



FUTURE COMPUTING 2017

The Ninth International Conference on Future Computational Technologies and
Applications

ISBN: 978-1-61208-530-2

February 19 - 23, 2017

Athens, Greece

FUTURE COMPUTING 2017 Editors

Carla Merkle Westphall, Federal University of Santa Catarina, Brazil

Woomin Hwang, National Security Research Institute, South Korea

Kendall E. Nygard, North Dakota State University - Fargo, USA

FUTURE COMPUTING 2017

Forward

The Ninth International Conference on Future Computational Technologies and Applications (FUTURE COMPUTING 2017), held between February 19-23, 2017 in Athens, Greece, continued a series of events targeting advanced computational paradigms and their applications. The target was to cover (i) the advanced research on computational techniques that apply the newest human-like decisions, and (ii) applications on various domains. The new development led to special computational facets on mechanism-oriented computing, large-scale computing and technology-oriented computing. They are largely expected to play an important role in cloud systems, on-demand services, autonomic systems, and pervasive applications and services.

The conference had the following tracks:

- Computational intelligence strategies
- Security and Privacy in Computing Environments
- Computing technologies

We take here the opportunity to warmly thank all the members of the FUTURE COMPUTING 2017 technical program committee, as well as all the reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and effort to contribute to FUTURE COMPUTING 2017. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations and sponsors. We also gratefully thank the members of the FUTURE COMPUTING 2017 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope that FUTURE COMPUTING 2017 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the area of future computational technologies and applications. We also hope that Athens, Greece provided a pleasant environment during the conference and everyone saved some time to enjoy the charm of the city.

FUTURE COMPUTING 2017 Committee

FUTURE COMPUTING 2017 Steering Committee

Cristina Seceleanu, Mälardalen University, Sweden

Hiroyuki Sato, The University of Tokyo, Japan

Kendall E. Nygard, North Dakota State University - Fargo, USA

Alex Wijesinha, Towson University, USA

Albert Zomaya, University of Sydney, Australia
Sergio Ilarri, University of Zaragoza, Spain
Dan Tamir, Texas State University, USA
Wail Mardini, Jordan University of Science and Technology, Jordan

FUTURE COMPUTING 2017 Industry/Research Advisory Committee

Francesc Guim, Intel Corporation, Spain
Yasushi Kambayashi, Nippon Institute of Technology, Japan
Jay Lofstead, Sandia National Laboratories, USA

FUTURE COMPUTING 2017

Committee

FUTURE COMPUTING Steering Committee

Cristina Seceleanu, Mälardalen University, Sweden
Hiroyuki Sato, The University of Tokyo, Japan
Kendall E. Nygard, North Dakota State University - Fargo, USA
Alex Wijesinha, Towson University, USA
Albert Zomaya, University of Sydney, Australia
Sergio Ilarri, University of Zaragoza, Spain
Dan Tamir, Texas State University, USA
Wail Mardini, Jordan University of Science and Technology, Jordan

FUTURE COMPUTING 2017 Industry/Research Advisory Committee

Francesc Guim, Intel Corporation, Spain
Yasushi Kambayashi, Nippon Institute of Technology, Japan
Jay Lofstead, Sandia National Laboratories, USA

FUTURE COMPUTING 2017 Technical Program Committee

Panagiotis D. Bamidis, Aristotle University of Thessaloniki, Greece
Rudolf Berrendorf, Bonn-Rhein-Sieg University of Applied Sciences Sankt Augustin, Germany
Abdelhani Boukrouche, University of Guelma, Algeria
Christos Bouras, University of Patras | Computer Technology Institute & Press <Diophantus>
Greece
Massimiliano Caramia, University of Rome "Tor Vergata", Italy
Chin-Chen Chang, Feng Chia University, Taiwan
Isabel Maria de Sousa de Jesus, ISEP-Institute of Engineering of Porto, Portugal
Leandro Dias da Silva, Instituto de Computação | Universidade Federal de Alagoas, Brazil
Qin Ding, East Carolina University, USA
Francesco Fontanella, Università di Cassino e del Lazio Meridionale, Italy
Félix J. García Clemente, University of Murcia, Spain
J Paul Gibson, Telecom Sud Paris, France
Apostolos Gkamas, University Ecclesiastical Academy of Vella of Ioannina, Greece
Victor Govindaswamy, Concordia University - Chicago, USA
George A. Gravvanis, Democritus University of Thrace, Greece
David Greenhalgh, University of Strathclyde, Scotland
Francesc Guim, Intel Corporation, Spain
Houcine Hassan, Universitat Politècnica de València, Spain
Tzung-Pei Hong, National University of Kaohsiung, Taiwan
Andrei Hossu, University Politehnica of Bucharest, Romania

Daniela Hossu, University Politehnica of Bucharest, Romania
Woomin Hwang, National Security Research Institute, South Korea
Sergio Ilarri, University of Zaragoza, Spain
Yuji Iwahori, Chubu University, Japan
Yasushi Kambayashi, Nippon Institute of Technology, Japan
Michihiro Koibuchi, National Institute of Informatics, Japan
Zbigniew Kokosinski, Cracow University of Technology, Poland
Piotr A. Kowalski, System Research Institute - Polish Academy of Sciences / AGH University of Science and Technology, Poland
Carlos Leon, Universidad de Sevilla, Spain
Vitus Leung, Sandia National Laboratories, USA
Lan Lin, Ball State University, USA
Jay Lofstead, Sandia National Laboratories, USA
Szymon Łukasik, AGH University of Science and Technology / System Research Institute - Polish Academy of Sciences, Poland
Stephane Maag, Telecom SudParis, France
Olaf Maennel, Tallinn University of Technology, Estonia
Giuseppe Mangioni, DIEEI - University of Catania, Italy
Francesco Marcelloni, University of Pisa, Italy
Wail Mardini, Jordan University of Science and Technology, Jordan
Yehya Mohamad, Fraunhofer FIT, Germany
Isabel Muench, German Federal Office for Information Security, Germany
Anand Nayyar, KCL Institute of Management and Technology, Jalandhar, Punjab, India
Kendall E. Nygard, North Dakota State University, USA
Gonzalo Pajares, University Complutense of Madrid, Spain
Prakash Ranganathan, University of North Dakota, USA
Carsten Röcker, Fraunhofer IOSB-INA, Germany
Božidar Šarler, University of Nova Gorica | Institute of Metals and Technology, Slovenia
Hiroyuki Sato, The University of Tokyo, Japan
Peter Schartner, Klagenfurt University - System Security Group, Austria
Elad Michael Schiller, Chalmers University of Technology, Sweden
Andrew Schumann, University of Information Technology and Management in Rzeszow, Poland
Friedhelm Schwenker, University of Ulm, Germany
Cristina Seceleanu, Mälardalen University, Sweden
Manik Sharma, DAV University - Jalandhar, India
Patrick Siarry, Université Paris-Est Créteil, France
Georgios Ch. Sirakoulis, Democritus University of Thrace, Greece
Anca-Juliana Stoica, Uppsala University, Sweden
Mu-Chun Su, National Central University, Taiwan
Young-Joo Suh, POSTECH, Korea
Dan Tamir, Texas State University, USA
Sim-Hui Tee, Multimedia University, Malaysia
Carlos M. Travieso-González, University of Las Palmas de Gran Canaria, Spain
Ugo Vaccaro, University of Salerno, Italy

Teng Wang, Florida State University, USA

Alex Wijesinha, Towson University, USA

Hongji Yang, Bath Spa University, UK

Peng-Yeng Yin, National Chi Nan University, Taiwan

Fernando Zacarias, Universidad Autonoma de Puebla, Mexico

Ales Zamuda, University of Maribor, Slovenia

Erliang Zeng, University of South Dakota, USA

Zhiyong (Zeroun) Zhang, Henan University of Science & Technology, China

Albert Zomaya, University of Sydney, Australia

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission to reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Real Power Loss Optimization for a Hydrocarbon Industrial Plant Using Genetic Algorithm and Differential Evolution Algorithm <i>Muhammad Alhajri, Mohamed Darwish, and Mohammed Abido</i>	1
Using BP Neural Network for Adapting Payout Time in Communication Networks <i>Sara Helmi, Niveen Mohamed Badra, and Mohamed Elkattan</i>	7
Electricity Price Forecasting in a Smart Grid <i>Md Mahbubur Rahman Khan, Md Minhaz Chowdhury, and Kendall Nygard</i>	14
A Class of Minimum-Time Synchronization Algorithms for 2D Rectangular Arrays Based on L-Shaped Mapping <i>Hiroshi Umeo, Takuya Yamawaki, Hiroki Uchino, and Kaori Ishida</i>	20
APT Detection with Host-Based Intrusion Detection System and Intelligent Systems <i>Seong Oun Hwang</i>	26
Memristive Implementation of Fuzzy Logic for Cognitive Computing <i>Martin Klimo and Ondrej Such</i>	28
Lyapunov-based Control Theory of Closed Quantum Systems <i>Shuang Cong and Fangfang Meng</i>	32
Performance Evaluation of DSDV and AODV Protocols for Green Corridor Management in a Metropolitan City <i>Sandeep Sagar Kariyappa Shivappa, Asha Hanumantharaya, and Narendra Kumar Gurumurthy</i>	38
Neural Signal Processing and Motion Capture as a Feedback Mechanism to Improve Interceptive Human Movement <i>Devanka Pathak, Tin-Kai Chen, Meiyu Shi, and Hongji Yang</i>	44

Real Power Loss Optimization for a Hydrocarbon Industrial Plant Using Genetic Algorithm and Differential Evolution Algorithm

M. T. Al-Hajri
Computer & Electronic Eng. Dept.
Brunel University
Uxbridge, United Kingdom
e-mail:muhammad.t.alhajri@gmail.com

M. A. Abido
Electrical Engineering Department.
King Fahad University (KFUPM)
Dhahran, Kingdom of Saudi Arabia
e-mail:mabido@kfupm.edu.sa

M. K. Darwish
Computer & Electronic Eng. Dept.
Brunel University, U.K.
Uxbridge, United Kingdom
e-mail:mohamed.darwish@brunel.ac.uk

Abstract- In this paper, a techno-economic assessment of a real life hydrocarbon facility electrical system real power loss optimization is addressed. This optimization was attained by using the Genetic Algorithm (GA) and the Differential Evolution Algorithm (DEA). The study is the first of its kind as none of the previous studies were conducted in the context of a real life hydrocarbon facility's electrical system. The hydrocarbon facility's electrical system examined in the study consisted of 275 buses, two gas turbine generators, two steam turbine generators, and large synchronous motors, with both rotational and static loads. For the real life hydrocarbon facility, the performance of the GA and the DEA were benchmarked in the course of optimizing the subject objective. The problem was articulated as a constrained nonlinear problem. The constraints were all real values reflecting the system equipment and components' limitations. The consequences obtained from the study showed the efficiency and prospects of the proposed algorithms in solving the described optimization case. Also presented in this study is the annual fuel cost avoidance.

Keywords- genetic algorithm; differential evolution algorithm; power loss optimization; hydrocarbon facility; millions of standard cubical feet of gas (MMscf).

I. INTRODUCTION

Most of the oil exporting developing countries are facing a challenge associated with the increasing demand for domestic electrical energy. This increase has reached such an alarming level that it mandates action from the governments of the subject countries. For example, in the Kingdom of Saudi Arabia, the average annual increase in electricity demand is 7.4% [1].

In fact, in these countries, a high percentage of electric generation comes from low efficiency power generation plants, such as the simple cycle steam turbine. This complicates the issue, creating an urgent requirement for the utilization of more efficient plants coupled with a reduction in loss in the transmission and distribution system. In Saudi Arabia, the distribution of plant capacity for electricity generation by technology illustrates that low efficiency simple cycle steam turbine generators make up 32% of the utility company's generation fleet, while the most efficient combined cycle generators are around 13.8% of the whole fleet [2].

The aforementioned challenges gave impetus to the idea of studying the potential to use intelligent algorithms in optimizing real hydrocarbon facility power loss. The study used the real values of the system parameters and practical constraints, which escalated the challenges in finding a global solution.

This paper considers an existing real life hydrocarbon central processing facility electrical power system model for assessing the potential of system real power loss minimization using the Genetic Algorithm (GA) and the Differential Evolution Algorithm (DEA) for two generation modes. In Section 2, the problem will be formulated as optimization problem with equality and inequality constraints. In Section 3, the GA and DEA will be employed to solve this problem. In Section 4, the paper study scenarios will be developed. Finally, in Section 5 techno-economic analysis of the results and discussion will be presented.

II. PROBLEM FORMULATION

The problem formulation consists of two parts: the development of the objective functions and the identification of the system electrical constraints to be met; equality and inequality constraints.

A. Problem Objective Function

The objective to be achieved is the minimization of the real power loss J_1 (P_{Loss}) in the transmission and distribution lines. This objective function can be expressed in terms of the power follow loss between two buses i and j as follows:

$$J_1 = P_{Loss} = \sum_{k=1}^{nl} g_k [V_i^2 + V_j^2 - 2 V_i V_j \cos(\delta_i - \delta_j)] \quad (1)$$

Where nl is the number of transmission and distribution lines; g_k is the conductance of the k^{th} line, $V_i \angle \delta_i$ and $V_j \angle \delta_j$ are the voltage at end buses i and j of the k^{th} line, respectively [3] [4].

B. Problem Equality and Inequality Constraints

The system constraints are divided into two categories: equality constraints and inequality constraints [3][5]. Details are as follows:

B.1 Equality Constrains

These constrains represent the power load flow equations. The balance between the active power injected P_{Gi} , the active power demand P_{Di} and the active power loss P_{li} at any bus i is equal to zero. The same balance applies for the reactive power Q_{Gi} , Q_{Di} , and Q_{li} . These balances are presented as follows:

$$P_{Gi} - P_{Di} - P_{li} = 0 \tag{2}$$

$$Q_{Gi} - Q_{Di} - Q_{li} = 0 \tag{3}$$

The above equations can be detailed as follows:

$$P_{Gi} - P_{Di} - V_i \sum_{j=1}^{NB} V_j [G_{ij} \cos(\delta_i - \delta_j) + B_{ij} \sin(\delta_i - \delta_j)] = 0 \tag{4}$$

$$Q_{Gi} - Q_{Di} - V_i \sum_{j=1}^{NB} V_j [G_{ij} \sin(\delta_i - \delta_j) - B_{ij} \cos(\delta_i - \delta_j)] = 0 \tag{5}$$

where $i = 1, 2, \dots, NB$; NB is the number of buses; P_G and Q_G are the generator real and reactive power, respectively; P_D and Q_D are the load real and reactive power, respectively; G_{ij} and B_{ij} are the conductance and susceptance between bus i and bus j , respectively.

B.2 Inequality Constrains

These constrains represent the system operating constrains posted in Table 1 and they are as follows:

- a. Generator and synchronous motor voltages; V_G and V_{Synch} ; their reactive power outputs; Q_G and Q_{Synch} .
- b. The transformers taps.
- c. The load buses voltages V_L .

Combining the objective function and these constrains, the problem can be mathematically formulated as a nonlinear constrained single objective optimization problem as follows:

Minimize J_f

Subject to:

$$g(x,u) = 0 \tag{6}$$

$$|h(x,u)| \leq 0 \tag{7}$$

Where:

x : is the vector of dependent variables consisting of load bus voltage V_L , generator reactive power outputs Q_G and the synchronous motors reactive Power Q_{Synch} . As a result, x can be expressed as

$$x^T = [V_{L1} \dots V_{LNL}, Q_{G1} \dots Q_{GNL}, Q_{Synch1} \dots Q_{SynchNSynch}] \tag{6}$$

u : is the vector of control variables consisting of generator voltages V_G , transformer tap settings T , and synchronous motors voltage V_{Synch} . As a result, u can be expressed as

$$u^T = [V_{G1} \dots V_{GNL}, T_1 \dots T_{NT}, V_{Synch1} \dots V_{SynchNL}] \tag{8}$$

g : are the equality constrains.

h : are the inequality constrains.

All the constraints posted in Table 1 are real values based on the system and equipment real data.

TABLE 1
SYSTEM INEQUALITY CONSTRAINS

Description	Lower Limit	Upper Limit
GTG Terminal Voltage (V_{GTG})	90%	105%
STG Terminal Voltage (V_{STG})	90%	105%
GTG Reactive Power (Q_{GTG}) Limit	-62.123 MVAR	95.72 MVAR
STG-1 Reactive Power (Q_{STG}) Limit	-22.4 MVAR	20.92 MVAR
STG-2 Reactive Power (Q_{STG}) Limit	-41.9 MVAR	53.837 MVAR
Captive Synch. Motors Terminal Voltage	90%	105%
Synch. Motors Terminal Voltage (V_{Synch})	90%	105%
Causeway downstream Buses Voltage	95%	105%
All Load Buses Voltage	90%	105%
Main Transformer Taps	+16 (+10%)	-16 (-10%)
Generators Step-Up Transformer Taps	+8 (+10%)	-8 (-10%)

III. THE PROPOSED APPROACH

A. Generic Algorithm (GA) Implementation

The implementation of the GA technique can be summarized in the following steps [6]-[12]:

- 1) Generate initial populations of chromosomes; each chromosome consists of genes and each of these genes represents either transformer tap settings, synchronous motors voltages or the generators voltages values.
- 2) Assign fitness to each chromosomes, as follows:
 - a. Use the Newton-Raphson method to calculate the real power losses for each population [8].
 - b. Identify if the voltage constrains are satisfied.
 - c. Identify if the synchronous machines (generators and motors) capacity limitations are met.
 - d. Assign fitness values to the populations that meet all the constrains; the population best power loss value (J_f).
- 3) Identify the best population with its associated chromosomes that has the best objective function value and store it.
- 4) Identify the chromosomes parents that will go to the mating pool for producing the next generation via the random selection method. This method works by generating two random integer numbers (each represents a chromosome). Then, these two randomly selected chromosomes fitness values are compared and the one with the better fitness value will go into the mating pool. This randomly selected chromosomes mechanism will be repeated until the population in the mating pool equals to the initial chromosomes population [13].
- 5) Perform genes crossover for the mating pool parents via the simple crossover method [13]. In this method, the offspring chromosomes are generated by establishing a vertical crossover position for parent's chromosomes and then crossover their genes.

- 6) Perform gene mutation for the mating pool parents after they have been crossed over; the random mutation method was implemented [13]. In this method, the offspring chromosomes genes are mutated to new ones randomly from the genes domain.
- 7) Go to Step #2 and repeat the above steps with the new populations generated from the original chromosome parents after being crossed over and mutated.
- 8) Each time, identify the best population and compare its fitness value with the stored one; if it is better (meeting the objective function), replace the best chromosomes with the new ones.
- 9) The loop of generation is repeated until the best population with its associated chromosomes, in terms of minimum real power loss, is identified or the maximum number of generations is met. The flow chart of the proposed approach is shown in Fig. 1.

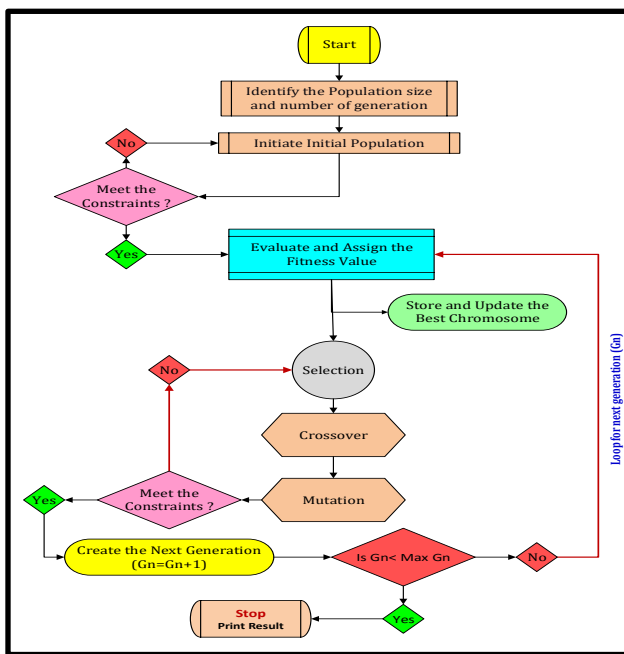


Figure 1. The GA algorithm evolution process flowchart

B. Differential Evolution Algorithm (DEA) Implementation

DEA utilizes special differential operators in creating the offspring individuals from the parent individuals' population in place of the classical crossover and mutation operators used in the GA. In DEA, there are two control parameters, which are the mutation constant F and the crossover constant C. Different from the GA, in the DEA the mutations are performed before the crossover and the selection is taken place after both the mutation and the crossover. DEA's first three evolutionary process steps are similar to the GA ones. [14]-[18]. The remaining steps are as described below:

- 4) In the DEA, mutations are performed using the DE/rand/1 mutation technique [17]. $V_i(t)$ - the mutated vector, is created for each population member $X_i(t)$ set by randomly selecting three individuals' x_{r1} , x_{r2} and x_{r3} and not

corresponding to the current individual x_i . Then, a scalar number F is used to scale the difference between any two of the selected individuals. The resultant difference is added to the third selected individual. The mutation process can be written as:

$$V_{ij}(t) = x_{r1j}(t) + F * [x_{r2j}(t) - x_{r3j}(t)] \quad (9)$$

The value of F is usually selected between 0.4 and 1.0. In this study, F was set to be 0.5 (50%). In [14], scaling mutation based on the frequency of successful mutations is applied.

