the option to talk to respondents and leave the completing itself up to them. There is no tendency to influence or direct respondents, so they do not feel the need to make themselves better during the interview.

An individual completing is suitable in a collective research when individual interviews would take too much time. Here after completing the questionnaire the researcher can ask additional and specifying questions.

A. M. Hickey [20] and his team say that the SEIQoL method is generally applicable and highly valid. He also presents the difference between the fully-fledged method and the simplified Direct Weighting. DW measures only the areas which the respondent realizes in comparison with the fully-fledged method measuring also implicit elements which are not realized by a respondent.

The quality of life was first monitored among sick patients. Thus the result showed the idea of a patient's view of his/her life in the period affected by the illness. At the same time it was possible to follow the development during the illness. But if a researcher is interested in a general quality of life of

the studied group, the result can be distorted by worsened cognitive abilities. It is true that *judgement analysis* is the most suitable way of detecting a respondent's preferences, however, due to its routine clinical use the authors have come up with the *direct weighting* (SEIQoL-DW), which is not as demanding for the patients with limited cognitive abilities [19].

It is important to point out that the SEIQoL method is used primarily for *individual diagnostics* of the selected person's situation. It is not suitable to apply it to large units when a social phenomenon or a human characteristic in the society is studied. It is possible to use the method to compare two groups of respondents. We either study one group in the range of time delay or two groups that differ in a specified feature.

## 2.1 SEIQoL method

SEIQoL method is conducted in the form of a semi-structured interview. First the questioner has to find out the most important areas in the respondent's life. To make it easier the researchers demand 5 areas bearing in mind the fact that the areas represent the most important ones and thus they should reflect the quality of respondent's life sufficiently.

Doing so the questioner has to be aware of a few pitfalls. The respondent does not choose specific objects or people. If the respondent mentions his wife as one of the areas important in his life, the questioner's role is to ask what in particular makes his life happy in connection with his wife. The answer will probably indicate the meaning of sharing time with the beloved person and it is suitable to enter the term "marriage" into the questionnaire. The questioner should avoid influencing the respondent, direct or unintentional inciting him/her or presenting examples in a larger extent than it is necessary. The respondent gives his/her five areas according to which he/she evaluates his/her life. The questioner's role is to find out by using suitable questions whether they are really those most important areas thanks to the respondent feels happier or because of which he/she is not in a good psychological condition. For example, the respondent can claim that one of the essential areas is religion. As a matter of fact the respondent does not mind the faith itself but he/she appreciates the sense of security that the religion gives him. Before both of them move to the part of evaluation the respondent should definitely know what he/she will be evaluating. For example, if the respondent selects work as one of his life themes, the following evaluation may be influenced by the

fact that work is important for the respondent but he/she receives insufficient financial reward for doing the work. The questioner can help the respondent to clarify his/her thoughts but the final decision is up to the respondent.

In the situation when the principal areas determining a person's quality of life are selected the respondent is asked to define how he/she is satisfied with each of the areas. The respondent identifies the amount of satisfaction with a particular area which is expressed as a percentage. The questioner ought to make sure that the respondent really expresses the *current perception of satisfaction* with the entered areas and does not think about the importance of the given areas instead.

The next step is *determining the importance of the given five areas*. We ask the respondent to allocate to each selected area a corresponding amount out of one hundred percent. The use of the disc divided into five sectors (see Fig. 1) is easier for the respondent. Each sector has a different colour and they are movable – overlapping. By shifting the sectors the respondent indicates the importance of each area which can be transferred into numeric values thanks to the scale along the perimeter of the disc. As in the previous step the questioner assures that the respondent really states the personal importance of the selected areas.

In the end we ask the respondent to identify the amount of satisfaction with his/her life in general. The best way how to do it is to prepare a line ("thermometer") where the respondent marks the place which according to him/her reflects his/her current life situation. To enable the respondent to have a better idea about the significance of the line it is suitable to tilt it at the angle of forty five degrees with the left end at the bottom and the right end at the top.