- 5) Perform the binomial crossover, which can be expressed as follows:

$$u_{i,j}(t) = \begin{cases} v_{i,j}(t) & \text{if } rand(0,1) < CR \\ x_{i,j}(t) & \text{else} \end{cases} \quad (10)$$

CR is the crossover control parameter, and it is usually set within the range [0, 1]. The child $u_{i,j}(t)$ will compete with its parent $x_{i,j}(t)$. CR is set equal to 0.9 (90%) in this study.

- 6) Perform the selection procedure as described below:

$$x_i(t+1) = u_i(t) \quad \text{condition } f(u_i(t)) \leq f(x_i(t)) \quad (11)$$

$$x_i(t+1) = x_i(t) \quad \text{condition } f(x_i(t)) \leq f(u_i(t)) \quad (12)$$

Where $f()$ is the objective function to be minimized.

- 7) Looping back for the terminating criteria. If the criteria are not fulfilled, then generate new offspring population and begin again.
- 8) If the termination criteria are met, identify the best population with its associated chromosomes, in terms of minimum real power loss. The DEA evolution process is shown in Fig. 2.

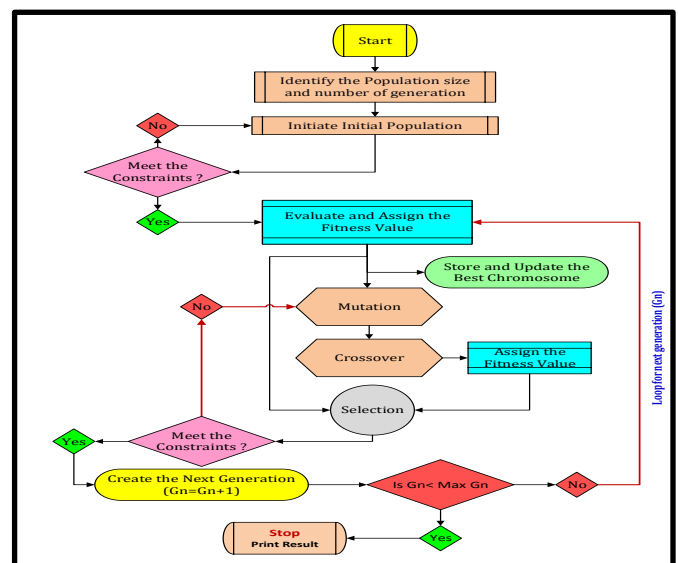


Figure 2: DEA in single-objective mode evolutionary process chart

IV. STUDY SCENARIOS

In this paper, three scenarios were studied: the base case scenario- business as usual (BAU), the optimal case scenario when all generators are online, and the optimal case scenario when two generators are offline. In the optimal cases, the best system parameters (chromosomes) that meet the minimum objective function (J_1) are obtained.

A. Base Case Scenario

The BAU scenario was simulated to be benchmarked with the two optimal scenarios. Following are some of the normal system operation mode parameters:

- 1) The utility bus and generators terminal buses were set at unity p.u. voltage.
- 2) All the synchronous motors were set to operate very close to the unity power factor.
- 3) All downstream distribution transformers and the captive synchronous motors transformers; off-load tap changers; were put on the neutral tap.
- 4) The causeway substations main transformers taps were raised to meet the very conservative voltage constrains at these substations downstream buses; ≥ 0.95 p.u. Refer to Table 2 below.

TABLE 2
THE SELECTED FEASIBLE TRANSFORMERS TAPS VALUE

Substation Number	Transformer Tap
Causeway Substation#1	+3 (1.019 p.u.)
Causeway Substation#2	Neutral (1.0 p.u.)
Causeway Substation#3	+3 (1.019 p.u.)
Main Substation Transformers	+1 (1.006 p.u.)

B. Optimal Case Scenario with All Generators Online

In this scenario, all the generators were assumed to be online. The initial 100 populations of feasible chromosomes (individuals), which meet both the buses voltage and synchronous machine reactive power constrains were identified. The feasible populations with their associated chromosomes were subject to the GA and DEA evolutionary process of 100 generations guided by the objective function J_1 . The GA process was set with 90% crossover probability and 10% mutation probability. In the DEA case, the mutation F constant was set equal to 0.5 (50%) and the CR is set equal to 0.9 (90%). The system parameters and the objective function value associated with the optimal solution of this scenario were identified.

C. Optimal Case Scenario with Two Generators Offline

In this scenario, two generators (one gas turbine generator and one steam turbine generator) were assumed to be offline. All others parameters are identical to the optimal case scenario with all generators are online.

V. RESULTS ANALYSIS AND DISCUSSIONS

The results from the three scenarios, base case, when all generators online and with two generators offline will be

analysed in two categories: the system parameters analysis and the economic analysis.

A. System Parameters Analysis

The hydrocarbon facility simplified electrical system model, which is studied in this paper, is shown in Fig. 3. The evolution of the objective function (J_1) values over the GA and DEA evolution process is captured in Fig. 4. In the all generators online scenario the GA converted to its optimal J_1 value of 1.892 MW after 53 generations while the DEA converted to better J_1 value of 1.885 MW but after 78 generations. So, DEA converted to a better J_1 value but after a higher number of generations. In the scenario with two generators offline, DEA converted to again better J_1 value of 2.926 MW at generation 72 while the GA converted to its optimal J_1 value of 2.933 MW after 91 generation. In this scenario, DEA produced a better J_1 value and within a lower number of generations compared to the GA. The evolution process for both the GA and the DEA were repeated many times to confirm the obtained results.

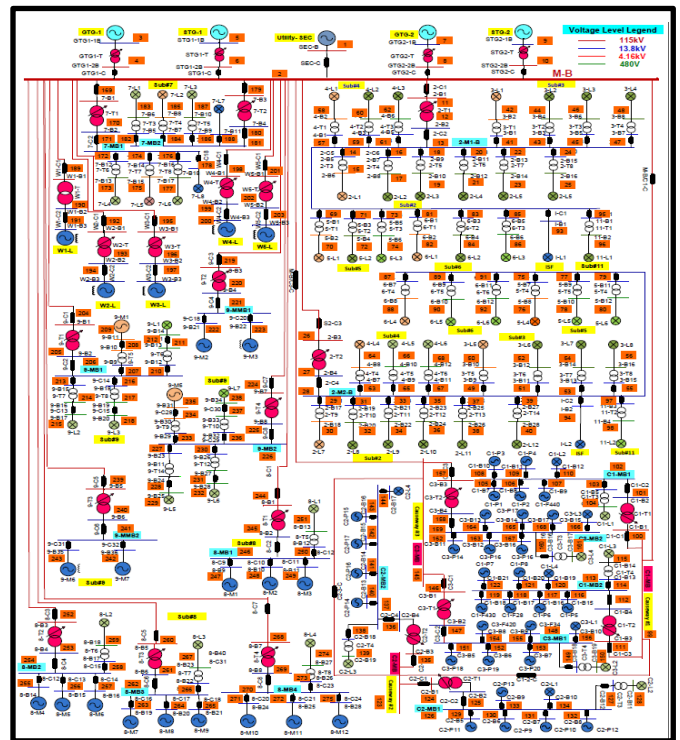


Figure 3. Simplified electrical system of the hydrocarbon processing facility

In Table 3, a comparison of the objective functions' values are posted for the three studied scenarios. The DEA performs better than the GA for the all generators' online scenario. J_1 was reduced by 11.67% when compared with the BAU values. Also, DEA produced better objectives' values than GA for the two generators' offline mode.

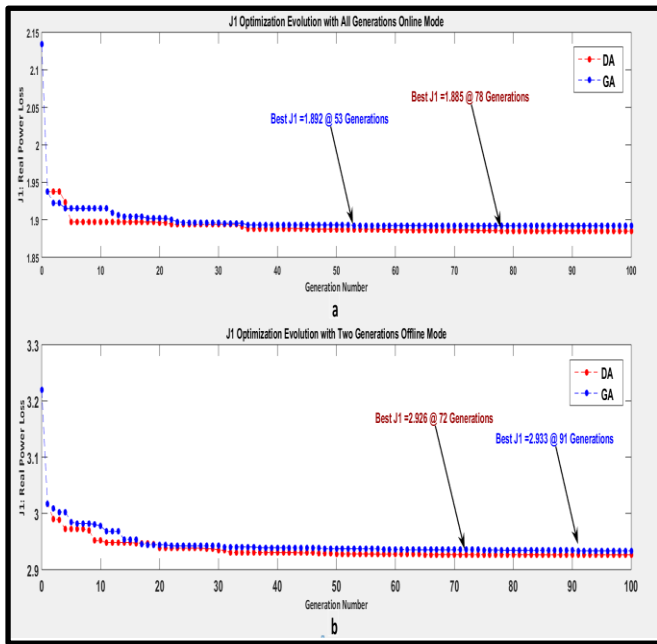


Figure 4. J_1 value convergent for the two generation scenarios

TABLE 3
THE J_1 VALUES FOR THE STUDIED THREE SCENARIOS

Generation Mode	Case #	J_1 Value	$\Delta J_1\%$
All online	BAU	2.134	0.00
Two offline	BAU	3.219	0.00
All online	GA	1.892	-11.34%
All online	DEA	1.885	-11.67%
Two offline	GA	2.933	-8.89%
Two offline	DEA	2.926	-9.10%

B. Economic Analysis

The annual cost due to the system real power loss calculated based on natural gas cost of \$3.5 per MM is demonstrated in Table 4. In the all generators online scenario, DEA produce better real power loss cost \$561,782 which is \$74,200 less than the BAU power loss cost and \$2,076 less when benchmarked with GA. Also, DEA produce better real power loss of \$872,124 which is \$87,361 less than the BAU power loss cost and \$2,161 less when compared with GA real power loss cost value.

TABLE 4
ECONOMIC ANALYSIS FOR THE STUDIED THREE SCENARIOS

Generation Mode	Case#	Real Power Loss Cost
All online	BAU	(635,982.30)
Two offline	BAU	(959,485.76)
All online	GA	(563,858.10)
All online	DEA	(561,782.11)
Two offline	GA	(874,285.01)
Two offline	DEA	(872,124.64)

VI. CONCLUSION AND FUTURE WORK

This paper presented the potential of minimizing the real system’s power loss for a real-life hydrocarbon facility using

the GA and DEA approach. Three scenarios were considered, the base case scenario and two generation scenarios. The technical and economic advantages of the optimal scenarios versus the base case scenario were highlighted in this paper. Also, the superiority of applying DEA to search for the optimal value of the objective function over the GA was highlighted. Future study may need to address the problem as a multi-objective problem considering the grid connection power factor as a second objective which competes with the real power loss minimization objective.

ACKNOWLEDGMENTS

The authors acknowledge the support of the Power System Operation Department/Saudi Aramco Oil Company, Brunel University and King Fahd University of Petroleum & Minerals for their support and encouragement throughout the study.

REFERENCES

- [1] “Saudi Arabia historical peak demand”, http://ecra.gov.sa/peak_load.aspx#.VOg-5105BKA, accessed March 2016.
- [2] “Saudi Electrical Company 2014 annual report”, <https://www.se.com.sa/en-us/Lists/AnnualReports/Attachments/14/AnnualReport2014En.pdf>, accessed March 2016.
- [3] D. Godwin Immanuel and C. Chritober Asir Rajan, “An genetic algorithm approach for reactive power control problem,” IEEE International Conference on Circuits, Power and Computing Technologies (ICCPCT), pp. 74–78, 2013.
- [4] A. A. El-Fergany, “Optimal capacitor allocations using evolutionary algorithms,” IET Generation, Transmission & Distribution, vol. 7, Iss.6, pp. 593–601, 2013.
- [5] Y. Zeng and Yanguang Sun, “Comparison of multiobjective particle swarm optimization and evolutionary algorithm for optimal reactive power dispatch problem,” IEEE Congress on Evolutionary Computation (CEC), Beijing, China, pp. 258–265, July 6-11, 2014.
- [6] W. N. W. Abdullah, H. Saibon and K. L. Lo, “Genetic Algorithm for Optimal Reactive Power Dispatch,” IEEE Catalog No: 98EX137, pp.160-164, 1998.
- [7] M. A. Abido and J. M. Bakhshwain, “Optimal VAR Dispatch Using a Multiobjective Evolutionary Algorithm,” International Journal of Electrical Power & Energy Systems, Vol.27, No. 1, pp.13-20, Jan. 2005.
- [8] M. A. Abido, “Intelligent Control Course Notes,” King Fahad University of Petroleum & Minerals, 2007.
- [9] M. R. Silva, Z. Vale, H. M. Khodr and C. Ramos, “Optimal Dispatch with Reactive Power Compensation by Genetic Algorithm,” Transmission and Distribution Conference and Exposition, 2010 IEEE PES.
- [10] P. Yonghong and L. Yi, “An Improved Genetic Algorithm for Reactive Power Optimization,” 30th Chinese Control Conference (CCC), China, pp.2105-2109, 2011.
- [11] R. K. Kapadia and N. K. Patel, “Reactive Power Optimization Using Genetic Algorithm,” IEEE, Engineering (NUiCONE), Ahmedabad, 2013.
- [12] M. T. Al-Hajri, M. Abido and M. K. Darwish, “Power optimization for a hydrocarbon industrial plant using a genetic algorithm,” IEEE, Universities Power Engineering Conference (UPEC), September 2-5, 2014, Romania, Cluj-Napoca.

- [13] M. T. Al-Hajri and M. Abido “Assessment of Genetic Algorithm Selection, Crossover and Mutation Techniques in Reactive Power Optimization,” IEEE, CEC 2009, pp.1005-1011, May 8-21, 2009.
- [14] K. P. Wong and ZhaoYang Dong, “Differential Evolution, an Alternative Approach to Evolutionary Algorithm” , Proceedings of the 13th International Conference on Intelligent Systems Application to Power Systems,Arlington, VA, Nov. 6-10, pp. 73 – 83, 2005.
- [15] M. Varadarajan and K. S. Swarup, “Solving multi-Objective optimal power flow using differential evolution”, IET Generation, Transmission & Distribution, vol. 2, No. 5, pp. 720–730, 2008.
- [16] S. R. Spea, A. A. Abou El Ela and M. A. Abido, “Multi-objective differential evolution algorithm for environmental-economic power dispatch”, IEEE International Energy Conference, Manama,pp. 841–846, 2010.
- [17] S. Das and P. N. Suganthan “Differential Evolution: A Survey of the State-of-the-Art”, IEEE Transactions on Evolutionary Computation, Volume 15, Issue.1, pp. 4 – 31, 22 February 2011.
- [18] M. T. Al-Hajri, M. Abido and M. K. Darwish, “Multiobjective Power Loss Optimization Versus System Stability Assessment for Hydrocarbon Industrial Plant Using Differential Evolution Algorithm”, The Eighth International Conference on Future Computational Technologies and Applications, March 20-24, 2016, Rome, Italy.

Using BP Neural Network for Adapting Playout Time in Communication Networks

Sara Helmi^{1,2}, Niveen Mohamed Badra², Mohamed Elkattan³

¹Department of Basic Science, Faculty of Engineering, British University in Egypt, Cairo, Egypt
 e-mail: Sara.helmy@bue.edu.eg

²Department of Engineering Physics and Mathematics, Faculty of Engineering, Ain Shams University, Cairo, Egypt
 e-mail: niveen_badra@eng.asu.edu.eg

³Nuclear Materials Authority, Cairo, Egypt
 e-mail: emtiazegf@hotmail.com

Abstract-New multimedia applications have a critical requirement on jitter. Network jitter is a serious problem in communication networks, especially in Voice over IP networks (VOIP). One of the proposed solutions to minimize jitter is to adapt the playout time. In our paper, we introduce an adaptive approach using Back Propagation (BP) neural network to identify the jitter and adjust the playout time according to several network conditions. The algorithm was tested using k-fold cross validation and the results show that the algorithm can achieve promising results under different delay conditions.

Keywords-Jitter, playout time, optimization, Neural Network, Back-propagation neural network, k-fold cross validation

I. INTRODUCTION

In VOIP communication, several requirements for Quality of Service (QOS) affect the speech quality [1][2]. There are three main performance aspects that characterize the quality of voice in communication networks over the Internet. The first aspect is the end-to-end delay, which is the time it takes for a packet to be transmitted across a network from source to destination. Acceptable end-to-end delay values are less than 100 ms for the one way delay [3][4]. The second aspect is the packet loss, as described in [5]. The acceptable range of the voice quality is when the packet loss is less than 2%. The last aspect is the delay jitter. It is defined as the difference in end-to-end one-way delay between selected packets in a flow with any lost packets being ignored.

Delay jitter is the result of network congestion and improper queuing is the delay jitter. In the sending host, the voice packets are transmitted at a steady rate, but packets are received at a disparate rate [6]. To be able to recover the initial steady rate, the played out packets should be at a steady rate. When the jitter is large, dropping of the delayed packets can occur, and that will lead to obvious audible gaps. The sense of hearing in humans is highly sensitive for short audio gaps. That is the reason behind keeping the jitter in to a minimum value (less than 30 ms). Other elements mentioned in [7] state that the acceptable jitter can be between 30 ms and 75 ms.

In VOIP applications, the mechanism that is used to make the rate of output packets constant is the play out

buffer, also known as the jitter buffer. The jitter buffer holds the late packets and then plays them out at a steady rate. There are two kinds of jitter buffers: static buffers and dynamic buffers [8]. An adaptive jitter buffer is the most applicable mechanism, because it uses several strategies in the sender and the receiver hosts. When network conditions are perfect (the variation in interspace delay is almost 0 ms), it adjusts to be at a minimum value to reduce latency. On the other hand, when network conditions are very hard (a high transient jitter and packet loss exist), it adjusts itself to a higher value. However, this has the side effect of increasing the latency [9].

Many other works have been proposed to improve jitter. In [10], Artificial Neural Network (ANN) was combined with a standard multi-layer perceptron (MLP) and a recurrent-MLP [11] and Wavelet Packet-MLP (WP-MLP) [12]. Our new approach is combining ANN with the algorithms Exponential-Average, Fast-Exponential-Average and Min-Delay [13]. We use back propagation (bp) neural network on the three algorithms and we train the (bp) neural network to choose which algorithm is suitable to be run in different network conditions. This offered a flexibility in the network and a wide range of conditions that the network can deal with. We will focus on optimization of the playout time to achieve the best trade-off between latency and dropping of a voice packet.

The following sub-section presents the background for the problem formulation in Section II. It is the basic concept behind sending packets with time i from a sender to a receiver [13]. Fig. 1 shows a schematic diagram that illustrates the packet timing from sender to receiver. Table 1 shows the description of each parameter in "Fig. 1".

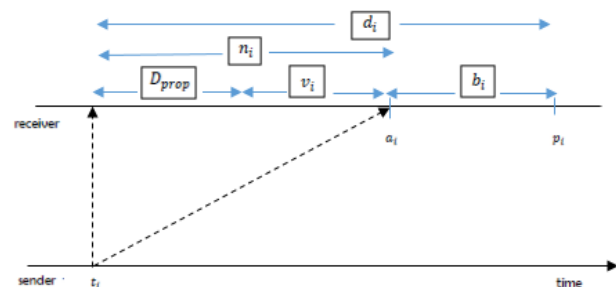


Figure 1. Packet i timing from sender to receiver [13]

TABLE 1. NOTATION PACKET i TIMING

Notation	Description
t_i	The time of sending i th packet.
a_i	The received time of packet i at the receiver. $a_i = t_i + n_i$
p_i	The played out time for packet i at the receiver.
D_{prop}	The delay of propagation from the sender to the receiver.
v_i	The delay of packet i from the sent to the destination.
b_i	buffer delay for packet i $b_i = p_i - a_i$
d_i	The time elapsed from transmitting from the source until it is played out at destination, (playout delay).
n_i	Packet network delay. $n_i = D_{prop} + \hat{v}_i$

In [13], three adaptive playout delay adjustment algorithms were defined:

- Algorithm 1 (Exponential-Average): This algorithm uses the mean \hat{d}_i to estimate the playout delay \hat{p}_i of its arriving packet so that:

$$\hat{d}_i = a * \hat{d}_{i-1} + (1 - a) * n_i \quad (1)$$

- Algorithm 2 (Fast-Exponential-Average): The only difference between this algorithm and the previous one is the new condition when the network delay n_i is larger than \hat{d}_{i-1} , β was chosen as 0.75 as in [13].

$$\begin{aligned} & \text{if } (n_i > \hat{d}_i) \\ & \quad \hat{d}_i = B * \hat{d}_i + (1 - B) * n_i \\ & \text{else} \\ & \quad \hat{d}_i = a * \hat{d}_i + (1 - a) * n_i \\ & \text{end} \end{aligned} \quad (2)$$

- Algorithm 3 (Min-Delay): The proposal of this algorithm is to minimize the delays. It uses the minimum delay of all packets received in the current talkspurt as \hat{d}_i to

predict the next talkspurt playout delay. Let \hat{S}_i be the set of all packets received in a single talkspurt $\hat{d}_i = \min_{j \in S_i} \{n_i\}$

$$\begin{aligned} & \text{if } (n_i < n_{i-1}) \\ & \quad \hat{d}_i = n_i \\ & \text{else} \\ & \quad \hat{d}_i = n_{i-1} \\ & \text{End} \end{aligned} \quad (3)$$

The remainder of this paper is structured as follows. Section II provides a description and overview of the problem formulation. Section III provides our modeling approaches. Section IV provides the optimization approach with back propagation (NN). Section V presents the validation test using k-fold cross validation and the results, and, finally, Section VI gives the conclusions.

II. PROBLEM FORMULATION

If the first packet i in a talkspurt is played out, then its calculation will be as below, where \hat{d}_i is the mean and \hat{v}_i is the variation in the end-to-end delay:

$$p_i = t_i + \hat{d}_i + \mu * \hat{v}_i \quad (4)$$

In this paper, we use a uniform distribution for random variance \hat{v}_i from 0 to 10 in normal conditions because the equation of variance [13] cancels the delay in the network, which is not realistic in a real network communication. A certain value 0.998002 of the weighting factor α was mentioned in [13] and, after applying some statistical analysis with small differences from the specific value 0.998002, α was settled as a value varying between $0.01 \leq \alpha \leq 0.998002$. The playout time for any posterior packet j in a talk spurt is computed as:

$$p_j = p_i + t_j - t_i \quad (5)$$

The term μ in the playout time is used to keep the playout time beyond the delay estimate so that only a small number of the packets in the receiver will be lost due to late arrivals [14]. In this paper, we set a range of values $1 \leq \mu \leq 20$. Also, we considered that, if the packet i arrived at the receiving host at the same time or after its playout time, the packet i will be played out:

$$\begin{aligned} & \text{if } (p_i \leq a_i) \\ & \quad p_i = a_i \end{aligned} \quad (6)$$

TABLE 2. NORMAL CONDITIONS OF NETWORKS ($\times 10^{-15}$)

α	0.01		0.26		0.51		0.76		0.875		0.9		0.925		0.95		0.998002	
d_1	$\mu = 10$	7.4	$\mu = 12$	8.9	$\mu = 6$	7.3	$\mu = 14$	8.3	$\mu = 12$	8.3	$\mu = 1$	9.1	$\mu = 17$	8.4	$\mu = 3$	9.7	$\mu = 1$	8.4
d_2	$\mu = 4$	8.3	$\mu = 9$	9.4	$\mu = 15$	9.1	$\mu = 6$	7.4	$\mu = 18$	8.4	$\mu = 18$	8.3	$\mu = 7$	8.7	$\mu = 2$	9.1	$\mu = 9$	7.2
d_3	$\mu = 14$	7.6	$\mu = 6$	9.6	$\mu = 18$	7.6	$\mu = 16$	7.4	$\mu = 12$	8.3	$\mu = 16$	8.9	$\mu = 6$	8.6	$\mu = 6$	7.4	$\mu = 2$	8.7
min	$\mu = 10$	7.4	$\mu = 12$	8.9	$\mu = 6$	7.3	$\mu = 6$	7.4	$\mu = 12$	8.3	$\mu = 18$	8.3	$\mu = 17$	8.4	$\mu = 6$	7.4	$\mu = 9$	7.2
	d1		d1		d1		d2		d1		d2		d1		d3		d2	

A new formula is proposed to calculate the old jitter at a_i by using:

$$|a_j - a_1| - NT \tag{7}$$

And the new jitter at p_i by using

$$|p_j - p_1| - NT \tag{8}$$

N: 1, 2, 3, 4...

We assumed that the packets will be sent at a constant rate. So, T refers to a constant time for sending the packets at the sending host, which will be 1.2.

To evaluate the best parameters for different network conditions, we set 100 packets with eight conversations, each conversation including 5 to 20 packets. Under normal condition of the network where variance changes from 0 to 10, and the initial value for the playout delay d_0 will be 29, and propagation delay 20. The mean and standard deviation for the proposed formula (8) have been calculated in the receiving host. After that, we evaluate the performance of the three algorithms d_1, d_2, d_3 in normal conditions. With variance (v), propagation delay (D_{prop}) and initial value where used, for each α value we changed μ from 1 to 20. With total simulation trials of 27, and after conducting those trials we choose the minimum mean for the calculated function and identify the corresponding μ to it.

III. IMPLEMENTATION

Our approach for changing the network conditions is summarized in five stages, as follows:

Stage 1. Comparing the results of the three adaptive algorithms under normal conditions (i.e. d_1, d_2, d_3 respectively) and selecting the minimum jitter value in the receiving host with the corresponding μ (see Table 2). The last row in Table 2 shows the selected algorithm for each set of trials that gives the minimum playout value between the three algorithms while comparing over a certain α .

Stage 2. Repeating this process for the three algorithms will be done now under different network conditions, first,

by varying the variance from 0 to 15, 0 to 20, and 0 to 25 respectively. This leads us to a really hard condition in queuing delay. Different results are shown for algorithms that give the minimum values.

Stage 3. Inserting a different value for the propagation delay 40, 60, and 80 to create a new set of network simulations.

Stage 4. Varying the initial value for playout delay d_0 from 50, 75, and 100.

Stage 5. In the final stage, we vary the variance v , mean d_0 and propagation delay D_{prop} simultaneously to evaluate the algorithms in terms of achieving the minimum jitter in the receiving host under rough conditions. The previous five stages results are summarized in Table 3.

TABLE 3. INPUTS AND OUTPUTS RESULTS

No. training set	Inputs			Outputs		
	v	D_{prop}	d_0	α	μ	Algorithm
1	0 to 10	20	29	0.998002	9	d2
2	0 to 15	20	29	0.875	18	d2
3	0 to 20	20	29	0.9	18	d2
4	0 to 25	20	29	0.26	6	d3
5	0 to 10	40	29	0.01	4	d2
6	0 to 10	60	29	0.01	4	d2
7	0 to 10	80	29	0.9	16	d3
8	0 to 10	20	50	0.76	16	d3
9	0 to 10	20	75	0.76	16	d3
10	0 to 10	20	100	0.76	16	d3
11	0 to 15	40	29	0.925	7	d2
12	0 to 15	60	29	0.9	16	d3
13	0 to 15	80	29	0.51	18	d3
14	0 to 20	40	29	0.998002	9	d2
15	0 to 20	60	29	0.925	7	d2
16	0 to 20	80	29	0.998002	2	d3
17	0 to 25	40	29	0.51	6	d1
18	0 to 25	60	29	0.76	14	d1
19	0 to 25	80	29	0.998002	2	d2

IV. OPTIMIZATION APPROACH

Neural networks are modeled on the mechanism of the brain [Kohen, 1989] [Hecht-Nielsen, 1990]. Neural networks can perform various duties like classification, identification, pattern recognition, control systems, speech and vision.

A supervised learning algorithm adjusts the strengths or weights of the inter-neuron connections according to the difference between the desired and actual outputs corresponding to a given input. Thus, supervised learning requires a "teacher" or "supervisor" to provide desired or target output signals.

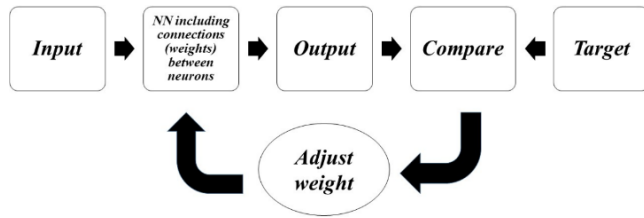


Figure 2. Training a neural network

The main objective of neural networks is to adjust a particular input to lead to a specific target output. This situation is illustrated in Fig. 2. A comparison has been adjusted in the network between outputs and targets, until the output matched the target in the network. In the proposed model, back propagation neural network (Rumelhart and McClelland, 1986) was used in Layered feed-forward ANNs, as in [15]. This means that the artificial neurons are organized in layers and send their signals "forward", and then the errors are propagated backwards. The network receives inputs by neurons in the input layer, and the output of the network is given by the neurons on an output layer. There may be one or more intermediate hidden layers, as in [15].

Neural Network Implementation

In our approach, we first conduct linear scaling of data (i.e., data normalization), which is important to put the data in interval between zero and one. This requires *mean* and *standard deviation* values associated with the facts for a single data input $x = (x - mean) / (standard deviation)$, and then the error (difference between actual and expected results) is calculated.

Our training data set in Table 3 consists of 19 input signals assigned with corresponding target (desired output). The neural network is then trained using back propagation algorithms. We separate the three desired outputs to construct three separate networks. We train the network with the 19 training data in each round a three input from d_0, D_{prop} will enter the three input neuron. Each BP neural network consists of input layer with three neurons, one hidden layer with five neurons, and output layer consist of one or three outputs based on the type of the network, two

of them with only one output and that will be for outputs α and μ in Table 3; see Fig. 3.

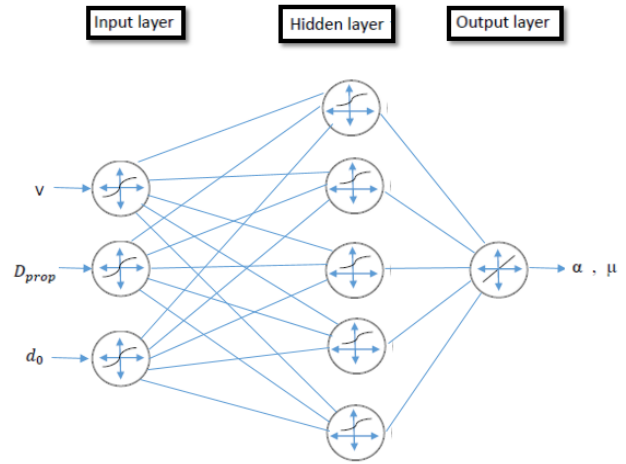


Figure 3. BP neural network for outputs α and μ

For the input layer, we use hyperbolic tangent sigmoid as a transfer function, which varies in the interval $[-1, 1]$. For the hidden layer, we used Log-sigmoid transfer functions to calculate a layer's output in the interval $[0, 1]$ from its net input. At the output layer, we used linear transfer function. Fig. 4 and Fig. 5 show the performance of the two BP neural networks after 1000 epoch. It can be seen that the curve converges to a minimum value after 1000 epoch.

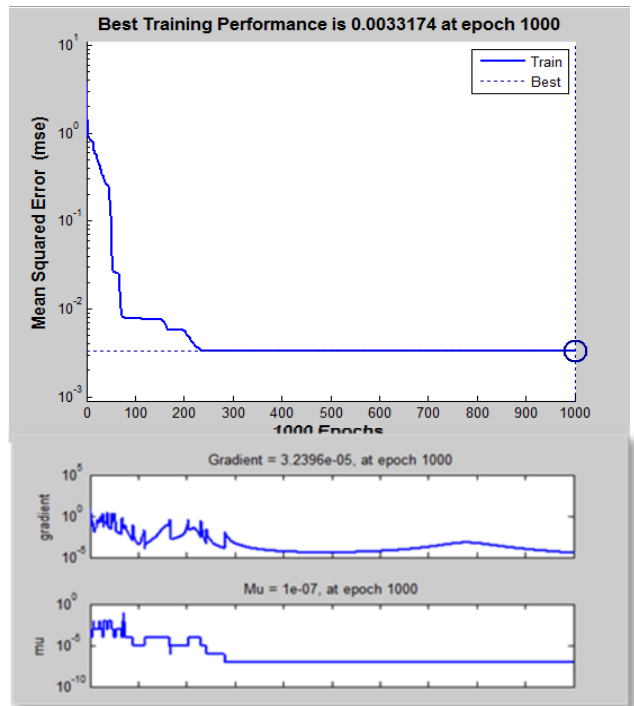


Figure 4. Performance of α

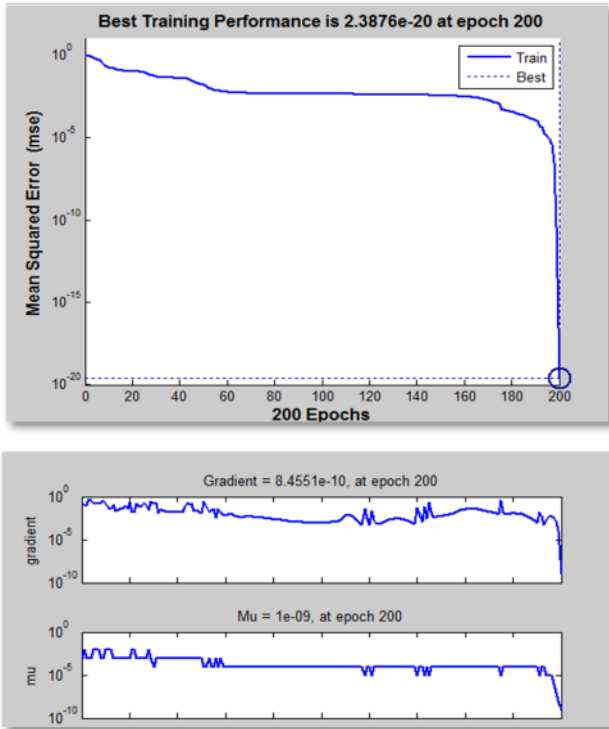


Figure 5. Performance of μ

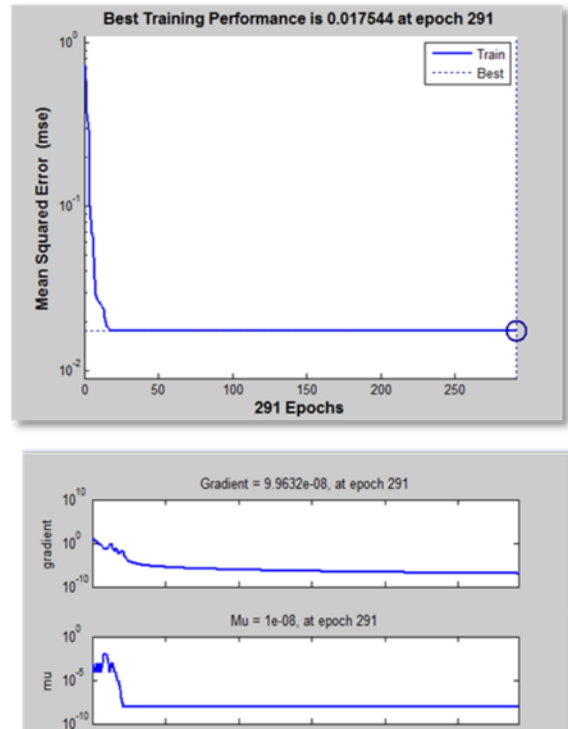


Figure 7. Performance of network algorithm

Classification

A classification problem occurs when an object needs to be assigned into a pre-defined group or class based on a number of observed attributes related to that object as in [16] [17]. Neural networks (NNs) are popular classification algorithms in computer-aided diagnosis because of their ability to “learn” classification rules from a set of training data (see Fig. 6).

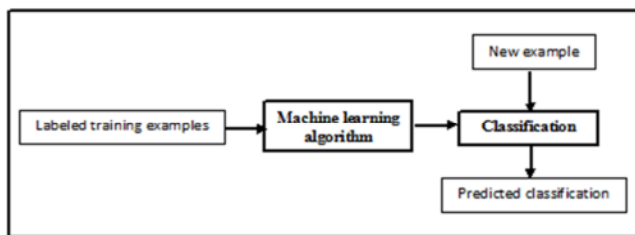


Figure 6. Classification in machine learning

The network with output algorithm is used to calculate the payout time related to the classification problem. We used a three-output network to sense the classification in d_1, d_2, d_3 . Fig. 7 shows the performance out of 1000 epoch. It can be seen that neural network succeeds in classifying each one to the appropriate algorithms.

Now, let us test and evaluate our model using the k -fold cross validation method. The test will give us a really good overview of the model performance.

V. K-FOLD CROSS VALIDATION

To be able to build a strong and useful machine learning solution, we need suitable analytical tools for evaluating the performance of our system. We can calculate the prediction errors (differences between the actual response values and the predictions) and summarize the predictive ability of the model by the mean squared prediction error (MSPE). This gives an indication of how well the model will predict in the future. Sometimes, the MSPE is rescaled to provide a cross-validation R^2 . However, most of the time, we cannot obtain new independent data to validate our model. An alternative is to partition the sample data into a training (or model-building) set, which we can use to develop the model, and a validation (or prediction) set, which is used to evaluate the predictive ability of the model. This is called cross-validation.

Here, we will clarify the reasons why k -fold cross validation is used in neural networks. The main goal of the classification problem is to find a group of weights and bias values that will lead to a perfect match between the targets and the output values, as in [18] [19].

Our approach would use all of the 19 data items to train the neural network. However, this approach would find weights and bias values that match the target value extremely, in fact, probably with 100 % accuracy, but when presented with a previously unseen set of input data, the neural network would fail to predict well. This case is called over-fitting. To overcome this problem (over-fitting), we use the k -fold cross validation. The idea behind k -fold

cross validation is to randomly sort your data and divide your data into k equally-sized sets, unless the data cannot be divided equally, like in our example. Each set is used one time as the test set while the rest of the data is used as the training set. The learner trains for k rounds, each round using one of the sets as the validation set and the remaining sets as the training set. We measure its accuracy on the validation set. Then, we average the accuracy over the k rounds to get a final cross validation accuracy, as in [20]. We used a 4-fold cross validation three folds with size 4 and the last fold sized 7, as shown in Fig. 8.

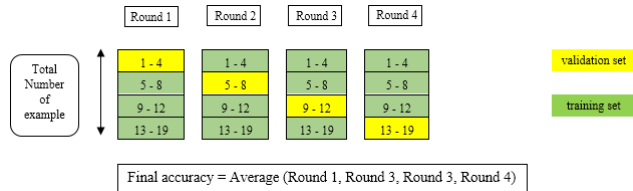


Figure 8. 4-fold cross validation

By using $E = \frac{\sum_{i=1}^k n_i}{m}$ (9)

where,

E Accuracy for the neural network.

n_i The number of examples in Fold i that were Correctly classified.

m Number of examples which is equal 19.

the calculated accuracy for the neural network with output μ is 0.26789 and for the neural network with output α is 0.315789; the neural network accuracy for the output algorithm is 0.263157.

Repeated k-fold cross validation

It will not be effective to take the mean of 4 samples. On the other hand, splitting our sample into more than 4 would greatly reduce the stability of the estimates from each cross validation. A way around this is to do repeated k -folds cross validation. To do this, we simply repeat the k -fold cross validation three times, each time with a different random arrangement and take the mean of this estimate. An advantage of this approach is that we can also get an estimate of the precision of this out-of-sample accuracy by creating a confidence interval. We will do three replications so we end up with a nice round: 12 out-of-sample accuracy estimates. By using the following formulas:

$$e = \frac{\sum_{j=1}^t E_j}{t}, v = \frac{\sum_{j=1}^t (E_j - e)^2}{t-1}, \sigma = \sqrt{v} \quad (10)$$

where,

e Mean of the three runs for each one of the three networks.

$E_1 \dots E_t$ Accuracy estimates obtained in t runs, which is in our case 3 runs.

v Variance (variability it shows for different samples)

σ Standard deviation (How much it deviates from the true value).

This time, we get an estimate for network μ , α and algorithm, which is quite close to our estimate from a single k -fold cross validation. We obtained the mean, variance, and standard deviation of the 3 runs for network μ , α and algorithm in Table 4.

TABLE 4. ACCUARACY IN THREE REPLICATIONS

Neural Network	Mean (e)	variance (v)	Standard deviation (σ)
Network μ	0.0877	6.463×10^{-3}	0.080393
Network α	0.3333	6.463×10^{-3}	0.080395
Network Algorithm	0.3157	2.769×10^{-3}	0.0526

The standard procedure for training a neural network involves training on the complete database by minimizing the accumulated misclassification of inputs in the dataset. Since the overall goal is not to minimize errors on the dataset, but rather to minimize misclassification on a much larger set of conceivable inputs, cross-validation gives a much better measure of expected ability to generalize. The estimate for the algorithm performance has an error of 0.0877 in network μ , 0.3333 in network α and 0.3157 in network algorithm, with standard-deviation of 0.080393 in network μ , 0.080395 in network α and 0.0526 in network algorithm. The validation error gives an unbiased estimate of the predictive power of a model. So, after comparing the mean of the error (e) of the three networks with the actual error from training the 19 data sets in the network, we conclude that more training samples need to be used to obtain better results and provide a good network performance, and this can be achieved by monitoring a real voice over IP network.

VI. CONCLUSION

In this paper, we have investigated the performance of our algorithm under different conditions of the communication network, starting from the normal conditions to congested conditions which have high queuing, propagation delay, and playout delay for adapting the buffering of packets at the receiver. The main objective is to build an algorithm that can transact with all these conditions and to keep the playout delay as small as possible. We proposed an optimization method for optimizing the playout delay of packets. Using back propagation neural network, the results were promising in dealing with the input data. In our future work, we will conduct further investigation of more realistic conditions, such as larger data sets with more complicated distributions.