The assessment of the questionnaire answers is carried out by the following calculation. The satisfaction with the first life area is multiplied by the importance of the first life area. The same calculation is done for the other four areas. These five results are added up and the total is divided by one hundred. This way the value between zero and one hundred is obtained and it can be presented as a *percentage value of the quality of life*. If everything is in order, the result should not differ too much from the value which the respondent marked on the "thermometer" at the end of the questionnaire when he/she was evaluating the overall quality of his/her life. [10]

### 3 Results of research

#### 3.1 Group of respondents

For the case study the data from four target groups were collected – university students at the beginning and at the end of their studies and university students of two regions – the capital Prague and the regional capital Hradec Králové. 156 people in total took part in the survey and completed the questionnaire:

- university students at the beginning of their studies: 30 students (age from 20 to 21)
- university students at the end of their studies: 30 (age from 23 to 24)
- university students from Prague: 30 (age from 20 to 25)
- university students from Hradec Králové: 30 (age from 20 to 25)
- university students from previous (2010-2012) research: 36

#### 3.2 Graphic processing of results and data interpretation

##### 3.2.1 Prague vs. Hradec Králové

Firstly we will compare the ranking of life values as it was done by students from the capital Prague and the regional capital Hradec Králové.

Students in both surveyed cities placed traditional values such as *family*, *studies*, *friends* and *work* in the first four positions and in Hradec Králové family was selected as the most important life theme even by all respondents.

Prague		Hradec Králové	
Life theme	Selected	Life theme	Selected
family	90%	family	100%
studies	77%	friends	67%
friends	70%	studies	60%
work	63%	work	60%
partner relationship	47%	health	53%
health	43%	sport	23%
money	30%	money	20%
sport	17%	free time	17%
interests	7%	hobbies	13%
travelling	7%	partner relationship	13%

Fig. 2 comparison of value system of students from Prague and Hradec Králové

From the fifth position on, the values vary both in the percentage representation and in the preference order. Prague students prefer more partner relationship (fifth position) which is placed by Hradec students as far as in the tenth position.

Students from Prague also included travelling among the most preferred life values (see Fig. 2)

##### 3.2.2 Beginning vs. end of studies

The following comparison shows the progression of preference development by the age of students, among the first year students at the age of 20-21 and students of the last third year of studies at the age of 23-24 (see Fig. 3).

1st year		3rd year	
Life theme	Selected	Life theme	Selected
family	83%	family	83%
school, education	67%	school, education	63%
work	47%	work	57%
financial security	43%	health	53%
own housing	37%	friends	50%
travelling	37%	hobbies	33%
be happy	33%	own housing	30%
friends	30%	career	30%
health	27%	travelling	27%
occupation	27%	money	23%

Fig. 3 comparison of value system of first year students and students of the last year of studies

Preferences of life values in both groups are identical in the first three positions: *family*, *education* and *work*. The values show that students at the beginning of their studies are primarily concerned about *financial security* and *own housing*, while at the end of studies they realize the value of their *health* and the importance of *contacts with friends*.

*Money* was shifted as far as the tenth position by senior students due to the fact that these students usually earn their own money and they do not have to rely on financial backing from their parents. First year students prefer the theme of *travelling* more than already quite settled students of the third year.

With advancing age the importance of *permanent employment and career growth* increases. More senior students already realize the degree of seriousness connected with the choice of their future profession.

##### 3.2.3 Year 2012 vs. 2015

Now there will be compared the results of a previous research, which was conducted in the years 2010–2012 with 36 university students [21], and the current research from the years 2013–2015, where the total of 120 students participated. Thus the overall number of students who were considering their value system is 156 (see Fig. 4).

2010-2012		2013-2015	
Life theme	Selected	Life theme	Selected
family	78%	family	89%
self-improvement	61%	school, education	67%
partner	53%	work	57%
health	42%	friends	54%
work	39%	health	44%
friendship	36%	money	28%
money	31%	love	23%
peace of mind	19%	travelling	18%
career	19%	own housing	17%
entertainment	14%	hobbies	14%

Fig. 4 comparison of students' value system from the researches in years 2010-12 & 2013-15

When comparing identical researches among university students in 2012 and 2015 the first position is permanently held by the traditional value of *family*. The second position is occupied by the value of *self-development and education*. Other positions are represented by the same values but in different order.

The value of *health* has the constant rate of preferences (42% and 44%) and the value of *money* has the rate slightly decreasing (31% => 28%). Newly the value of *own housing* occurs in the research. It was not represented among the first ten values at all but nowadays it is a concern for 17% of respondents and also the value of *travelling* is newly added by 18% of respondents.

The importance of *close relationship* held a significant 3rd position in the previously conducted research while in the current group of students it occurs in the 7th position.

### 3.2.4 Satisfaction with life

When assessing the quality of life the SEIQoL method [10] enables to find out also the subjective perception of life satisfaction. Thus students in the 2012-2015 research could totally evaluate their feelings, attitudes and fulfilled wishes in the certain moment of their life. (120 respondents in total, see Fig. 5)

	Estimate	Reality	Difference
Hrade Králové	76,0%	67,8%	8,2%
Prague	82,5%	76,7%	5,8%
1st year	68,6%	58,9%	9,7%
3rd year	67,7%	57,9%	9,8%

Fig. 5 comparison of estimated and real life satisfaction among four different groups in the 2012-2015 research

On an imaginary thermometer (*Amount of satisfaction with life* - see Fig. 6) the perception was recorded in the range from pessimistic, negative to optimistic, positive life-tuning (so called the state of *well-being*). Looking at Fig. 5 we can say that

students from Hradec Králové have the estimated rate of life satisfaction by 8.2% higher than the real rate. Students from Prague have this rate higher by 5.8%. It means that in both cases predominantly optimistic mood of university students has been proved.

This fact is also confirmed in the students who start studying as well as students finishing their university studies – in this group of students the estimated reality differs more from the real one (by 9.7% and by 9.8%) in comparison with the group of students from the selected cities.

### 3.3 Summary of research results

In the area of life values preferences surveyed students from both cities prefer the identical traditional values of *family*, *studies*, *friends* and *work* in the first 4 places. In Hradec Králové *family* was selected the most important life theme even by all respondents.

When comparing the situation at the beginning and the end of university studies the preference of values in the first three positions is the same – *family*, *education* and *work*. The results show that students at the beginning of their studies are mainly concerned about *financial security and own housing*, while at the end of their studies they realize the value of their *health* and the importance of *contacts with friends*.

In the group of senior students the value *money* occurs as far as in the tenth position which is related to the fact that these students mostly earn their own money and thus are not dependent only on financial security from their parents. The first year students prefer *travelling* more than already rather settled students of the third year. With the advancing age the importance of *permanent job and career growth* is also increasing. More senior students are more aware of the degree of seriousness when they choose their future profession.

When identical researches in university students from the years 2012 and 2015 are compared the traditional value of *family* permanently occurs in the first position. The second position is occupied by the value of *self-development and education*. Other positions are represented by the same values but in a different order. The value *health* has a constant rate of preferences (42% and 44%) and the value *money* has a slightly decreasing rate (31% => 28%). Newly the value of *own housing* occurs in the research. Previously it was not represented at all among the first ten values but nowadays 17% of respondents are concerned about it, and the value



*travelling* is also newly included by 18% of respondents.

The importance of *close relationship* had a significant third position in the previously conducted research, while in the current group of students it occurs as far as in the seventh position.

The applied SEIQoL method enabled to record the perception of life satisfaction on an imaginary thermometer in the range from pessimistic, negative to optimistic, positive life mood (so called state of *well-being*). Students from Hradec Králové show the higher estimated rate of life satisfaction. It is by 8.2% higher than the real rate. Students from Prague show this estimated rate higher by 5.8%. Thus it means that in both cases we can prove *predominantly optimistic mood of university students*.

This fact is also confirmed by students who start their university studies and those who finish them – however the estimated reality differs more from the real one in this group of students (by 9.7% and by 9.8%) – in comparison with the same group from selected cities.

FORM TO FIND OUT ABOUT QL BY MEANS OF SEIQoL METHOD

First name and surname: \_\_\_\_\_ Today's date: \_\_\_\_\_ Year of birth: \_\_\_\_\_

Importance of the theme in %	Life theme – What are you after in life above all? (5 life goals)	Amount of satisfaction in % (in each line from 0 to 100 %)
	1.	
	2.	
	3.	
	4.	
	5.	

Total of percentage in all five lines in the left column must equal 100%.

Amount of satisfaction with life:

<-----|-----|-----|-----|-----|----->

It is as bad as it is possible It is as good as it can be

Total value of QL: \_\_\_\_\_

Total amount of satisfaction: \_\_\_\_\_

Fig. 6 quality of life questionnaire [10]

## 4 Application for self-testing

Quality of life measuring by means of SEIQoL method is carried out under the assistance of a trained person who explains the procedure to the

evaluated person, monitors the validity of all data and categorizes carefully described life themes (especially due to the subsequent statistical assessment). The assistance of this evaluator guarantees better accuracy but on the other hand it makes this test inaccessible to a random applicant.

For this reason we have created an web application [22] for quality of life measuring via SEIQoL method, which can be used not only by professionals for their patients, clients and respondents, but also by anybody who is interested in trying out this quality of life test individually and just for themselves. It can be found on the address <http://qol.alltest.eu/seiqol>.