REFERENCES

- [1] J. Zhang, H. J. Kim and D. H. Ahn, "Analysis of Streaming Service Quality Using Data Analytics of Network Parameters, " Proc. 2012 DATA ANALYTICS 2012 : The First International Conference on Data Analytics
- [2] O. Obafemi, T. Gyires and Y. Tang, " An Analytic and Experimental Study on the Impact of Jitter Playout Buffer on the E-model in VoIP Quality Measurement , " Proc. 2011 ICN 2011 :The Tenth International Conference on Networks.
- [3] E. TIPHON, "End-to-End Quality of Service in TIPHON Systems,"Proc. 2000 Definition of Quality of Service (QoS) Classes,Vols. Part 2: 101 329-2.
- [4] Dr. H. A. Mohammed, Dr. A. H. Ali and H. J. Mohammed, "The Affects of Different Queuing Algorithms within the Router on QoS VoIP application Using OPNET," Proc. 2013 International Journal of Computer Networks & Communications (IJCNC) Vol.5, No.1
- [5] T. Uhl, "QUALITY of service in VOIP communication," Proc. 2004 Int.J.Electron.Commun 58 (3), pp. 178-182.
- [6] H. Dahmouni, A. Girard and B. Sansò, "An analytical model for jitter in IP networks," Ann. Telecommun. Proc. 2012 67:81–90.
- [7] E. DTR /TIPHON, "Telecommunication and Internet Protocol Harmonization Over Networks (TIPHON)," General Aspects of Quality of Service (QOS), Proc. 1998 vol. tr 101 329 , no. Ver. 1.2.5
- [8] R. J. B. Reynolds, A. W. Rix, 2001"Quality VOIP - an engineering challenge," Proc. 2001 BT Technol Journal Vol.19 Issue:2 pp. 23-32.
- [9] I. Klimek, M. Čajkovský and F.Jakab, "Novel methods of utilizing Jitter for Network Congestion Control," Proc. 2013 Acta Informatica Pragensia, 2(2), 1–24, DOI: 10.18267/j.aip.20
- [10] Y. Zhang, D. Fay and L. Kilmartin, "An application of neural networks to adaptive playout delay in VoIP," Proc. 2007 Ireland Conference on Information and Communication Technologies
- [11] S. Haykin, "Neural Networks: A Comprehensive Foundation," Proc. 1998 Prentice. Hall, Upper Saddle River, NJ. ISBN:0132733501
- [12] Q. Zhang and A. Benveniste, Wavelet networks. IEEE Trans. Neural Networks, Proc. 1992 3:889–898
- [13] R. Ramjee, J. Kurose, D. Towsley and H. Schulzrinne, "Adaptive playout mechanism for packetized audio application in wide-area networks," Proc. 1994 Proseeding of IEEE infocom, vol. 2, no. ISBN: 0-8186-5570-4, pp. 680-688.
- [14] V. Jacobson, "Congestion avoidance and control," Proc.1988 ACM SIGCOMM Conf Stanford, pp. 314-329, August, 1988.
- [15] Z. Reitermanoy, "Feedforward Neural Networks - Architecture Optimization and Knowledge Extraction," Proc.2008 WDS'08 Proceedings of Contriuted Papers, Part 1, pp. 159-164.
- [16] R. P. Lippmann, "Pattern classification using neural networks," Proc. 1989 IEEE Communications Magazine, vol.27, no. Issue: 11.
- [17] Y. L. Murphey and G. Ou, "Multiclass pattern classification using neural networks," Proc. 2007 science direct journal Pattern Recognition, vol.40, Issue:1 pp. 4-18.
- [18] S. Haykin, "Neural Network sand Learning machines,"Proc. 2009.
- [19] A. Perez, J. A. Lozano and J. D. Rodriguez, "Sensitivity Analysis of K-Fold Cross Validation in Prediction Error Estimation," Proc.2010 IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.32, Issue: 3 pp. 569-575.
- [20] T. Fushiki, "Estimation of prediction error by using K-fold cross-validation," Proc. 2009 Springer Science Business Media, LLC 2009.

Electricity Price Forecasting in a Smart Grid

Md Mahbubur Rahman Khan, Md Minhaz Chowdhury and Kendall E. Nygard

Department of Computer Science

North Dakota State University

Fargo, ND, USA

Email: Mahbuburrahman.Khan@ndsu.edu, Md.Chowdhury@ndsu.edu, Kendall.Nygard@ndsu.edu

Abstract— A smart grid refers to a digitized and intelligently controlled electrical power system. Intelligent monitoring and communication of digital information can support two-way between consumers and providers. The work of this paper concerns modeling of dynamic pricing, potentially helping to improve efficiency of electricity consumption and delivery. A simple variant of collaborative filtering is applied for dynamically predicting prices. Information on power consumption periods, and history of purchase levels and prices are used as input. The collaborative filtering approach is compared with a naïve forecasting method and the Winter method for incorporating seasonality. Actual price data is available for use in validating the models, which reveals that the collaborative filtering method provides the best results.

Keywords- Collaborative Filtering; Winter Method; Dynamic Pricing; Smart Grid; Price Forecasting.

I. INTRODUCTION

Collaborative filtering (CF) is becoming a popular technique of filtering for information or patterns using collaboration among multiple data sources, viewpoints, and agents, for large data sets. Collaborative filtering techniques have been applied in applications such as analyses of environmental data from multiple sensors, google news recommendations [1], Netflix movie recommendations, personalized pricing recommendations [2], financial data that integrates multiple financial sources, and electronic commerce user data. In this work, we apply a simple variant of collaborative filtering for prediction of electricity prices. The application environment is that of a fully instrumented and networked smart grid.

A smart grid is a digitally enabled electrical grid that gathers, distributes, and acts on information across energy suppliers and consumers within grid infrastructure. A smart grid holds promise to improve the efficiency, reliability and sustainability of electricity services [3]. A smart grid opens opportunity for changing the traditional electric pricing system that is typically based on peak and off peak hourly rates. The existing price model for electricity hides the temporal deviation in the cost of electricity depicted in the pattern of consumption of electricity during peak times and off peak times.

In this paper, collaborative filtering is applied to forecast prices of electricity consumed across regions. Our test day is from the ten regions of the New York Independent System Operator (NYISO). The effectiveness of applying collaborative filtering is compared with Winter Method of seasonality and a basic forecasting model.

The remainder of this paper is organized as follows: Section 2 describes the related works. Section 3 describes current electricity market. Section 4 describes steps related to data used in this work e.g. data source and data pre-processing steps. Section 5 describes the three approaches used in this work, for predicting the price of electricity. This section also describes dynamic pricing input output model. Section 6 describes the results of these predictions. Section 7 concludes this paper by mentioning limitations, suggests future use and possible improvement of collaboration filtering.

II. LITERATURE REVIEW

Collaborative filtering is based upon patterns that can be identified with limited details concerning the items or users under analyses. For the grid, the patterns can be such things as ratings, usage or purchases [4]. The idea is that a selective group of consumers of the same service shares a similar opinion with judgments based on their personal preferences.

Filtering proceeds by matching the available information from a domain in which information shares certain similarities in nature. For example, the movie recommendation system Netflix uses collaborative filtering that finds people with similar tastes in movies, called nearest-neighbors. Based on their history of movie ratings, the collaborative filtering approach recommends movies and predicts the rating for a movie. There are recursive application of the filtering for predicting such neighbors that is more effective in long term reusability [5].

Collaborative filtering falls into the general category of Recommender Systems [6]. Content-based information filtering is effective in identifying items similar to that a consumer preferred by analyzing textual similarity from user data [7]. However, the work presented in this paper focuses on finding similarities between customers pattern of electricity consumption that cannot be identified by keyword based searching [1]. Hence, the approach used for price forecasting applies collaborative filtering more broadly over content-based information.

There are several state-of-the-art works dealing with prediction of prices. For example, one recent study predicts hourly day-ahead electricity prices are using features like long memory, positive and negative price spikes, and seasonality [8]. However, this method is weak in capturing nonlinear patterns of price.

Other reported works uses approaches such as Auto-Regressive Fractionally Integrated Moving Averages with Feedforward Neural Networks [9], 3-Regime Markov Regime-Switching [10], and Hodrick–Prescott filters [11]. The work reported in [12] summarizes multiple methods. Concerning seasonality, there has been work reported on ways to identify and model seasonality [12]. As a benchmark seasonality method, the Winter method is applied in this work [13].

III. ELECTRICITY MARKET

Consumption levels for electricity increases during the working day, peaks in the late afternoon or early evening, and is at a low point by midnight [14]. The graph in Fig. 1 and Fig. 2, shows average consumption per hour, not instantaneous power. One interesting point is that the required maximum power is nearly twice as high as the average power consumption [15]. In much of North America, the problem is especially pronounced during the top 60 to 100 hours of the year, which may account for as much as 10–18 percent of the system peak load [16].

The yearly consumption increases in summer, when people use air conditioners. Figure 3 shows that in the summer, the power requirement doubles for several days [17]. To meet this critical peak load, expensive combustion turbines are purchased and installed, which raises rates for all customers. A prediction model predicting real time price of electricity can address this problem and increase economic efficiency.

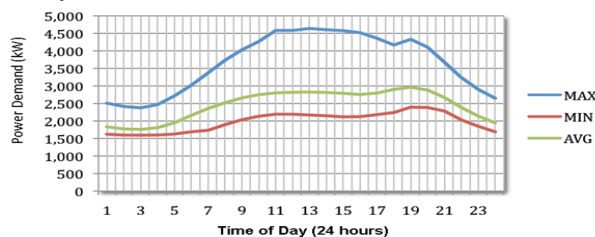


Figure 1. Daily electricity consumption pattern in San Diego

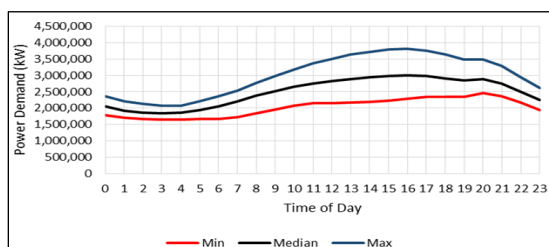


Figure 2. San Diego dynamic load profile from June to August 2014

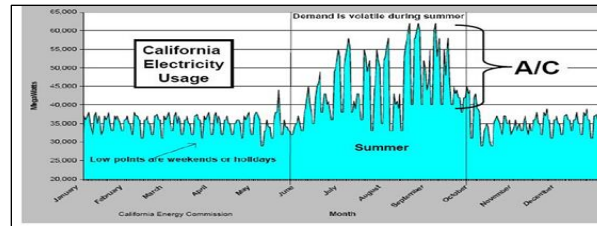


Figure 3. Yearly electricity consumption pattern in San Diego

Dynamic pricing that is aligned with demand response can reduce the maximum annual peak load to save investment in expensive large power plants. Large investments in generation capacity by power companies propagate to increase prices of electricity. Power plant reliability and the uncertainty of not achieving maximum utilization of resources increases supply side volatility and contributes to increasing the price of electricity. Conceptually, the dynamic pricing model for electricity is a mechanism that mitigates uncertainties in the electric grid by reacting to real-time fluctuations. The real time price reflects the capacity of the power generation system through price sensitive demand and supply. Design of a sustainable model that reflects consumer preferences, behavior and response is a challenge for researchers in modeling supply side uncertainties.

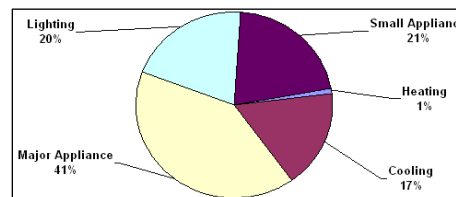


Figure 4. San Diego county average household electricity consumption: 6,300kw

These preferences, behavior and response of consumer can improve the use of electricity consumed. For example, Fig. 4 shows that a household consumes 62% of electricity for appliances and 20% for lighting. This 82% could be controlled and scheduled to use by making the consumer responsive to the price of the electricity [17].

IV. ELECTRICITY PRICE DATA

A. Consumed electricity price Data

NYISO is a not-for-profit organization based on New York’s Capital Region to govern the New York’s electricity market. It administers and monitors the wholesale electricity market, conducts planning, assesses long term projects and develops and deploys state-of-the-art technology for a sustainable and efficient power grid in the New York State. The NYISO publishes the wholesale price of consumed electricity every day on their website. The data (Table 1) used in this research is from that published data. The total area of New York is divided into 15 regions, each region is addressed as one node in this paper. For the analysis in this paper, data

for 10 nodes are used. However, we could have done with 15 nodes as there are 15 regions in New York electricity market.

TABLE I. PRICE OF ELECTRICITY FOR 10 NODES IN NEW YORK

ID	Date	Node ID	HR00	HR01	HR02	HR03	HR04	HR23
000001	6/27/2011	61757	23.97	35.84	35.36	10.78	14.54	46.75
000002	6/27/2011	61754	22.96	34.34	33.87	10.35	13.86	45.28
000003	6/27/2011	61760	24.75	36.9	36.33	11.02	14.87	48.95
000004	6/27/2011	61753	22.32	33.2	32.71	10	13.22	43.68
000005	6/27/2011	61844	22.72	34.01	33.54	10.24	13.77	44.79
000006	6/26/2011	61757	37.64	35.56	35.07	27.83	12.96	31.91
000007	6/26/2011	61754	36.63	34.64	34.14	27.09	12.58	31.07
000008	6/26/2011	61760	42.27	44.38	46.79	46.68	35.5	98.26
.....

B. Formatting extracted electricity price Data

The price of electricity is published for each node every hour. A day is divided into K time slots within a range of 0 to 24.

V. THE DEVELOPMENT ENVIRONMENT

The simulation is developed by in Visual Studio 2010 using C# as a programming language. The reason for choosing C# as a programming language is to benefit from powerful .NET framework. The Visual Studio 2010 makes it simple and quick to develop and deploy a software project. Two Graphical User Interfaces (GUI) are used in this software. The Window Forms Designer provides the flexibility to control the layout that houses controls (textbox, label, list box, etc.). The Windows Presentation Foundation (WPF) helps to control the GUI by event driven programming and the Extensible Application Markup Language (XAML) file. For simplicity and better visualization, Microsoft Excel 2010 is used to hold the raw data. This provides quicker processing of data as the National Grid demand data is published in Microsoft Excel format.

VI. METHODS FOR PREDICTING PRICE OF ELECTRICITY

In this research, three methods for predicting the price of electricity are applied. They are based on history (basic or naive forecasting method), collaborative filtering and the Winter method for seasonality. The Winter Method is considered to be a difficult competent of collaborative filtering. This is because of the capacity for capture the variation of price throughout the day.

A. Basic or naive forecasting method

For time series data, naive forecasting is the simplest way to forecast by making forecasted value equal to the last observed value. It is easy to use Naive Forecast and it can

handle seasonality effect. However, if there is an unusual change in the last period, this method will produce significantly inaccurate results.

B. Collaborative Filtering as dynamic pricing model

In dynamic pricing model, the price of electricity will depend on demand, and the demand will in turn depend on several qualitative and quantitative variables, such as temperature, number of appliances, user sleeping times, and user consumption preferences. Consumer behavior in each variable is not known and response could vary by clustering one or more variables together. One of the benefits of collaborative filtering is that it models the behavior of consumers based on their response without such complex details. By complex details it means dealing with the mentioned variables. Also, this collaborating filtering approach works well for predicting user recommendation for movie ratings [1]. Hence the use of collaborative filtering in predicting electricity price is beneficial in dynamic price domain.

Now, in the dynamic pricing model, consumers respond to the real-time price of electricity [18]. To calculate the dynamic price, the demand and supply must be forecasted, including information regarding generation capacity for fulfilling unexpected high demand (Fig. 5). The forecasted demand is based on user categories, such as household, commercial, and industrial. The user utility function is needed to provide a smooth estimate. The utility function considers the user level of satisfaction and behavioral patterns. By using collaborative filtering, user patterns are reflected in the choices they have made in their consumption of electricity. The output from the dynamic pricing model is the price of electricity for each group of users as well as load per power generator.

Classical time-window or instance-decay approaches are inappropriate in such scenario, as they lose signals when discarding data instances [19]. The factor and neighborhood models can be merged smoothly to predict more accurately [20]. Researchers at Yahoo applied collaborative filtering with bilinear predictive model for many of their predictions [21].

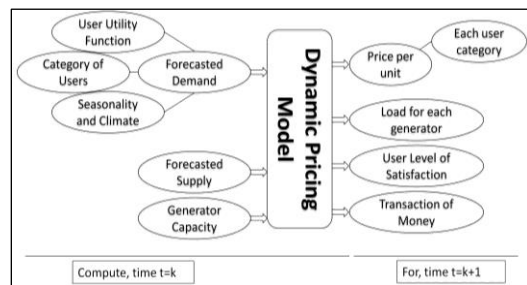


Figure 5. Dynamic pricing input output model

In the collaborative filtering approach a similar node is selected by nearest neighborhood search. The selection of

nearest neighbor is done by calculating mean square deviation (MSD).

Each node in the previous time (yesterday) has its price for 24 hours (0...23). By applying the MSD calculation shown below, a Node is selected to have the closest behavior of the node for which price is need to be forecasted (Table 2).

TABLE II. MSD CALCULATION TO FIND THE BEST MATCH FOR EACH NODE

Node	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7
Node 1	-	0.707	983.944	5.224	28.275	6.043	5.378
Node 2	3.178	-	1027.595	2.098	26.340	10.811	9.924
Node 3	938.299	1027.595	-	1104.065	1101.345	867.937	873.540
Node 4	10.387	2.098	1104.065	-	26.946	22.311	21.032
Node 5	32.031	26.340	1101.345	26.946	-	41.974	41.010
Node 6	2.285	10.811	867.937	22.311	41.974	-	0.023
Node 7	1.891	9.924	873.540	21.032	41.010	0.023	-
Node 8	0.903	3.178	938.299	10.378	32.031	2.285	1.891
Node 9	1.907	9.949	873.232	21.058	41.130	0.023	0.00086
Node 10	57.256	76.715	684.736	96.495	116.437	45.648	46.483
Minimum	0.903	0.707	684.736	2.098	26.340	0.023	0.00086
Match Node	Node 8	Node 1	Node 10	Node 2	Node 2	Node 5	Node 9

The Price of the most similar node in the previous time is applied to provide the forecasted price for the current time.

C. The Winter Method for Seasonality

The Winter method of seasonality is applied to calculate the forecast, as shown below [16].

$$F(k) = \alpha \frac{A(k)}{C(k-K)} + (1 - \alpha)[F(k-1) + T(k-1)]$$

$$T(k) = \beta[F(k) - F(k-1)] + (1 - \beta)T(k-1)$$

$$C(k) = \delta \frac{A(k)}{F(k)} + (1 - \delta)C(k-K)$$

$$f(k + \tau) = [F(k) + \tau T(k)]C(k + \tau - K)$$

This method updates a smoothed estimate F(k), a smoothed trend T(k), a seasonal factor C(k) and compares with actual demand A(k). The forecast period, τ is used to forecast more than one period in the future. The first and second equation calculates the smoothed estimate and the smoothed trend respectively by using exponential smoothing with a linear trend. These two equations capture the linear trend over recent days and the trend during the last couple of hours in consideration. The factor of seasonality is incorporated in the first equation above, to get the data about last time's demand as C(k-K).

For example, considering a day as a domain, then K =24 (24 hours a day) and considering a year, K=12 (12 months in a year). The parameters α, β, and δ are smoothing constants between 0 and 1 either chosen or defined by the lowest mean square deviation (MSD) for the best performance in the test data. In this experiment, α = 0.10, β = 0.10, and δ = 0.10 are applied. Different combination of values of α, β and δ was applied to find the lowest RMS value and the 0.1 value for α, β and δ provides the lowest

RMS. That's why 0.1 was used. Other combinations of α, β and δ could be used but they won't be optimized.

In this work, a machine learning algorithm, to determine smoothing constants dynamically, has been applied.

The following equations give the formula for MSD and Root Mean Square (RMS) values with k = 1, 2...K.

$$MSD = \frac{\sum_1^K [f(k) - A(k)]^2}{K}$$

$$RMS = \sqrt{\frac{\sum_1^K [f(k) - A(k)]^2}{K}}$$

As a benchmark, a $\epsilon \in A$ and A is set of actual price of electricity used to measure the effectiveness of the forecasted price. Since price could vary hence is an element of A e.g. a. Also, $b \in B$ and B represents the price for the day before (Fig. 6). Again, $n \in N$ and N represents number of nodes which are regions in the study area (in this case New York). M represents matched node with lowest MSD with the node for consideration and W is the forecasted price by applying the Winter Method for seasonality. R is the MSD values calculated while searching for the match node. For comparing the MSD values among three forecasting method $s \in S$ is used. For applying the Winter Method for seasonality, $e \in E$ represents the smoothed estimate, $t \in T$ and T represents the smoothed trend which is the seasonality effect over the period of time, $c \in C$ represents the seasonal factor and $f \in F$ represents the forecasted price of electricity.

```

1 function bestForecastingMethod(α,β,δ,B,A)
2 {
3     S[0] ← getMSD(B,A); // basic forecasting
4     M←getMatchNode(B); // CF
5     S[1] ← getMSD(M,A);
6     //Winter Mthod
7     W ← forecastByWinterMethod(α,β,δ,B,A);
8     S[2] ← getMSD(W,A);
9     return forecasting method for lowest S;
10 }
11 function getMSD(A,B)
12 {
13     sumOfSquare ← 0;
14     for k:=0 to K
15         sumOfSquare +=(A[k]-B[k])2;
16     end for
17     return sumOfSquare/K;
18 }
19 function getMatchNode(B)
20 {
21     N ← getAllNodeExceptGivenNode(B);
22     for each n in N do
23         R ← getMSD(n,B);
24     end for
25     return n for min R;
26 }
27 function forecastByWinterMethod(α,β,δ,B,A)
28 {
29     E(K-1) ← ΣB/K;
30     for k:=0 to K
31         C(k) ← B(k)/E(K-1);
32     end for
33     for k:=K to 2K
34         E(k) ← αA(k-K)/C(k-K)+(1-α)[E(k-1)+T(k-1)];
35         T(k) ← β[E(k)-E(k-1)]+(1-β)T(k-1);
36         C(k) ← δA(k-K)/E(k)+(1-δ)C(k-K);
37         f(k+1) ← [E(k)+T(k)]*C(k+1-K)
38     end for
39     return f;
40 }

```

Figure 6. Algorithm for finding best forecasting method for forecasting the price of electricity.