### 4.1 Interface


The interface of the application is user-friendly (see Fig. 7) and its use is clear immediately to more experienced users thanks to the brief captions of individual form elements. In case of difficulties in understanding there is more detailed help available under the icon .

Fig. 7 application form for entering data

As the data processing is immediate and fully automated, the user cannot select the life themes through an extensive description but he/she has to categorize them individually at once. That is why the menu of the previously entered categories of life themes is used instead of the common text box. However this menu also enables to enter an entirely new theme. During typing it filters and offers similar expressions (see Fig. 8) in order to minimize the possibility of duplicate entering the same themes under slightly different designations (e.g. “health” and “state of health”).

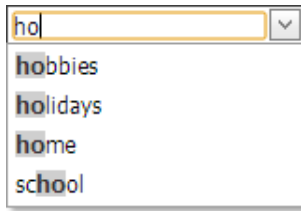


Fig. 8 entering life themes with “insinuation” of previously entered categories

Another critical step is the correct completion of importance weight of each life theme. At each importance weight it is entered the value that is to express this importance but their total must be exactly 100. This is to make the user think about his/her preferences and with consideration divide this limited number of points among the individual themes. The validation function of application secures that the total is really 100. This function does not allow to evaluate the form unless this and other data are entered correctly. Nevertheless not everyone is able to use independently the sufficient amount of imagination and divide these points appropriately.

One of the variants of SEIQoL-DW method uses five-colour disc for this purpose (see Fig. 1). The respondent indicates the importance of the themes by shifting the sectors – the larger sector is set the higher importance the given theme has [20]. This application uses a similar functionality which does not enable to set sectors directly by means of colour disc. However, to make the division of points among the themes more illustrative the application draws the disc in the form of a pie chart (see Fig. 9). This happens during the process of completion without disrupting AJAX functions, immediately after all five themes have been selected and the value of their importance has been entered. This graphic representation thus helps the user get a better idea about the division of points and enables him/her to modify it afterwards.

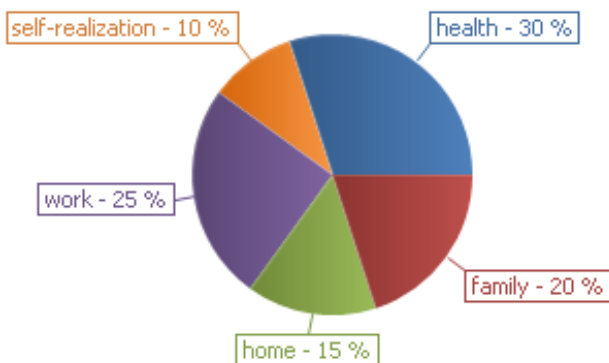


Fig. 9 pie chart illustrating importance of life themes

## 4.2 Evaluation

After completing all data correctly they are evaluated. The user's satisfaction with life is calculated by means of SEIQoL method and it is compared with the result which the user estimated (see Fig. 10). The difference between both values is identified and verbally evaluated.

Your estimated satisfaction with life	73%
Satisfaction with life by SEIQoL	64%
Difference	9%

Fig. 10 table evaluation of estimated and calculated life satisfaction

Both values of satisfaction have their graphic representation in the disc diagram (see Fig. 11). Here we can see very well how the individual values are ranked in this scale and how they compare to each other. For better orientation the scale is at the same time divided into five basic zones which categorizing the quality of life according to the satisfaction with life.

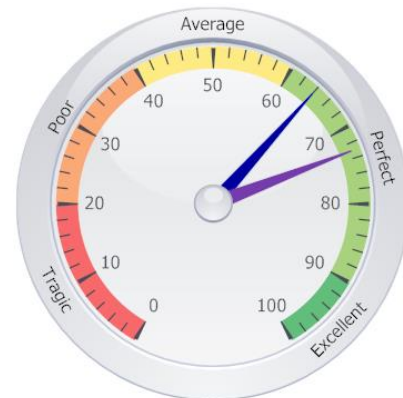


Fig. 11 graphic evaluation of estimated and calculated quality of life

The difference between the estimated and calculated life satisfaction is also depicted. This difference has its significance and according to its extent it is included into an appropriate category expressing the user's ability of self-evaluation. The extent of this difference and its categorization is illustrated by this picture (see Fig. 12).