In the algorithm for finding best forecasting method for forecasting the price of electricity (Fig. 6), best forecasting method is determined by finding the forecasting method with lowest MSD. Collaborative Filtering (CF) and the Winter Method is compared. For collaborative Filtering method, “getMatchNode” method by providing a Node to find out the best matching node with lowest MSD. Inside “getmatchNode” each node was compared except the provided node to find MSD by using “getMSD” method. The “getMSD” method takes two nodes calculate difference for all time period data available. The “forecastByWinterMethod” is used to forecast by using winter method.

VII. ANALYSIS OF RESULTS FROM FORECASTING METHODS

All three forecasting methods are applied in 10 Nodes. Table 3 shows results of calculation of MSD for three forecasting methods for node 1. The optimum column shows the optimum value of MSDs resulting from each of these methods. It shows that collaborative filtering (CF) provides lowest MSD for node 1. This table is summarized and then extended into Table 4.

TABLE III. CALCULATION OF MSD FOR FORECAST, CF AND WINTER METHOD FOR NODE 1

Method	HR00	HR01	HR02	HR03	HR22	HR23	MSD	Optimum
Actual	23.97	35.84	35.36	10.78	48.84	46.75	-	
Forecast	37.64	35.56	35.07	27.83	49.85	31.91	140.615	
CF	38.65	36.46	36.06	28.58	51.23	32.9	133.068	133.068
Winter	34.93	34.49	27.30	11.83	41.55	47.53	141.802	

Table 4 shows the listed MSDs of Table 3 for all 10 Nodes. This table 4 also shows that collaborative filtering gives the best results for 6 occurrences. The Winter Method shows best results in 3 out of 10 Nodes and for one Node the general forecasting method is best. The average MSD for CF is 148.99 (shown in the average row for CF MSD), which is about 39% of average MSD for the Winter Method. While calculating the Root Mean Square (RMS) value for CF, the average is 12.21 in a day. This means \$0.51 deviation for each forecast price while applying collaborative filtering. The Winter Method, the deviation of each forecast is \$0.815. Based on these results shown in Table 4, it can be concluded that collaborative filtering provides a better forecasting of the price of electricity than the Winter Method for Seasonality and basic forecasting.

TABLE IV. FINDING THE BEST FORECASTING METHOD FOR EACH NODE

Node	Best match	Matching MSD	Forecasting MSD	CF MSD	Winter Method	Optimum
Node 1	Node 8	0.9037	140.615	133.068	141.802	CF
Node 2	Node 1	0.7074	130.032	123.866	132.632	CF
Node 3	Node 10	684.736	631.198	173.702	401.333	CF
Node 4	Node 2	2.098	117.858	107.594	125.414	CF

Node 5	Node 2	26.345	121.381	93.238	53.802	Winter
Node 6	Node 5	0.0239	169.941	198.054	163.704	Winter
Node 7	Node 9	0.00086	3729.436	171.466	2191.85	CF
Node 8	Node 1	0.903	127.986	134.77	135.34	Forecast
Node 9	Node 7	0.00086	173.71	173.86	167.292	Winter
Node 10	Node 6	45.468	190.95	180.260	316.545	CF
Average		76.118	553.310	148.987	382.971	
Root Mean Square		8.724	23.522	12.206	19.569	

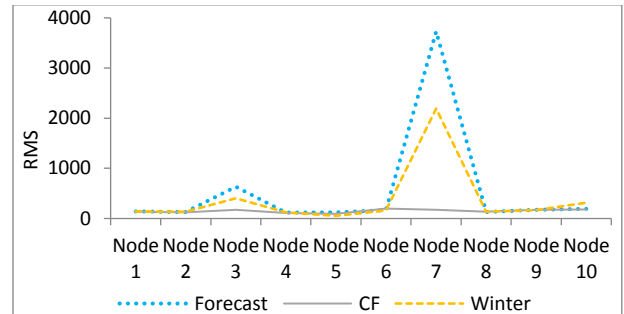


Figure 7. MSD for different forecasting method for each node

In Fig. 7, MSD for three forecasting methods are plotted. This figure shows that CF provides an excellent estimation of the price. Node 7 shows a significantly higher MSD. The data shows that the price of electricity in the day before was unusually high. By using collaborative filtering, such unusual behavior can be avoided.

To compare these methods, a program was developed, shown in fig. 6. Screen shots for the developed program are provided in Fig. 8, Figure 9, and Figure 10.

Figure 8. Data input for calculating forecast of the price of electricity

Winter Method		
Period	Actual	Forecast
0	23.97	34.9366983699871
1	35.84	34.4996548873408
2	35.36	27.3044892018985
3	10.78	11.837202123417
4	14.54	9.78411151497416
5	38.65	30.8825877460063
6	41.65	26.4313893444546
7	40.5	33.5539688705311
8	45.94	53.7992231383779
9	48.38	58.2835204960329
10	48.99	60.4010807619692
11	51.11	61.0556849620902
12	43.75	59.725917949658
13	44.56	58.779649401727
14	45.15	58.0590293831749
15	46.91	57.0781899187859
16	46.07	57.0110375309788
17	49.21	56.7428166958921
18	45.42	56.4727621607123
19	41.48	60.1340939180185
20	46.27	68.4245200757151
21	43.41	66.7525905858175
22	48.84	41.5583941159769
23	46.75	47.5378825215703

RMS value is: 141.802034344981]

Figure 9. Forecasted price by using winter method of seasonality

Forecast With Collaborative Filtering		
Period	Actual	Forecast
0	23.97	38.65
1	35.84	36.46
2	35.36	36.06
3	10.78	28.58
4	14.54	13.33
5	38.65	10.87
6	41.65	26.08
7	40.5	21.13
8	45.94	24.95
9	48.38	37.39
10	48.99	39.85
11	51.11	40.94
12	43.75	41.07
13	44.56	40.51
14	45.15	40.18
15	46.91	40
16	46.07	39.55
17	49.21	39.81
18	45.42	39.76
19	41.48	40.04
20	46.27	43.55
21	43.41	50.56
22	48.84	51.23
23	46.75	32.9

MSD value is: 133.0682

Figure 10. Forecasting by using collaborative filtering method for forecasting the price of electricity

VIII. CONCLUSION AND FUTURE WORK

Collaborative filtering is an effective method of predicting prices. However, there may be limitations related to covering the full factorial set of other possible factors that may influence price. Further testing is expected to include much larger data sets. Another limitation is that it does not considered price changes due such things as natural disasters, power outage for maintenance, transmission device failure, and generator scarcity. Working with such effect of disasters or their combination in pricing is of interest for future research. In this research, only one neighbor is selected (best match), i.e. $K=1$ for K -nearest neighbor (k -NN). Higher values of k could be applied to obtain more than one match. Finally, user preferences on the source of power (e.g., coal, nuclear, hydro, solar, wind) or an open market with pay as you go can be considered for future work.

REFERENCES

[1] Das, A. M., Google news personalization: scalable online collaborative filtering: industrial practice and experience, 2007.

[2] T. Kamishima, S. Akaho, "Personalized pricing recommender system", The 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems, pp 57-64, October 2011.

[3] W. Ketter, J. Collins, and C. Bloc, "Smart grid economics: policy guidance through competitive simulation", November 2010, ERIM Report Series Reference No. ERS-2010-043-LIS.

[4] Y. Koren, R. Bell, "Advances in collaborative filtering", recommender systems handbook, pp 77-118, 2015.

[5] J. Zhang, P. Pu, "A recursive prediction algorithm for collaborative filtering recommender system", The ACM conference on recommender systems, pp 57-64, 2007.

[6] F.O. Isinkaye, Y.O. Folajimi, and B.A. Ojokoh, "Recommendation systems: principles, methods and evaluation", Egyptian informatics journal, Volume 16, Issue 3, pp 261-273, November 2015.

[7] M. D. Ekstrand, J. T. Riedl, and J. A. Konstan, "Collaborative filtering recommender systems", Foundations and trends in human-computer interaction, Volume 4, Issue 2, pp 81-173, 2011.

[8] E. Fanone, A. Gamba, and M. Prokopczuk, "The case of negative day-ahead electricity prices", Energy economics, Volume 35, pp 22-34, 2013.

[9] C. Najeh, "A hybrid ARFIMA and neural network model for electricity price prediction", International journal of electrical power & energy systems, Volume 55, pp 187-194, February 2014.

[10] J. Janczura, "Pricing electricity derivatives within a Markov regime-switching model: a risk premium approach", Mathematical methods of operations research, Volume 79, Issue 1, pp 1-30, February 2014.

[11] R. Weron, M. Zator, "A note on using the Hodrick- Prescott filter in electricity markets", Energy economics, Volume 48, pp 1-6, March 2015.

[12] R. J. Hyndman, G. Athanasopoulos, "Forecasting: principles and practice", OTexts, October 2013.

[13] W. J. Hopp, M. L. Spearman, "Factory Physics," McGraw-Hill Higher Education, third edition, August 2011.

[14] J. Smart, "What is the controllable electrical demand from residential EVSE in the San Diego region?", Idaho National Laboratory, April 2015.

[15] P. D., "San Diego gas & electric power demand (load) in Kw", San Diego, CA, 2009.

[16] A. Faruqui, R. Hledik, and J. Tsoukalis, "The power of dynamic pricing," The electricity journal, Volume 22, Issue 3, pp 42-56, April 2009.

[17] Bartley, T, "Effect of air conditioning", retrieved on December 11, 2011, from annual pattern of daily peak demand: http://www.mpoweruk.com/electricity_demand.htm

[18] M. Roozbehani, M. A. Dahleh, S. Mitter, "Dynamic pricing and stabilization of supply and demand in modern electric power grids", Smart grid communications (SmartGridComm), first IEEE international conference, pp 543-548, 2010.

[19] Y. Koren "Collaborative filtering with temporal dynamics", Communications of the ACM magazine, Volume 53, Issue 4, pp 89-97, April 2010.

[20] Y. Koren, "Factorization meets the neighborhood: a multifaceted collaborative filtering model", The 14th ACM SIGKDD international conference on knowledge discovery and data mining, pp 426-434, 2008.

[21] Chu, "Personalized recommendation on dynamic content using predictive bilinear models", The 18th international conference on World Wide Web, pp 691-700, 2009.

A Class of Minimum-Time Synchronization Algorithms for 2D Rectangular Arrays Based on L-Shaped Mapping

Hiroshi Umeo[†], Takuya Yamawaki[†], Hiroki Uchino[†], and Kaori Ishida[†]

[†]School of Information Engineering

University of Osaka Electro-Communication

Neyagawa-shi, Hastu-cho, 18-8, Osaka, Japan

corresponding email address: umeo@cyt.osakac.ac.jp

{yamawaki, uchino, ishida}@cyt.osakac.ac.jp

Abstract—We introduce a new class of minimum-time FSSP (Firing Squad Synchronization Problem) algorithms for two-dimensional (2D) rectangular arrays. The algorithms in the class are all based on L-shaped mapping, where the synchronized configurations on 1D arrays are mapped in an L-shaped form onto 2D arrays efficiently, yielding minimum-time FSSP algorithms. We also present a comparative study of their recent implementations. Several state-efficient implementations, new insights into 2D synchronization and multi-dimensional extensions are also discussed.

Keywords—cellular automata; FSSP; synchronization.

I. INTRODUCTION

The synchronization in ultra-fine grained parallel computational model of cellular automata, known as the Firing Squad Synchronization Problem (FSSP), has been studied extensively for more than fifty years, and a rich variety of synchronization algorithms has been proposed. In the present paper, we attempt to answer the following questions:

- First, how many variations of 2D FSSP algorithms and their implementations which are based on L-shaped mapping exist?
- How can we synchronize 2D arrays using the L-shaped decomposition in minimum-time?
- What is the exact rule set for the real implementations of minimum-time FSSP algorithms on 2D arrays?
- How do the algorithms compare with each other?
- Can we extend those 2D FSSP algorithms to 3D arrays, or more generally, to multi-dimensional arrays?

In Section 2, we give a description of the 2D FSSP and review some basic results on the 2D FSSP algorithms. Section 3 introduces a new class of FSSP algorithms based on L-shaped mapping for 2D arrays. In the last section, we give a summary of the paper.

II. FIRING SQUAD SYNCHRONIZATION PROBLEM

A. FSSP on 2D Cellular Arrays

Fig. 1 shows a finite two-dimensional (2D) cellular array consisting of $m \times n$ cells. Each cell is an identical (except the border cells) finite-state automaton. The array operates in lock-step mode in such a way that the next state of each cell (except border cells) is determined by both its own present

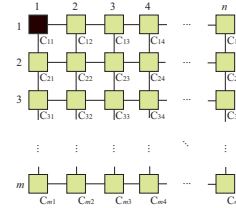


Fig. 1. A two-dimensional (2D) rectangular cellular automaton of size $m \times n$ arranged in m rows and n columns.

state and the present states of its north, south, east, and west neighbors. All cells (*soldiers*), except the north-west corner cell (*general*), are initially in the quiescent state at time $t = 0$ with the property that the next state of a quiescent cell with quiescent neighbors is the quiescent state again. At time $t = 0$, the north-west upper corner cell $C_{1,1}$ is in the *fire-when-ready* state, which is the initiation signal for the array. The FSSP is to determine a description (state set and next-state function) for cells that ensures all cells enter the *fire* state at exactly the same time and for the first time. The tricky part of the problem is that the same kind of soldier having a fixed number of states must be synchronized, regardless of the size $m \times n$ of the array. The set of states and transition rules must be independent of m and n .

A formal definition of the 2D FSSP is as follows: A cellular automaton \mathcal{M} is a pair $\mathcal{M} = (\mathcal{Q}, \delta)$, where

- 1) \mathcal{Q} is a finite set of states with three distinguished states $G, Q,$ and F , each in \mathcal{Q} . G is an initial general state, Q is a quiescent state, and F is a firing state, respectively.
- 2) δ is a next state function such that $\delta : \mathcal{Q} \times (\mathcal{Q} \cup \{*\})^4 \rightarrow \mathcal{Q}$. The state $*$ $\notin \mathcal{Q}$ is a pseudo state of the border of the array.
- 3) The quiescent state Q must satisfy the following conditions:

$$\delta(Q, Q, Q, Q, Q) = \delta(Q, *, Q, Q, *) = \delta(Q, *, Q, Q, Q) =$$

$$\delta(Q, *, *, Q, Q) = \delta(Q, Q, Q, *, Q) = \delta(Q, Q, *, *, Q) =$$

$$\delta(Q, Q, *, Q, Q) = \delta(Q, Q, *, Q, *) = \delta(Q, Q, Q, Q, *) =$$

$$Q.$$

A 2D cellular automaton of size $m \times n$, $\mathcal{M}_{m \times n}$ consisting of $m \times n$ copies of \mathcal{M} , is a 2D array of \mathcal{M} . Each \mathcal{M} is referred to as a cell and denoted by $C_{i,j}$, where $1 \leq i \leq m$ and $1 \leq j \leq n$. We denote a state of $C_{i,j}$ at time (step) t by $S_{i,j}^t$, where $t \geq 0, 1 \leq i \leq m, 1 \leq j \leq n$. Each tuple in the

next state function δ means that:

$$S_{\text{itself}}^{t+1} = \delta(S_{\text{itself}}^t, S_{\text{north}}^t, S_{\text{south}}^t, S_{\text{east}}^t, S_{\text{west}}^t).$$

A configuration of $\mathcal{M}_{m \times n}$ at time t is a function $\mathcal{C}^t : [1, m] \times [1, n] \rightarrow \mathcal{Q}$ and denoted as:

$$\begin{array}{cccc} S_{1,1}^t & S_{1,2}^t & \dots & S_{1,n}^t \\ S_{2,1}^t & S_{2,2}^t & \dots & S_{2,n}^t \\ S_{3,1}^t & S_{3,2}^t & \dots & S_{3,n}^t \\ \vdots & \vdots & \vdots & \vdots \\ S_{m,1}^t & S_{m,2}^t & \dots & S_{m,n}^t. \end{array}$$

A computation of $\mathcal{M}_{m \times n}$ is a sequence of configurations of $\mathcal{M}_{m \times n}$, $\mathcal{C}^0, \mathcal{C}^1, \mathcal{C}^2, \dots, \mathcal{C}^t, \dots$, where \mathcal{C}^0 is a given initial configuration such that:

$$S_{i,j}^0 = \begin{cases} G & i = j = 1 \\ Q & \text{otherwise.} \end{cases} \quad (1)$$

A configuration at time $t+1$, \mathcal{C}^{t+1} is computed by synchronous applications of the next transition function δ to each cell of $\mathcal{M}_{m \times n}$ in \mathcal{C}^t such that:

$$S_{i,j}^{t+1} = \delta(S_{i,j}^t, S_{i-1,j}^t, S_{i+1,j}^t, S_{i,j+1}^t, S_{i,j-1}^t).$$

A *synchronized configuration* of $\mathcal{M}_{m \times n}$ at time t is a configuration \mathcal{C}^t , $S_{i,j}^t = F$, for any $1 \leq i \leq m$ and $1 \leq j \leq n$.

The FSSP is to obtain an \mathcal{M} such that, for any $m, n \geq 2$,

- 1) A synchronized configuration at time $t = T(m, n)$, $\mathcal{C}^{T(m,n)} : S_{i,j}^{T(m,n)} = F$, for any $1 \leq i \leq m$ and $1 \leq j \leq n$, can be computed from an initial configuration \mathcal{C}^0 in equation (1).
- 2) For any t, i such that $1 \leq t \leq T(m, n) - 1$, $1 \leq i \leq m$, $1 \leq j \leq n$, $S_{i,j}^t \neq F$.

No cells fire before time $t = T(m, n)$. We say that the array $\mathcal{M}_{m \times n}$ is synchronized at time $t = T(m, n)$ and the function $T(m, n)$ is the time complexity for the synchronization.

B. Lower-Bound and Optimality in 2D FSSP Algorithms

As for the time optimality of the 2D FSSP algorithms, Beyer [1] and Shinahr [4] gave a lower bound and minimum-time FSSP algorithm.

Theorem 1 There exists no cellular automaton that can synchronize any 2D array of size $m \times n$ in less than $m + n + \max(m, n) - 3$ steps, where the general is located at one corner of the array.

Theorem 2 There exists a cellular automaton that can synchronize any 2D array of size $m \times n$ at exactly $m + n + \max(m, n) - 3$ steps, where the general is located at one corner of the array.

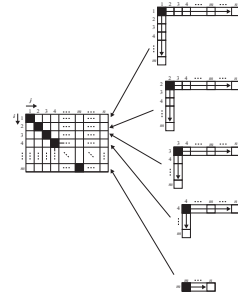


Fig. 2. A 2D array synchronization scheme based on L-shaped mapping developed in Beyer [1] and Shinahr [4].

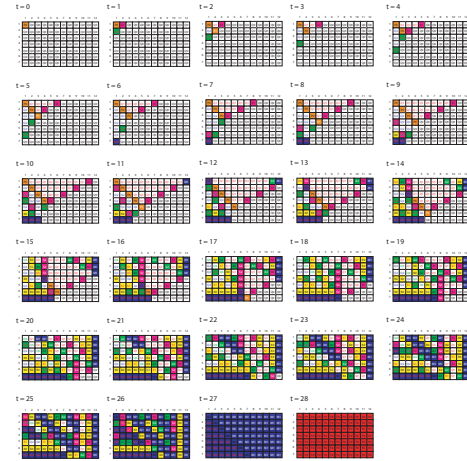


Fig. 3. Snapshots of the configurations of the Shinahr's 28-state synchronization algorithm on a rectangle array of size 7×12 .

III. ALGORITHM \mathcal{A}_1

The first minimum-time FSSP algorithm \mathcal{A}_1 , developed independently by Beyer [1] and Shinar [4], is based on rotated L-shaped mapping which maps configurations of generalized FSSP solutions on 1D L-shaped arrays onto 2D arrays. A rectangular array of size $m \times n$ is regarded as $\min(m, n)$ rotated L-shaped 1D arrays, where each rotated L-shaped 1D array is synchronized independently by using the Generalized FSSP (GFSSP) algorithm. The GFSSP algorithm [3], [8] is stated as follows:

Theorem 3 There exists a cellular automaton that can synchronize any 1D array of length n in optimum $n + \max(k, n - k + 1) - 2$ steps, where the general is located on the k th cell from left end.