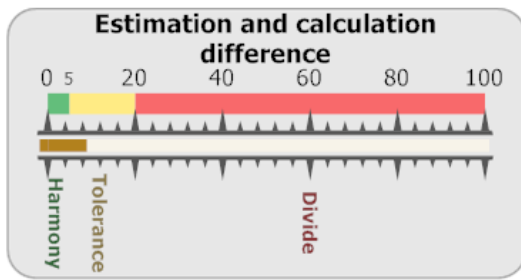


Fig. 12 graphic evaluation of difference between estimated and calculated life satisfaction

The data entered by users are simultaneously saved in a database. Then the user obtains a generated unique code added to the link where he/she can see his/her evaluation at any time later. Moreover the saved data will be used for their later statistical assessment and further research in this field.

For the future we plan to extend the assessment of each measurement by adding the comparison of individual values with the averages of results from the other respondents. The user thus gets information how his/her values compare to those of the whole population, the same age group, the same sex, etc.

## 5 Conclusion

The search for answers to the question “How to lead a good life?” already occupied ancient philosophers. However, quality of life exploring started as late as in the 20th century. Currently there are three basic approaches to the quality of life. The economic approach reduces the quality of life to satisfying consumer preferences. The sociological approach equates the quality of life with the fulfilment of the idea about a desirable social development. According to the psychological approach the quality of life is defined on the basis of subjective ideas of individuals. Interdisciplinary approach to the quality of life and its research is desirable and it is nowadays applied mainly by various independent research centres – e.g. Australian Centre on Quality of Life, Quality of Life Research Unit and others. All the effort connected with recording the quality of life is socially beneficial because a high-quality life is the goal for each of us.

Our research in university students at the age of 20–24 confirmed the preference of traditional life values – family, education and work. The value of health has a constant rate of preference and the value of money shows a decreasing rate of preference. The values of own housing and travelling occur newly in the latter research. In the

area of life satisfaction a predominantly optimistic mood of the surveyed students was proved, both of students from various university cities and of students beginning or finishing their studies.

Until now measuring the quality of life by means of SEIQoL method has been going on only through the subjective questioning in the printed form and with the assistance of a questioner. Now a new application has been created. It enables to carry out this measurement in the electronic form and quite individually by a surveyed person. Another advantage is an immediate automated assessment with the possibility to compare one's result with the other users of the application. Moreover, thanks to this form of data collection our research can be extended to a large number of other respondents.

### Acknowledgments:

This article is supported by the project No. CZ.1.07/2.3.00/20.0001 Information, cognitive, and interdisciplinary research support, financed from EU and Czech Re-public funds. The paper is supported by the Excellence Project N. 2208. *The ICT reflection within the cognitive processes development.*

### References:

- [1] Diener, E., Assessing subjective well-being: Progress and opportunities, *Social Indicators Research* 31, 1994, pp. 103–157.
- [2] Gilman, R., Huebner, E. S., A review of life satisfaction research with children and youth, *School Psychology Quarterly*, 2003, 18, pp. 192–205.
- [3] Šimková E., Psychology and its application in tourism, in: *Social and Behavioral Sciences: Proceeding of the World Conference on Psychology, Counseling and Guidance (WCPCG'13)*, vol. 114, 2014, p. 317–321, ISSN 1877-0428.
- [4] Huebner, E. S., Valoris, R. F., Paxton, R. J., Drane, J. W., Middle school students' perceptions of quality of life, *Journal of Happiness Studies*, vol. 6, issue 1, Mar. 2005, p. 15–24, DOI 10.1007/s10902-004-1170-x.
- [5] Mikulecká, J., Bureš, V., Shatil, E., VM D7\_5\_Verification\_v9(Revised), *The Vital Mind Consortium*, 17th Aug. 2011, 68 p.
- [6] Strnadová, V., *Interpersonal communication*, reviewed monograph, vol. 1, Hradec Králové: Gaudeamus, 2011, 542 p., ISBN 978-80-7435-157-0.