We overview the algorithm \mathcal{A}_1 operating on an array of size $m \times n$. Configurations of the generalized synchronization processes on 1D array can be mapped on the rotated L-shaped array. We refer it as L-array. See Fig. 2. At time $t = 0$, the north-west cell $C_{1,1}$ is in general state and all other cells are in quiescent state. For any i such that $1 \leq i \leq \min(m, n)$, the cell $C_{i,i}$ will be in the general state at time $t = 3i - 3$. A special signal which travels towards a diagonal direction is used to generate generals on the cells $\{C_{i,i} | 1 \leq i \leq \min(m, n)\}$. For each i such that $1 \leq i \leq \min(m, n)$, the cells $\{C_{i,j} | i \leq j \leq n\}$ and $\{C_{k,i} | i \leq k \leq m\}$ constitute the i th L-shaped array. Note that the i th general generated at time $t = 3i - 3$ is on

the $(m - i + 1)$ th cell from the left end of the i th L-array. The length of the i th L-array is $m + n - 2i + 1$. Thus, using Theorem 3, the i th L-array can be synchronized at exactly $t_i = 3i - 3 + m + n - 2i + 1 - 2 + \max(m - i + 1, n - i + 1) = m + n + \max(m, n) - 3$, which is independent of i . In this way, all of the L-arrays can be synchronized simultaneously.

Shinahr [4] presented a 28-state implementation of the algorithm, where most (97%) of the transition rules had *wild cards* which can match any state. Later, H. Umeo, K. Ishida, K. Tachibana, and N. Kamikawa [7] showed that the rule set consists of 12849 transition rules and it is valid for the synchronization for any rectangle arrays of size $m \times n$ such that $2 \leq m, n \leq 500$. Fig. 3 illustrates snapshots of the configurations on an array of size 7×12 based on our new 28-state 12849-rule implementation. Thus, we have:

Theorem 4 The algorithm \mathcal{A}_1 can synchronize any $m \times n$ rectangular array in optimum $m + n + \max(m, n) - 3$ steps.

IV. ALGORITHM \mathcal{A}_2

In this section, we develop a new minimum-time FSSP algorithm \mathcal{A}_2 based on a new L-shaped mapping. First, we introduce a *freezing-thawing* technique that yields a delayed synchronization algorithm for 1D array. The technique was developed by Umeo [5] for designing several fault-tolerant FSSP algorithms for 1D arrays.

Theorem 5 Let t_0, t_1, t_2 and Δt be any integer such that $t_0 \geq 0, t_1 = t_0 + n - 1, t_1 \leq t_2$ and $\Delta t = t_2 - t_1$. We assume that a usual synchronization operation is started at time $t = t_0$ by generating a special signal at the left end of 1D array of length n . We also assume that the right end cell of the array receives another special signals from *outside* at time $t_1 = t_0 + n - 1$ and $t_2 = t_1 + \Delta t = t_0 + n - 1 + \Delta t$, respectively. Then, there exists a 1D cellular automaton that can synchronize the array at time $t = t_0 + 2n - 2 + \Delta t$.

We can freeze the entire configuration during Δt steps and delay the synchronization on the array for Δt steps. We refer the scheme as the *freezing-thawing technique*. Here, the freezing-thawing technique is employed efficiently.

A. Segmentation of rectangular array

First, we consider the case where $m \leq n$. We regard a 2D array of size $m \times n$ as consisting of m rotated (90° in counterclockwise direction) *L-shaped* 1D arrays. Each L-shaped array is denoted by $L_i, 1 \leq i \leq m$, shown in Fig. 4. Each L_i is divided into three segments, referred to as 1st, 2nd and 3rd segments. The length of each segment of L_i is $i, n - m$, and i , respectively.

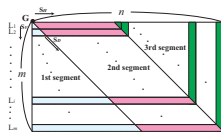


Fig. 4. A 2D array of size $m \times n$ is regarded as consisting of m rotated (90° in counterclockwise direction) L-shaped 1-D arrays. Each L-shaped array is divided into three segments.

B. Starting the synchronization process

At time $t = 0$, a 2D array M has a *general* at $C_{1,1}$ and any other cells of the array are in quiescent state. The *general* G (denoted by \bullet in Fig. 4) generates three signals s_V, s_D and s_H , simultaneously, each propagating at $1/1$ -speed in the vertical, diagonal and horizontal directions, respectively. See Fig. 4. The s_V - and s_H -signals work for generating wake-up signals for the 1st and 3rd segments on each L-shaped array. The s_D -signal is used for printing a delimiter between the 1st and 2nd segments. Their operations are as follows:

- **Signal s_V :** The s_V -signal travels along the 1st column and reaches $C_{m,1}$ at time $t = m - 1$. Then, it returns there and begins to travel at $1/2$ -speed along the 1st column towards $C_{1,1}$. On the return's way, the signal initiates the synchronization process for the 1st segment of each L_i . Thus, a new *general* G_{i1} for the synchronization of the 1st segment of each L_i and its wake-up signal for the 1st segment are generated at time $t = 3m - 2i - 1$ for $1 \leq i \leq m$.
- **Signal s_H :** The s_H -signal travels along the 1st row at $1/1$ -speed and reaches $C_{1,n}$ at time $t = n - 1$. Then it reflects there and returns the same route at $1/2$ -speed. Each time it visits a cell of the 1st row on its return way, it generates a *general* G_{i3} at time $t = 2m + n - 2i - 1$ to initiate a synchronization for the 3rd segment on each $L_i, 1 \leq i \leq m$.
- **Signal s_D :** The s_D -signal travels along a diagonal line by repeating a zig-zag movement: going one cell to the right, then going down one cell. Each time it visits a cell $C_{i,i}$ on the diagonal, it marks a special symbol to denote a delimiter between the 1st and 2nd segments. The symbol on $C_{i,i}$ is marked at time $t = 2i - 2$ for any $i, 1 \leq i \leq m$. Note that the wake-up signal of the 1st segment of L_m knows its right end by the arrival of s_D -signal, where they meet at $C_{m,m}$ at the very time $t = 2m - 2$.

The wake-up signal generated by G_{i1} reaches its right end at time $t = 3m - i - 2$ and generates a new *general* G_{i2} for the 2nd segment. The new *general* G_{i2} generates a wake-up signal. The wake-up signals for the 2nd and 3rd segments of L_i meet on $C_{i,n-m+i}$ at time $t = 2m + n - i - 2$. The collision of the two signals acts as a delimiter between the 2nd and 3rd segments in the case where $m \leq n$. Note that the synchronization operations on the 1st and 2nd segments are started at the left end of each segment. On the other hand, the synchronization on the 3rd segment is started at the right (upper) end of the segment.

C. Synchronization of L_m

Now, we consider the synchronization on L_m . Fig. 5 (left) shows a space-time diagram for synchronizing L_m . As was mentioned in the previous subsection, the synchronization of the 1st, 2nd and 3rd segments of L_m are started by the generals G_{m1}, G_{m2} and G_{m3} at time $t = m - 1, t = 2m - 2$ and $t = n - 1$, respectively. Each General always generates a wake-up and a *pre-thawing* signal, each propagating at $1/1$ - and $1/2$ -speed in the same direction. The wake-up signal wakes up cells in the segment itself, however, the pre-thawing signal

generates a thawing signal at its *neighboring* segment that it encounters first. Precisely, the pre-thawing signal generated by G_{m1} reaches the left end of the 2nd segments at time $t = 3m - 3$. In the case $m \leq n$, the configuration of the 2nd segment have not been frozen yet, and the signal doesn't work. The pre-thawing signal generated by G_{m2} arrives at the delimiter between the 2nd and 3rd segments at time $t = 2n - 2$ and generates a thawing signal for the 3rd segment. In a similar way, the pre-thawing signal generated by G_{m3} initiates its thawing operation for the 2nd segment at time $t = 2m + n - 2$.

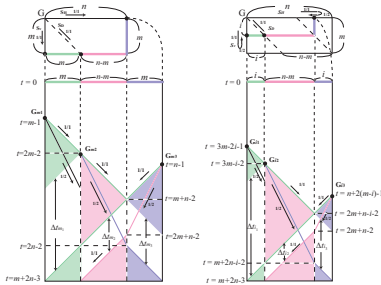


Fig. 5. Space-time diagram for synchronizing L_m (left) and L_i (right).

At time $t = 2m - 2$, the wake-up signal of the 1st segment reaches its right end and generates a freezing signal for the segment. Simultaneously, it initiates synchronization on the 2nd segments. The freezing signal for the 1st segment propagates in left direction at $1/1$ -speed, and freezes the configuration on it. At time $t = 2n - 2$, a thawing signal is generated at its right end, that is initiated by the arrival of the freezing signal for the 2nd segment, which will be explained again below. The wake-up signal for the 2nd and 3rd segments meets $C_{m,n}$ at time $t = m + n - 2$, where $C_{m,n}$ acts as an end of the both two segments. A new freezing signal is generated there for the 2nd and 3rd segments. It propagates in right and left directions at $1/1$ -speed to freeze the synchronization operations on the 2nd and 3rd segments, respectively. The freezing signal for the 2nd segment reaches at its left end at time $t = 2n - 3$, which generates a thawing signal for the 1st segment at time $t = 2n - 2$. The thawing signals for the 1st, 2nd and 3rd segments are generated at time $t = 2n - 2$, $t = 2m + n - 2$ and $t = 2n - 2$, respectively. Synchronization operations on j th ($1 \leq j \leq 3$) segment are delayed for Δt_{m_j} steps such that:

$$\Delta t_{m_j} = \begin{cases} 2(n - m) & j = 1 \\ m & j = 2 \\ n - m & j = 3 \end{cases} \quad (2)$$

The synchronization for the 1st segment is started at time $t = m - 1$ and its operations are delayed for $\Delta t = \Delta t_{m_1} = 2(n - m)$ steps. Now, letting $t_0 = m - 1$, $\Delta t = \Delta t_{m_1} = 2(n - m)$ in Theorem 5, the 1st segment of L_m can be synchronized at time $t = t_0 + 2m - 2 + \Delta t = m + 2n - 3 = m + n + \max(m, n) - 3$. In a similar way, the 2nd and the 3rd segments can be synchronized at time $t = m + 2n - 3 = m + n + \max(m, n) - 3$. Thus, L_m can be synchronized at time $t = m + n + \max(m, n) - 3$.

D. Synchronization of L_i

Now, we discuss the synchronization for L_i , $1 \leq i \leq m$. Fig. 5 (right) shows a space-time diagram for synchronizing L_i . The wake-up signals for the three segments of L_i are generated at time $t = m + 2(m - i) - 1$, $3m - i - 2$, and $n + 2(m - i) - 1$, respectively. Generation of the corresponding freezing and thawing signals are done in a similar way as employed in L_m . Synchronization operations on j th ($1 \leq j \leq 3$) segment are delayed for Δt_{i_j} steps such that:

$$\Delta t_{i_j} = \begin{cases} 2(n - m) & j = 1 \\ i & j = 2 \\ n - m & j = 3 \end{cases} \quad (3)$$

The synchronization for the 1st segment of L_i is started at time $t = m + 2(m - i) - 1$ and its operations are delayed for $\Delta t = \Delta t_{i_1} = 2(n - m)$ steps. Now, letting $t_0 = m + 2(m - i) - 1$, $\Delta t = \Delta t_{i_1} = 2(n - m)$ in Theorem 5, the 1st segment of L_i can be synchronized at time $t = t_0 + 2i - 2 + \Delta t = m + 2n - 3 = m + n + \max(m, n) - 3$. In a similar way, the 2nd and the 3rd segments can be synchronized at time $t = m + 2n - 3 = m + n + \max(m, n) - 3$. Thus, L_i can be synchronized at time $t = m + n + \max(m, n) - 3$.

E. Synchronization of rectangle longer than wide

In the case where $m > n$, a two-dimensional array of size $m \times n$ is regarded as consisting of n rotated L-shaped arrays. Segmentation and synchronization operations on each L-shaped array can be done almost in a similar way. It is noted that the right end cell of the 2nd segment works as a *General* for the 2nd segment. Any rectangle of size $m \times n$ can be synchronized at time $t = 2m + n - 3 = m + n + \max(m, n) - 3$.

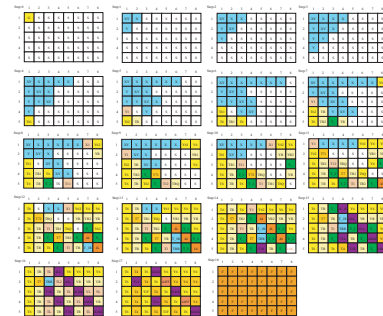


Fig. 6. Snapshots of the synchronization process on 5×8 array.

Fig. 6 shows some snapshots of the synchronization process operating in optimum-steps on a 5×8 array. Now, we can establish the next theorem.

Theorem 6 The algorithm \mathcal{A}_2 can synchronize any $m \times n$ rectangular array in optimum $m + n + \max(m, n) - 3$ steps.

V. ALGORITHM \mathcal{A}_3

In this section, we develop a new minimum-time firing squad synchronization algorithm \mathcal{A}_3 based on a new L-shaped mapping. The overview of the algorithm \mathcal{A}_3 is as follows:

- 1) A 2D array of size $m \times n$ is regarded as $\min(m, n)$ rotated or mirrored *L-shaped* 1-D arrays, each consisting of a horizontal and a vertical segment.
- 2) The shorter segment is synchronized by the freezing-thawing technique with $\Delta t = |m - n|$ steps delay. The longer one is synchronized in the usual way without using the freezing-thawing technique.
- 3) All of the *L-shaped* arrays fall into a special firing (synchronization) state simultaneously and for the first time.

A. Segmentation of rectangular array of size $m \times n$

First, we consider the case where $m \leq n$. We assume that the initial general G is on the north-west corner denoted by \bullet in Fig. 7. We regard a two-dimensional array of size $m \times n$ as consisting of m rotated (90° in counterclockwise direction) *L-shaped* one-dimensional arrays. Each bending point of the *L-shaped* array is the farthest one from the general. Each *L-shaped* array is denoted by $L_i, 1 \leq i \leq m$, shown in Fig. 7. Each L_i is divided into two segments, that is, one horizontal and one vertical segment, each referred to as 1st and 2nd segments. The length of each segment of L_i is $n - m + i$ and i , respectively.

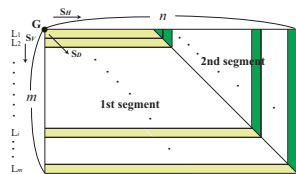


Fig. 7. A 2D array of size $m \times n$ ($m \leq n$) is regarded as consisting of m rotated (90° in counterclockwise direction) *L-shaped* 1-D arrays. Each *L-shaped* array is divided into two segments.

B. Starting the synchronization process

At time $t = 0$, a two-dimensional array M has a general at $C_{1,1}$ and any other cells of the array are in quiescent state. The general G (denoted by \bullet in Fig. 7) generates three signals s_V, s_D and s_H , simultaneously, each propagating at $1/1$ -speed in the vertical, diagonal and horizontal directions, respectively. See Fig. 3. The s_V - and s_H -signals work for generating wake-up signals for the 1st and 2nd segments on each *L-shaped* array. The s_D -signal is used for printing a special marker "■" for generating a thawing signal that thaws frozen configurations on shorter segment. Their operations are as follows:

- **Signal s_V :** The s_V -signal travels along the 1st column and reaches $C_{m,1}$ at time $t = m - 1$. Then, it returns there and begins to travel again at $1/2$ -speed along the 1st column towards $C_{1,1}$. On the return's way, the signal initiates the synchronization process for the 1st segment of each L_i . Thus, a new general G_{i1} for the synchronization of the 1st segment of each L_i is generated, together with its wake-up signal, at time $t = 3m - 2i - 1$ for $1 \leq i \leq m$.
- **Signal s_D :** The s_D -signal travels along a principal diagonal line by repeating a zigzag movement: going one cell to the right, then going down one cell. Each

time it visits cell $C_{i,i}$ on the diagonal, it marks a special symbol "■" to inform the wake-up signal on the segment of the position where a thawing signal is generated for the neighboring shorter segment. The symbol on C_{ii} is marked at time $t = 2i - 2$ for any $i, 1 \leq i \leq m$. Note that the wake-up signal of the 1st segment of L_m knows the right position by the arrival of the s_D -signal, where they meet at $C_{m,m}$ at the very time $t = 2m - 2$.

- **Signal s_H :** The s_H -signal travels along the 1st row at $1/1$ -speed and reaches $C_{1,n}$ at time $t = n - 1$. Then it reflects there and returns the same route at $1/2$ -speed. Each time it visits a cell of the 1st row on its return way, it generates a general G_{i2} at time $t = 2m + n - 2i - 1$ to initiate a synchronization for the 2nd segment on each $L_i, 1 \leq i \leq m$.

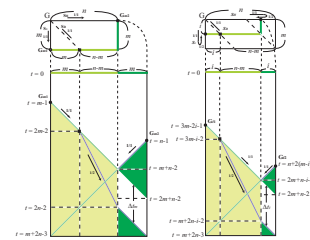


Fig. 8. Space-time diagram for synchronizing L_m (left) and L_i (right).

The wake-up signals for the 1st and 2nd segments of L_i meet on $C_{i,n-m+i}$ at time $t = 2m + n - i - 2$. The collision of the two signals acts as a delimiter between the 1st and 2nd segments. Note that the synchronization operations on the 1st segment are started at the left end of the segment. On the other hand, the synchronization on the 2nd segment is started at the right (upper) end of the segment. The wake-up signal generated by G_{i1} reaches a cell where the special mark is printed at time $t = 3m - i - 2$ and generates a thawing signal which travels at $1/2$ -speed along the 1st segment to thaw the configuration on the 2nd segment.

C. Synchronization of L_m

Now, we consider the synchronization on L_m . Fig. 8 (left) shows a space-time diagram for synchronizing L_m . As was mentioned in the previous subsection, the synchronization of the 1st and 2nd segments of L_m are started by the generals G_{m1} and G_{m2} at time $t = m - 1$ and $t = n - 1$, respectively. Each general generates a wake-up signal propagating at $1/1$ -speed. The wake-up signal for the 1st and 2nd segments meets $C_{m,n}$ at time $t = m + n - 2$, where $C_{m,n}$ acts as an end of the both two segments. A freezing signal is generated simultaneously there for the 2nd segment at time $t = m + n - 2$. It propagates in upper (right in Fig. 8 (left)) direction at $1/1$ -speed to freeze the synchronization operations on the 2nd segment. At time $t = 2m - 2$, the wake-up signal of the 1st segment reaches the symbol "■" and generates a thawing signal for the 2nd segment. The thawing signal starts to propagate from the cell at $1/2$ -speed in the same direction. Those two signals reach at the left end of the 2nd segment with time difference $n - m$ which is equal to the delay for the 2nd segment. The synchronization for the 1st segment is started

at time $t = m - 1$ and it can be synchronized at time $t = m + 2n - 3 = m + n + \max(m, n) - 3$ by the usual way. On the other hand, the synchronization for the 2nd segment is started at time $t = n - 1$ and its operations are delayed for $\Delta t = \Delta t_m = n - m$ steps. Now, letting $t_0 = n - 1, \Delta t = n - m$ in Theorem 1, the 2nd segment of length m on L_m can be synchronized at time $t = t_0 + 2m - 2 + \Delta t = m + 2n - 3 = m + n + \max(m, n) - 3$. Thus, L_m can be synchronized at time $t = m + n + \max(m, n) - 3$.

D. Synchronization of L_i

Now, we discuss the synchronization for $L_i, 1 \leq i \leq m$. Fig. 8 (right) shows a space-time diagram for synchronizing L_i . The wake-up signals for the two segments of L_i are generated at time $t = 3m - 2i - 1$ and $n + 2(m - i) - 1$, respectively. Generation of freezing and thawing signals is done in a similar way as employed in L_m . Synchronization operations on the 2nd segment are delayed for $\Delta t_i = n - m$ steps. The synchronization for the 1st segment of length $n - m + i$ is started at time $t = 3m - 2i - 1$ and it can be synchronized by a usual method at time $t = m + 2n - 3 = m + n + \max(m, n) - 3$. On the other hand, the synchronization for the 2nd segment is started at time $t = n + 2(m - i) - 1$ and its operations are delayed for $\Delta t = \Delta t_i = n - m$ steps. Now, letting $t_0 = n + 2(m - i) - 1, \Delta t = n - m$ in Theorem 5, the 2nd segment of length i of L_i can be synchronized at time $t = t_0 + 2i - 2 + \Delta t = m + 2n - 3 = m + n + \max(m, n) - 3$. Thus, L_i can be synchronized at time $t = m + n + \max(m, n) - 3$.

E. Synchronization of rectangle longer than wide

In the case where $m > n$, a two-dimensional array of size $m \times n$ is regarded as consisting of n mirrored *L-shaped* arrays. Segmentation and synchronization operations on each *L-shaped* array can be done almost in a similar way. It is noted that the thawing signal is generated on the 2nd segment to thaw frozen configurations on the 1st segment. Any rectangle of size $m \times n$ can be synchronized at time $t = 2m + n - 3 = m + n + \max(m, n) - 3$.

One notes that the algorithm needs no a priori knowledge on side length of a given rectangle, that is, whether wider than long or longer than wide. Fig. 9 shows some snapshots of the synchronization process operating in optimum-steps on a 6×9 arrays. Now, we can establish the next theorem.

Theorem 7 The synchronization algorithm \mathcal{A}_3 can synchronize any $m \times n$ rectangular array in optimum $m + n + \max(m, n) - 3$ steps.

VI. CONCLUSION AND FUTURE WORK

In the present paper, we gave a survey on recent developments in FSSP algorithms for 2D cellular arrays. We focused our attention on a new class of the 2D minimum-time FSSP algorithms based on L-shaped mapping. It is shown that the L-shaped mapping presents a rich variety of 2D minimum-time FSSP algorithms. As a future work, the mapping could be applied to the design of 3D minimum-time FSSP algorithms. The procedure would be as follows: First, a 2D version of the freezing-thawing technique has to be developed. Secondary, decompose a given 3D array into many thin layers, each consisting of three 2D faces. Last, apply the minimum-time 2D FSSP algorithm developed in this paper to each face.

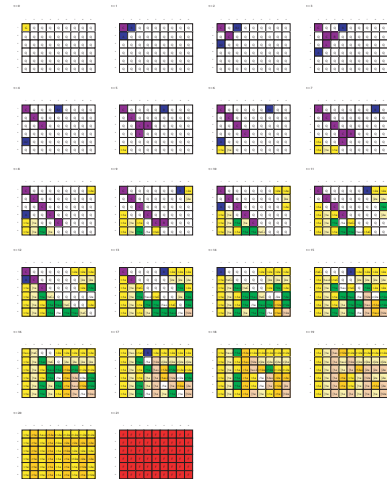


Fig. 9. Snapshots of the synchronization process on 6×9 array.

REFERENCES

- [1] W. T. Beyer, "Recognition of topological invariants by iterative arrays." Ph.D. Thesis, MIT, (1969), pp. 144.
- [2] E. F. Moore, "The firing squad synchronization problem." In *Sequential Machines, Selected Papers* (E. F. Moore, ed.), Addison-Wesley, Reading MA., pp. 213-214(1964).
- [3] F. R. Moore and G. G. Langdon, "A generalized firing squad problem." *Information and Control*, 12, pp.212-220(1968).
- [4] I. Shinahr, "Two- and three-dimensional firing squad synchronization problems." *Information and Control*, vol. 24, pp. 163-180(1974).
- [5] H. Umeo, "A simple design of time-efficient firing squad synchronization algorithms with fault-tolerance." *IEICE Trans. on Information and Systems*, Vol. E87-D, No.3, pp.733-739(2004).
- [6] H. Umeo, "Firing squad synchronization problem in cellular automata." In *Encyclopedia of Complexity and System Science*, R. A. Meyers (Ed.), Springer, Vol.4, pp. 3537-3574(2009).
- [7] H. Umeo, K. Ishida, K. Tachibana, and N. Kamikawa, "A transition rule set for the first 2-D optimum-time synchronization algorithm." *Proc. of the 4th International Workshop on Natural Computing*, PICT 2, Springer, pp.333-341(2009).
- [8] H. Umeo, N. Kamikawa, K. Nishioka, and S. Akiguchi, "Generalized firing squad synchronization protocols for one-dimensional cellular automata - a survey." *Acta Physica Polonica B, Proceedings Supplement*. Vol.3, pp.267-289(2010).
- [9] H. Umeo and H. Uchino, "A new time-optimum synchronization algorithm for rectangle arrays." *Fundamenta Informaticae*, Vol.87, No.2, pp.155-164(2008).

APT Detection with Host-Based Intrusion Detection System and Intelligent Systems

Seong Oun Hwang

Department of Computer and Information Communication Engineering, Hongik University
 Sejong-ro 2639, Jochiwon, Sejong, Republic of Korea
 Email: sohwan@hongik.ac.kr

Abstract—Recently, Advanced Persistent Threat (APT) attacks have targeted many institutions, such as governments and companies. APT refers to a type of offensive attacks, which have been performed for a long time using unique attack vectors and malware specifically developed for the target organization. Due to its complicated and sophisticated nature, this threat can be very hard to detect compared to other types of attacks. In this paper, we propose a new method to detect APT attacks by profiling user activities based on Indicator of Compromise (IOC) and chasing malware activities.

Keywords—OSSEC; APT; IOC; HIDS.

I. INTRODUCTION

OSSEC [2] is a free, Open Source host-based intrusion detection system (HIDS) SEcURITY. It performs functions such as integrity checking (i.e., check if a file is modified, deleted, or added to the agent and send the information to the server), process information checking (i.e., check if a process is terminated, or started), and APT [1] monitoring (i.e., monitor the APT attacks). OSSEC systems largely consist of two parts: agent and manager. The agent is a small program, or collection of programs, installed on a system to be monitored. The agent collects information and forwards it to the manager for analysis and correlation. The manager is the central piece of the OSSEC deployment. It stores the file integrity information, logs, events, and system auditing entries. All the rules, decoders and major configuration options are stored centrally in the manager, which makes it easier to administer in case of a large number of agents.

The rest of this paper is organized as follows. In Section II, we propose our OSSEC-based intrusion detection system. In Section III, we describe how to combine the proposed system with other intelligent systems. Finally, we give conclusions in Section IV.

II. PROPOSED SYSTEM

In this section, we implement our host-based intrusion detection system using the OSSEC system model.

A. File Integrity Check

- Get a new MD5 (Message Digest 5) value of a file when it is modified.
- Submit the new MD5 value to Virustotal [3], as shown in Figure 1. In order to scan a file using Virustotal, we need to have both the MD5 hash value and Virustotal key.

```
import simplejson
import urllib
import urllib2
url = "https://www.virustotal.com/vtapi/v2/file/rescan"
parameters = ("resource": "99017f6eebbac24f351415dd410d522d",
             ,"9542385c2c440a0e561aa5c1e4ac5b84307428589d854a969d52ab2daccff5e27")
data = urllib.urlencode(parameters)
req = urllib2.Request(url, data)
response = urllib2.urlopen(req)
json = response.read()
print json
```



Figure 1. MD5 Scan by Virustotal API.

B. Monitoring of Running Processes

Step 1: Accepting remote commands. The first step is to configure the agent log collector option to accept remote commands from the manager. That can be done by editing "internal_options.conf" file (usually located at "C:\Program Files(x86)\ossec-agent\internal_options.conf") and setting the variable logcollector.remote_commands to 1.

Step 2: Specifying the command to list running processes. This is a configuration that can be done both at the agent and at the manager side (using the shared directory). It only depends on how many agents you want to use this command. In our case, we edit "/var/ossec/etc/shared/agent.conf" configuration file, and have these settings pushed to our Windows agents. The command we use to list processes in Windows Operating

```
<agent_config os="windows">
  <localfile>

  <log_format>full_command</log_format>
  <command>tasklist</command>
  <frequency>60</frequency>
</localfile>
```

Figure 2. Agent Configuration.

System is "tasklist", as shown in Figure 2. There are other options, such as using wmic, but this one is sufficient. For Unix systems, "ps" can be used.

Step 3: Creating local rules. In this step, we edit our "/var/ossec/rules/local_rules.xml" file to add rules that will trigger an alert if our critical process is not running. For the purpose of this example, we will use "wordpad.exe", as shown in Figure 3, but, of course, it could be any other name.

The first rule (id "100050") will trigger a level "7" alert every time tasklist command is executed, unless (as defined in rule "100051") the output matches the string "wordpad.exe". If this is the case, the alert level is set to "0", which means that no alert would be triggered.

Now, we just need to save these changes and restart the

```
<rule id="100050" level="7">
  <if_sid>530</if_sid>
  <match>ossec: output: 'tasklist'</match>
  <description>Critical process not found.</description>
  <group>process_monitor,</group>
</rule>
<rule id="100051" level="0">
  <if_sid>100050</if_sid>
  <match>wordpad.exe</match>
  <description>Processes running as expected.</description>
  <group>process_monitor,</group>
</rule>
```

Figure 3. Local Rules.

manager for them to be applied. We can do that by running the "ossec-control restart" command.

C. APT Monitoring

APT monitoring can be performed using the APT detection system model, as shown in Figure 4.

Input part: This system is fed with an MD5 hash value of file handling processes which are collected through OSSEC agents along with URL. These MD5 inputs are not determined to be malicious or not yet, and will be under further testing at the Analyzing step.

Audit part: In this part, the MD5 values of files and processes are submitted to Virustotal and the Virustotal's responses are passed to Analyzing part.

Analyzing part: In this part, we check if a suspicious MD5 matches any of existing detection information in Normal Profile database or IOC database [4].

Report part: The report module provides information on the suspicious MD5/file to the security administrator and stores the analysis result in Signature Profile information database. In addition, it stores suspicious URLs to malicious URL (MURL) database.

Output part: It provides signature information in the form of OpenIOC [5] to security organizations or vendors. Figure 5 shows one of the FireEye publicly shared IOCs [4].

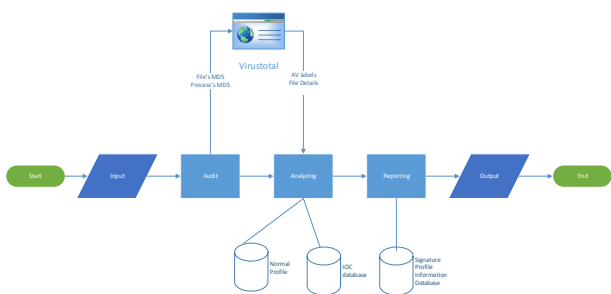


Figure 4. APT Detection System Model.

III. INTEGRATION WITH INTELLIGENT SYSTEMS

The above OSSEC-based intrusion detection system only cannot efficiently cope with malware which are so frequently updated and distributed in a short time that the existing anti-virus vendors including Virustotal cannot make the corresponding malware signatures appropriately. To overcome this limitation, we combine the above system with our intelligent systems such as Payload Chase System (PCS) and Malicious Site Detection System (MSDS) together. PCS periodically

```
xmlns="http://schemas.mandiant.com/2010/ioc" id="7b9e87c5-b619-4a13-b862-0145614d359a"
last-modified="2015-05-13T08:16:33Z"
<short description>BLACKOFFER (FAMILI)</short description>
<description>This IOC contains indicators detailed in the whitepaper "Hiding in Plain Sight: FireEye and Microsoft Expose Chinese APT Group's Offlocation Tactic". The whitepaper can be read here: https://www.fireeye.com/blog/threat-research/2015/05/hiding_in_plain_sight.html. This IOC contains indicators for the BLACKOFFER malware family that is attributed to APT7.</description>
<keywords/>
<authored by>FireEye</authored by>
<authored date>2014-10-15T21:02:19</authored date>
<links>
  <link rel="category">Backdoor</link>
  <link rel="threatcategory">APT</link>
  <link rel="threatcategory">APT7</link>
  <link rel="license">Apache 2.0</link>
</links>
<definition>
  <indicator id="5d0b0faa-aa4f-4d07-9698-4b8806279ab4" operator="OR">
    <indicator item id="8d0e7a7e-470e-4d1e-b4ee-9a070e4629" condition="is">
      <content document="FileItem" search="FileItem/MD5sum" type="mal"/>
    </indicator item>
    <indicator item id="f8dca8db-39e0-4e30-907f-cd3f0e9948f3" condition="is">
      <content document="FileItem" search="FileItem/MD5sum" type="mal"/>
    </indicator item>
    <indicator item id="522fcb8bd4271249975da2081a0ca0c" condition="is">
      <content type="md5">522fcb8bd4271249975da2081a0ca0c</content>
    </indicator item>
    <indicator item id="69526072-b5ba-4efc-9247-821ddf1a93cb" condition="is">
      <content document="ProcessItem" search="ProcessItem/HandleList/Handle/Name" type="mal"/>
      <content type="string">358bd08946</content>
    </indicator item>
    <comment>Process Handle Type: events</comment>
    </indicator item>
    <indicator item id="77521051-2713-4389-ab19-e7c86f020aba" condition="is">
      <content document="DnsEntryItem" search="ProcessItem/HandleList/Handle/Name" type="mal"/>
      <content type="string">Psp_No_Management</content>
    </indicator item>
    <comment>Process Handle Type: events</comment>
    </indicator item>
    <indicator item id="dc9c0e69-8b9d-4412-ab38-2ad7117442" condition="contains">
      <content document="DnsEntryItem" search="DnsEntryItem/RecordName" type="mal"/>
      <content type="string">translate.wordreference.com</content>
    </indicator item>
    <indicator item id="82b1a0e-af1e-4a6f-9e0f-902bf90b7c" condition="contains">
      <content document="DnsEntryItem" search="DnsEntryItem/RecordName" type="mal"/>
      <content type="string">news.juachad.net</content>
    </indicator item>
    <indicator item id="98b14a1d-3840-4d47-87d2-c1a180c49e77" condition="is">
      <content document="FileItem" search="FileItem/MD5sum" type="mal"/>
      <content type="md5">0516a730b3729a70e46563b3c384</content>
    </indicator item>
    <indicator item id="7902a58e-f85e-48a0-8447-04b495b004" condition="is">
      <content document="PortItem" search="PortItem/remotIP" type="mal"/>
      <content type="IP">69.80.72.165</content>
    </indicator item>
    <indicator item id="0c4522e0-a19d-411a-9867-cb3bd6c31e5" condition="is">
      <content document="PortItem" search="PortItem/remotIP" type="mal"/>
      <content type="IP">110.45.151.43</content>
    </indicator item>
  </indicator>
</definition>
</ioc>
```

Figure 5. APT IOC.

checks the MURL database and collects new malware downloadable at the distribution page by malware producers in advance, executes them in a controllable environment if they are determined to be malicious by MSDS, monitors their changes in time and records the detailed information such as period of alteration, IP address of the distribution, geographical information, attack patterns to identify and track further the underlying attacker group.

IV. CONCLUSION AND FUTURE WORK

Since the OSSEC system model is based on the signature approach, it cannot efficiently cope with new malware, such as malware variants whose signatures are not published yet. To overcome this problem, we proposed a new host-based intrusion detection method by profiling user activities based on IOC and chasing malware activities periodically. Our future work is to completely develop the OSSEC system model and collect more APT samples enough to classify them using machine learning techniques. If we finally classify APT malware with acceptable accuracy and performance, we could overcome the limitation of the signature approach taken by the conventional anti-virus vendors.

ACKNOWLEDGMENT

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2014R1A1A2054174).

REFERENCES

- [1] R. Jasek, M. Kolarik and T. Vymola, "Apt detection system using honeypots," In Proceedings of the 13th International Conference on Applied Informatics and Communications (AIC'13), WSEAS Press, pp. 25-29, 2013.
- [2] <http://ossec.github.io>, accessed January 2017.
- [3] <https://virustotal.com>, accessed January 2017.
- [4] <https://github.com/fireeye/iocs>, accessed January 2017.
- [5] <http://openioc.org>, accessed January 2017.

Memristive Implementation of Fuzzy Logic for Cognitive Computing

Martin Klimo, Ondrej Šuch
 Department of InfoComm Networks
 University of Zilina
 Zilina, Slovakia
 e-mail: {martin.klimo, ondrej.such}@fri.uniza.sk

Abstract—Today’s digital computers are based on three cornerstones: von Neumann architecture, Boolean algebra, and transistor as the basic element. With a history of approximately 70 years, this concept has demonstrated to be a success for algorithmic computing. However, at present, its disadvantages can be seen in real-time cognitive computing. This contribution presents the concept in which cognitive computing acts as a support for algorithmic computing, and the cognitive part is based on non-von Neumann architecture, Zadeh fuzzy logic, and resistive switch as the basic element.

Keywords—resistive switch; memristive circuits; fuzzy logic; non-von Neumann architecture; cognitive computing.

I. INTRODUCTION

Computers, as we know them, are based on three cornerstones: von Neumann architecture, Boolean logic, and transistor as a switching element. All these aspects limit computer performance to energy consumption ratio compared to the human brain. The man is not so fast in algorithmic thinking using exact terms, but he can think intuitively with the fuzzy meaning of words. Boolean values for True and False are implemented as analog values, while the brain uses spike trains. Von Neumann computer architecture has liberated programs from hardware processing circuits and stored them as data to the memory, but this transfer from the memory into the processing unit takes some time. On the other hand, the neural network in the brain is naturally a massive parallel structure, where processing and memory are both located in the same place. Since 1947, the transistor has been a principal component in computing implementation. The volatility of the transistor leads to enormous energy consumption when compared to the human brain. Non-volatile elements and spike-like computing give some promise to achieve brain efficiency in the future. However, the future does not mean replacing old architectures with new ones. It is about the coexistence of the algorithmic computing with the cognitive computing. These general trends are mirrored in recent experiments with a Central Processing Unit (CPU) support by hardware accelerators, and, in this paper, we present several examples in this regard.

Since 2008, when the HP Lab realised a memristor [1], nanotechnology has offered new ways to overcome traditional computer limits. Energy savings and higher densities can be obtained and memristor crossbar can unify the memory and the processing unit together.

Elementary circuits with resistive switches can give results for Min, Max, Avg functions in the voltage domain [2]. This idea has a significant impact on the fuzzy computer architecture. Comparing memristor-based computing with

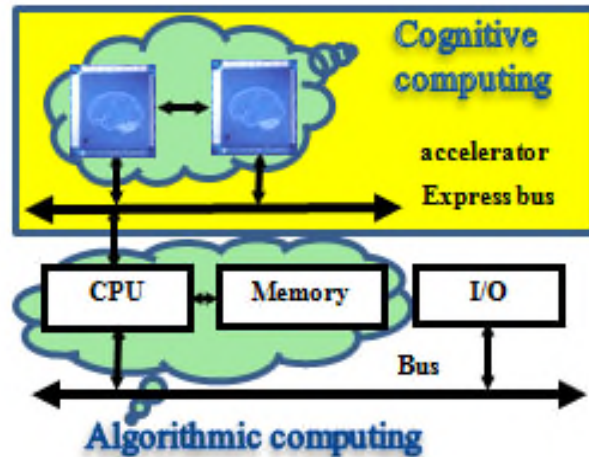


Figure 1. von Neumann architecture accelerated by non-von Neumann architecture

quantum or bio-cellular computing, the memristor technology is more mature, and several approaches (metal-oxide, ferromagnetic, grapheme-oxide) can be compared. The memristor crossbar, unlike synapse emulation and stateful Boolean implication, can provide the full range of the Zadeh logic functions in voltage input-output space. Fuzzy computing maintains continuity with classical computing where humans write programmes following the given logic requirements (implications). This is the basic approach used in fuzzy logic programming and Fuzzy Prolog-like programming languages [3], or specialised languages based on Haskell [18] e.g., Bluespec SystemVerilog [19].

The rest of the paper is structured as follows.

Section II gives the motivation for supporting algorithmic computing in von Neumann architecture by cognitive computing.

Section III supports the idea of memristive and fuzzy logic based cognitive computing [2], [10] with an experimental proof of concept by implementing elementary Zadeh fuzzy logic functions [17].

Section IV extends the proof of concept also to minterms (and analogically maxterms) in the fuzzy logic approximation by finite normal forms.

This short paper presents the progress in the memristive implementation of fuzzy cognitive computing.

II. ALGORITHMIC COMPUTING ACCELERATED BY COGNITIVE COMPUTING

Problem-solving is as old as life itself, and nature has created, by evolution, structures like the brain to do that. At the time prior to Homo sapiens, problems were less complex, but solutions had to be found in real time to save human life. Homo sapiens solved more complex problems that could be decomposed to a sequence of well-defined steps, but the process was not time critical. Among the first examples of such problem-solving would be dividing spoils in a group, growing crops and building houses. The sequence of steps that solves a problem in finite time is nowadays called an algorithm. Mainly, a history of mathematics gives an abundant supply of algorithms in numeric operations, geometry, algebra, and computer science (even for evolution). To conclude, the algorithmic solution of problems, from our perspective, is a result of human culture. Although algorithms have been implemented in the past by specific instruments (e.g., abacus, straightedge and compass, Antikythera mechanism), Turing found a universal algorithmic machine that von Neumann implemented as a digital computer. While the brain has the ability to solve problems through the topology of the neuronal network, this ability in the von Neumann computer is ascribed to the program (data). Until recently, it seemed that the von Neumann computer could solve all problems solvable by the brain. Today, we can see bottlenecks in the von Neumann architecture (transfer of data between the memory and CPU) and the advantages of programming by topology (naturally massive parallelism). As Amdahl’s law [20] pointed out, even a small part of a program serially executed can suppress the advantage of the parallel connection of CPUs. Therefore, naturally massive parallel computation is needed for time-critical applications. As mentioned before, nature has found such a structure through evolution in which the circuit topology gives the program. We call it cognitive computing in this paper. This approach was applied in the ENIAC and analog computers, but it was forgotten due to the low flexibility of programming compared to the program stored as data. Even if today there is a large demand for massively parallel computing, we do not think that cognitive computing will replace algorithmic computing. From our perspective, algorithmic computing has the same importance as cognitive computing, but they have different missions. They have to be combined: algorithmic computing with von Neumann architecture should be supported by cognitive computing executed by non-von Neumann architecture (see Fig. 1). This is not a new approach, and CPUs supported by accelerators (e.g., Intel Xeon Phi coprocessor) contain a well-known architecture.