- [7] Wei-Ching, W., Chin-Hsung, K., Tzung-Cheng, H., Chung-Chi, W., Free Time Management Contributes to Better Quality of Life: A Study of Undergraduate Students in Taiwan, *Journal of Happiness Studies*, vol. 12, issue 4, Aug. 2011, pp. 561-573, ISSN: 1389-4978, DOI: 10.1007/s10902-010-9217-7.
- [8] Barcaccia, B., Esposito, G., Matarese, M., Bertolaso, M., Rojo, E. M., De Marinis, M. G., Defining Quality of Life, *Europe's Journal of Psychology*, vol. 19, issue 1, Feb. 2013, pp. 185-203, DOI: 10.5964/ejop.v9i1.484.
- [9] WHOQOL Group, The World Health Organization quality of life assessment (The WHOQOL): Position paper from the World Health Organization, *Social Science and Medicine*, vol. 41, issue 10, Nov. 1995, pp. 1403-1409.
- [10] Křivohlavý, J., *Health psychology*, Prague: Portál, 2001, 279 p., ISBN 80-7178-551-2.
- [11] Slezáčková, A., *Guide of positive psychology*, Prague: Grada Publishing, 2012, 304 p., ISBN 978-80-247-3507-8.
- [12] Křivohlavý, J., *Psychology of diseases*, Praha: Grada Publishing, 2002, ISBN 80-247-0179-0.
- [13] Csikszentmihalyi, M., Csikszentmihalyi, I. S. (Eds), *A life worth living*, New York: Oxford University Press, 2006, ISBN 978-0195176797.
- [14] Efklides, A., Moraitou, D., A positive psychology perspective on quality of life, *Social indicators research series*, vol. 51, VIII, Dordrecht, Heidelberg, New York, London: Springer, 2013, 299 p., ISBN 978-94-007-4962-7.
- [15] Land, K. C., Michalos, A. C., Sirgy, J. M., *Handbook of social indicators and quality of life research*, Dordrecht, Heidelberg, London, New York: Springer, X, 2012, 593 p., ISBN 978-94-007-2420-4.
- [16] Knoop, H. H., Fave, A. D., Well-being and cultures: perspectives from positive psychology, *Cross-cultural advancements in positive psychology serie*, vol. 3, Dordrecht, Heidelberg, New York, London: Springer, VI, 2013, 221 p., ISBN 978-94-007-4610-7.
- [17] Huebner, D. E., Valois, R. F., Raheem, P. J., Wanzer, D. J., Middle school students' perceptions of quality of life, *Journal of Happiness Studies*, vol. 6, is. 1, Mar. 2005, pp. 15-24.
- [18] Křivohlavý, J., *Positive psychology*, Praha: Portál, 2004, 200 p., ISBN 80-7178-835-X.
- [19] O'Boyle, C. A., McGee, H., Joyce, C. R. B., Quality of Life: Assessing the individual, *Advances in Medical Sociology*, 5, 1994, pp. 159-180.
- [20] Hickey, A. M., Bury, G., O'Boyle, C. A., A new short form individual quality of life measure (SEIQoL-DW): application in a cohort of individuals with HIV/AIDS, *British Medical Journal*, vol. 313. 29 p., on-line [5/22/2013], <http://goo.gl/j7NR3> (6th Jul. 1996)
- [21] Strnadová, V., Voborník, P., Tools to Develop Quality of Life for Technical Thinkers in Various Generation Periods, in: *International conference of the Portuguese Society for engineering and education*, Porto, Portugal, pp. 23-30, ISBN 978-989-95907-8-6 (Nov. 2013)
- [22] Šedivý, J., Chromý, J., Hubálovský, Š., Šedivá, K., Research of web tools and mobile devices in education, *Recent Advances in Computer Engineering, Communications and Information Technology*, 2014, pp. 231-235, ISBN: 978-960-474-361-2.

# Authors Index

Abdalla, M.	81
Abenaou, A.	76
Bassiouni, M.	26
Beiraghi, M.	39
Bouyahi, M.	53
Dehri, K.	69
El Fahssi, K.	76
El-Dahshan , El-S. A.	26
Elmoufidi, A.	76
Ezzedine, T.	53
He, Z.	65
Hendy, H.	14
Jai-Andalousi, S.	76
Khalifa, W.	14, 26
Kwak, K.	11
Lee, G.	32
Lee, J.	32
Lee, Y.	32
Lim, H.	11
Lim, Y.	11
Mandourah, S. B. M.	47
Mansour, H. B.	69
Nouri, A. S.	69
Park, S.	32
Provazníková, K.	88
Ranjbar, A. M.	39
Rezig, H.	53
Roushdy, M.	14
Sadouni, K.	58
Salem, A.-B. M.	14, 26, 47
Sekkaki, A.	76
Shim, Y.-C.	18
Sidaoui, B.	58
Strnadová, V.	88
Voborník, P.	88
Yun, M.	11
Zhang, Y.	65
Zheng, D.	65