While the von Neumann architecture is a well-established concept for algorithmic computing that has been improving over the past sixty years, it is still an open question as to which concept and which inorganic technology is the best for non-von Neumann architecture. Game applications showed graphics bottlenecks, and Graphical Processing Units (GPU) have been developed to overcome this. But the potential of a GPU is much broader, and GPUs are used as accelerators for

CPUs. Dell added NVIDIA GPU coprocessors and Tesla K80 to accelerate Intel Xeon CPUs in PowerEdge servers [4]; SGI has done the same in SGI Rackable Servers [5]. If general purpose GPU (GPGPU) has its roots in graphics processing, the Field-Programmable Gate Array (FPGA) has its roots in the Digital Signal Processor (DSP). As the name FPGA indicates, programming means creating digital circuit topology within a set of gates that performs the goal of the program. All gates work in a naturally massively parallel way and can provide cognitive computing in real time. The main goal of the von Neumann part of this hybrid architecture is to configure/reconfigure FPGA digital circuit topology to perform cognitive computing. Microsoft has studied FPGAs as accelerators under the project Catapult since 2010 [6], and the last results in “Configurable Cloud” architecture were published [7]. The acquisitions of Altera by Intel or Xilinx by IBM have shown movement in the same direction [8]. IBM has also developed a different kind of accelerator. They introduced TrueNorth as an accelerator [9] in which pulses are switched over the crossbar. Pulses run over the crossbar in parallel, which allows building a naturally massive parallel system from crossbars.

This paper presents the concept of a coprocessor built on fuzzy logical circuits implemented by the memristive structure. As an idea, it was published at this conference in 2012 [10]. According to Fig. 1, it can be redrawn as presented in Fig. 2.

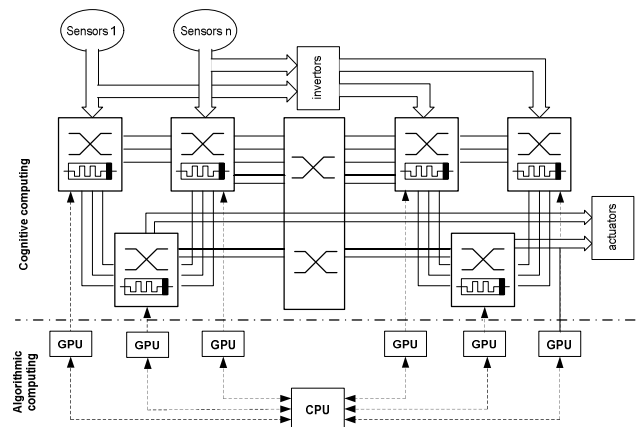


Figure 2. von Neumann architecture accelerated by memristive based cognitive computation

This paper presents an experimental proof of this concept on the level of the elementary fuzzy logic functions: minimum (Min) and maximum (Max).

III. IMPLEMENTATION OF FUZZY LOGIC BY THE MEMRISTIVE CIRCUIT

Hardware implementation of Min, Max functions is not new and fuzzy logic circuits based on Complementary Metal–Oxide–Semiconductor (CMOS), Field-Effect Transistor (FET) or FPGA implementations have been used before. However, there are two principal advantages of memristor-based implementations:

1. Energy supplies the circuit only over inputs, and no extra source of energy is needed to be compared with the transistor based implementations mentioned above.
2. The memristive implementation also introduces a memory function to these elementary functions. This property needs further research.

Fig. 3 shows an example of the implementation of $Y = \text{Max}(0, X)$ function using electrochemical metallization memory (ECM) resistive switches NEURO-BIT BT10001B14 [11]. To interpret this figure with respect to fuzzy logic, the input X after normalisation from $\langle -1.5V, 1.5V \rangle$ interval into $\langle -1, 1 \rangle$ interval represents the difference $X = a - b$ in $y = \text{Max}(a, b)$; $a, b \in \langle 0, 1 \rangle$ function. The accuracy of the mathematical function implementation depends mainly on the switching threshold (approximately 0.2V for the measured resistive switches), and measurement repeatability.

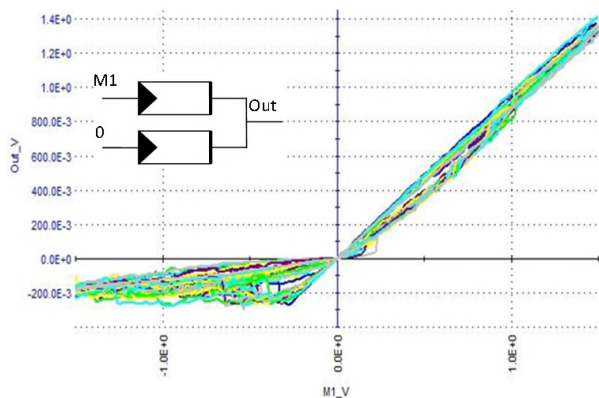


Figure 3. The input/output characteristic of the Max circuit implemented by ECM memristors NEURO-BIT BT10001B14

On the one hand, non-volatility is useful, but on the other hand, the preservation of the switch state causes a memorylessness in the input – output relationship in the Max circuit. More precisely, fuzzy logic circuits have to be assumed as state automata. Everything mentioned above regarding the implementation of Max functions is also valid for Min functions.

However, Min, Max functions allow only the building of a monotone fuzzy logic system [12]. In general, inverters are needed, but they cannot be implemented as a passive element by resistive switches. Here, we present an architecture using the property of the de Morgan’s law that a dual logic function with inverted inputs results in an inverted function. This dual system approach allows for simpler implementation because inverters are located only in the first stage of the Min/Max based circuit, and can be implemented by active elements.

IV. IMPLEMENTATION OF THE FUZZY LOGIC FUNCTION IN A NORMAL FORM

Our experience with memristive-based implementation of fuzzy logic functions shows [13] that in deep memristive networks, there pairs of states with no direct transition between them may occur. This is caused mainly by switching thresholds, and this phenomenon should be studied in the future. For the moment, flat memristive networks can be

instrumental in fuzzy logic implementation. This flat topology means a structure corresponding to the normal form within the classical Boolean logic (BL) algebra. But disjunctive or conjunctive finite normal forms are universal approximation formulas for any BL-algebra [14].

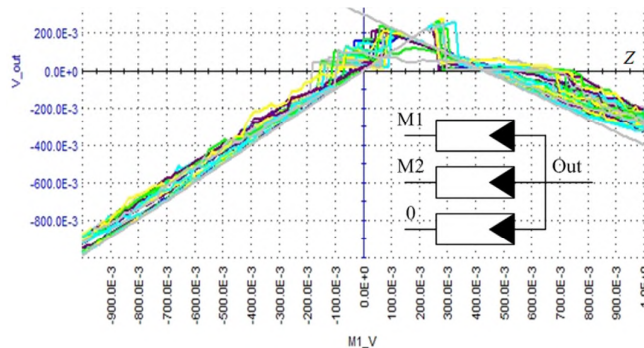


Figure 4. The input/output characteristic of the 3-input Min circuit implemented by ECM memristors NEURO-BIT BT10001B14

Max and Min functions implement disjunctions and conjunctions in the Zadeh fuzzy logic with more inputs. It is assumed again that input variables and their negations are available as inputs.

Fig. 4 shows an example of the implementation of the function

$$Y = \text{Min}(0, X1, X2).$$

As the independent input is taken $M1 = X1$, $X1 \in \langle -1, 1 \rangle$, the second input is set into

$$M2 = 1 - 0.7(1 + X1),$$

and the third one is a zero reference input. The impact of the switching threshold is visible here even more so than in Figure 2.

V. CONCLUSION AND THE FUTURE SCOPE

The paper sets forth the idea of how cognitive massively parallel computing based on the Zadeh fuzzy logic can be implemented using memristive circuits. These results support an idea of cognitive computing based on fuzzy logic accelerators implemented via memristive circuits.

Looking at Fig. 3 and Fig. 4, we can see the importance of the non-deterministic behaviour of memristors and the influence of the switching threshold. The impact of these properties on the accuracy of fuzzy memristive computing needs further research. Another extension of the field of research would be the large area of non-fuzzy memristive computing [15], and applications of memristive circuits [16].

ACKNOWLEDGMENT

APVV-14-0560 grant funded this research.

REFERENCES

- [1] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature* 453, pp. 80-83, May 2008, doi:10.1038/nature06932.
- [2] M. Klimo and O. Such, "Memristors can implement fuzzy logic," <http://arxiv.org/abs/1110.2074>.
- [3] S. Guadarrama, S. Muñoz, and C. Vaucheret, "Fuzzy Prolog: a new approach using soft constraints propagation," *Fuzzy Sets and Systems*, vol. 144, Issue 1, pp. 127-150, May 2004.
- [4] S. Iqbal, M. Deshmukh, and N. Dandapanthula, "Game-changing Extreme GPU computing with The Dell PowerEdge C4130," A Dell Technical White Paper, Dell HPC Engineering, May 2015.
- [5] "SGI with NVIDIA GPU Accelerators Complete Solutions", Data Sheet, www.sgi.com, Accessed September 30, 2016.
- [6] E. S. Chung, J. C. Hoe, and K. Mai, "CoRAM: An In-Fabric Memory Architecture for FPGA-based Computing," International Symposium on Field-Programmable Gate Arrays (FPGA), Monterey, CA, pp. 97-106, Feb. 2011, doi:10.1145/1950413.1950435
- [7] Project Catapult, <https://www.microsoft.com/en-us/research/project/project-catapult/>, Accessed October 17, 2016.
- [8] G. Lacey, G. Taylor, and S. Areibi, "Deep Learning on FPGAs: Past, Present, and Future", arXiv:1602.04283 [cs.DC].
- [9] P. A. Merolla et al., "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, Issue 6197, pp. 668-673, Aug. 2014, doi:10.1126/science.1254642.
- [10] M. Klimo and O. Such, "Fuzzy Computer Architecture Based on Memristor Circuits," Proc. FUTURE COMPUTING 2012, pp. 84-87, July 2012, Nice, France, ISBN: 978-1-61208-217-2.
- [11] "Neuro-bit user manual," <http://www.bioinspired.net/>, Accessed April 19, 2016.
- [12] M. Blum, R. Rue, and K. Yang, "On the complexity of MAX/MIN/AVRG circuit," Carnegie Mellon Technical Report CMU-CS-02-110, 2002, <http://repository.cmu.edu/compsci/150>
- [13] L. Nielen et al., "Memristive Sorting Networks Enabled by Electrochemical Metallization Cells," *International Journal of Unconventional Computing*, vol. 12, No. 4, pp. 303-317, 2016, ISSN: 1548-7199 (print), ISSN: 1548-7202 (online).
- [14] I. Perfilieva, "Normal forms for fuzzy logic functions," Proc. 33rd Int. Symp. Multi-valued logic, ISMVL'03, pp. 59-64, 2003, doi:10.1109/ISMVL.2003.1201385.
- [15] I. Vourkas and G. Ch. Sirakoulis, "Memristor-Based Nanoelectronic Computing Circuits and Architectures", Springer, 2016, ISBN 978-3-319-22646-0.
- [16] A. Chens, J. Hutchby, V. Zhirnov, and G. Bourianoff, editors, "Emerging Nanoelectronic Devices", Wiley, 2015, ISBN 978-1-118-44774-1.
- [17] L. A. Zadeh, "Fuzzy logic and its application to approximate reasoning," In *Information Processing 74*, Proc. IFIP Congr. 1974 (3), pp. 591-594.
- [18] S. Thompson, "Haskell: The Craft of Functional Programming," Addison-Wesley, 1996.
- [19] Y. S. Shao and D. Brooks, "Research infrastructures for hardware accelerators," *Synthesis Lectures on Computer Architecture*, 2015, Vol. 10, No. 4, pp. 1-99, Morgan & Claypool Pub, doi: 10.2200/S00677ED1V01Y201511CAC034.
- [20] G. M. Amdahl, "Validity of the Single Processor Approach to Achieving Large-Scale Computing Capabilities," AFIPS Conf. Proc. (30), 1967, pp.483-485, doi:10.1145/1465482.1465560

Lyapunov-based Control Theory of Closed Quantum Systems

Shuang Cong

Department of Automation, University of Science and
Technology of China
Hefei, China
e-mail: scong@ustc.edu.cn

Fangfang Meng

Department of electronic information and electrical
engineering, Hefei University
Hefei, China
e-mail: mengffo@hfu.edu.cn

Abstract—According to whether the internal Hamiltonians are strong regular and/or the control Hamiltonians are full connected, the quantum systems can be considered as ideal closed quantum systems or the quantum systems in degenerate cases. In this paper, we propose a unique formation of quantum Lyapunov-based control method, which is suitable for both ideal closed quantum systems and the systems in degenerate cases. This Lyapunov-based control method of closed quantum systems with unique formation is realized by means of introducing implicit Lyapunov functions into the control laws, which make the control system become strong regular and full connection. The proposed Lyapunov-based control theory can transfer from arbitrary initial states to arbitrary final states in the way of asymptotic stability. The paper gives the complete design procedure of control laws. At last, a numerical experiment of the state transfer between two mixed states in degenerate cases is given to demonstrate the effectiveness of the Lyapunov-based control theory proposed.

Keywords- ideal closed quantum systems; quantum systems in degenerate cases; implicit Lyapunov-based control method.

I. INTRODUCTION

From the perspective of system control, a quantum system can be considered as a closed or an open quantum system. The closed quantum system is an isolated system or without interaction with the environment. The majority of actual quantum systems are open quantum systems. However, the closed quantum systems have their own characteristics, namely, they are simpler to be analyzed and studied, and the research results of closed quantum systems are the foundations of open quantum systems. The role of a closed system in quantum systems is similar to that of the system which is a linear, definite and time-invariant in macroscopic systems. Even so, the control task of state transfer in closed quantum systems is quite difficult because there are eigenstates, superposition states and mixed states, in which only the eigenstate corresponds to the classical state in

macroscopic systems, while other two states do not exist in the macroscopic world.

The solutions of the control problems obtained by means of the system control theory are generally the control laws in an N dimensional quantum system, which can be easily applied to the high dimensional quantum systems without increasing control cost and design difficulty. Therefore, the closed quantum system control theory has a guiding significance for the realization of the actual experiments, especially for complex quantum systems. In the last 30 years, the control theory of quantum systems has developed rapidly. Many quantum control methods have been developed, such as coherent control [1]-[3], Bang-bang control and geometrical control [4][5], dynamical decoupling control [6]-[8], sliding mode control [9][14], robust control [10], optimal control [11]-[14], Lyapunov-based control [15]-[18], feedback control [19]-[21]. Among all the quantum control theories, optimal quantum control is the most widely used in quantum system control. Like the optimal quantum control method, the Lyapunov-based quantum control is also a very powerful control method. By means of the Lyapunov stability theorem, this control method designs an asymptotically stable controller by making the first time derivative of the Lyapunov function constructed not great than zero. Unlike the way it is being used in the macroscopic engineering field, in which the controller is only required to be designed as a stable one, the Lyapunov-based control method used in quantum fields should be designed as a convergent one in order to guarantee the control system to reach the target state with 100% probability. This is because the variable controlled in quantum systems is usually the density matrix, which is a probability. A general model of an N dimensional closed quantum system can be described by the Liouville equation: $i\dot{\rho}(t) = [H_0 + \sum_{k=1}^r H_k u_k(t), \rho(t)]$, in which $\rho(t)$ is the density matrix; H_0 is the internal Hamiltonian; $u_k(t)$ are external control fields; H_k are the control Hamiltonians. The eigenvalue (or spectrum) of the internal Hamiltonian

$H_0 = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ of the system, in which $\lambda_j, j = 1, 2, \dots, N$ indicate the energy levels of the system, while $\omega_{jl} = \lambda_j - \lambda_l$ are the transitions (Bohr) frequencies between the energy levels of the system. We define a non-degenerate quantum system if all the energy levels of a quantum system are not the same and a quantum system without degenerate transition, which means all Bohr frequencies are not the same. A quantum system is called full connection if $(H_k)_{jk} \neq 0, j < l \in \{1, 2, \dots, N\}$ for $\forall k \in \{1, 2, \dots, N\}$ holds. Based on the Lyapunov stability theorem, the analytical expressions of control laws can be designed by means of the construction of a suitable Lyapunov function $V(\rho)$, and under the condition of $\dot{V}(\rho) < 0$. The Lyapunov function is not unique. A general form of Lyapunov function is $V(\rho) = \text{tr}(P\rho)$, in which P is a positive definite Hermitian operator to be determined, which is one part of control laws design. P can be regarded as an imaginary mechanical value of the system. In mathematics, $V(\rho) = \text{tr}(P\rho)$ is a trace calculation. In physics, $V(\rho)$ is an expectation value of Hermitian operator P . By calculating the first order derivatives of $V(\rho)$ for the time, one can obtain $\dot{V}(\rho) = -i\text{tr}([P, H_0], \rho) - i\sum_{k=1}^r \text{tr}([P, H_k] \rho) u_k$. Because the first term in the right side of $\dot{V}(\rho)$ is independent of the control laws this term can be eliminated by $[P, H_0] = 0$, which also provides a condition of designing P . When H_0 is non-degenerate, P is a diagonal matrix. The control laws can be obtained by letting $\dot{V}(\rho) = 0$, and the expressions of control laws are $u_k = i\varepsilon_k \text{tr}([P, H_k] \rho), k = 1, 2, \dots, r$, in which ε_k is used to regulate the amplitude of the control laws. According to the LaSalle invariant set principle, the control system can be guaranteed that any trajectory converges to a maximum invariant set.

Now that the control laws is obtained by $\dot{V}(\rho) = 0$, besides the target state, generally there are many other states which can also make $\dot{V}(\rho) = 0$, all of which are the state points of the Lyapunov function $V(\rho)$. The number of these state points is even un-numerical in the cases when the target state is a supposition state or mixed state. In order to make the control system converge to the desired target state, one must add the constraint conditions to narrow the invariant set. For the different kinds of the target state, the conditions the system needs to meet are different. Generally speaking, the convergence conditions of a quantum system by using the Lyapunov control method based on the average value of an imaginary mechanical quantity P are three points, which are the requirements of internal Hamiltonian, control Hamiltonians, and target state, respectively. These three conditions are:

- i) The internal Hamiltonian is strongly regular, i.e., the transition energies between two different levels are clearly identified;

- ii) The control Hamiltonians are full connected, i.e., any two levels are directly coupled [18];
- iii) The target state must be diagonal, which makes $[\rho_f, H_0] = 0$ hold.

Fig. 1 is an example that satisfies the conditions i) and ii).

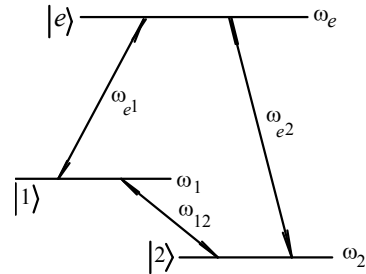


Figure 1. Quantum system that satisfies the conditions i) and ii).

The system which satisfies above mentioned conditions i) and ii) is called ideal quantum system. Under the conditions i) and ii), condition iii) is the condition of the state transfer of closed quantum systems from arbitrary initial state to an arbitrary diagonal target state, which can be an eigenstate, supposition state, or mixed state.

However, many quantum systems in practice do not satisfy the conditions i) or/and ii). For example,

$$H_0 = \begin{bmatrix} 0.3 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.9 \end{bmatrix}$$

or/and

$$H_1 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Fig. 2 shows that the V-type and Λ -type quantum systems we often encounter in practice do not satisfy conditions i) and ii). Because the convergence conditions obtained are so extremely rigorous, the designed control laws have little practical application value.

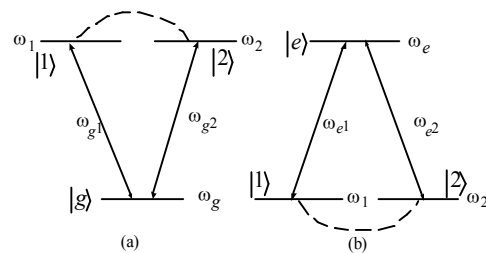


Figure 2. (a) V-type; (b) Λ -type quantum systems.

In order to obtain Lyapunov-based quantum control methods which have practical application value, one needs to solve the problems which appear in the quantum systems in degenerate cases, so as to establish the Lyapunov quantum control theory. Up to now, many researches have been done. Zhao et al. utilized an implicit Lyapunov control to solve the problem of convergence for the single control Hamiltonian systems governed by the Schrödinger equation [22]. We once investigated the implicit Lyapunov control method of multi-control Hamiltonian systems governed by the Schrödinger equation based on the state distance and state error, both of which are only suitable for the control of the pure states [23]. We also studied the implicit Lyapunov quantum control method based on the imaginary mechanical quantity for pure states and mixed states, in which the stricter convergence proof was given [24].

In this paper we propose a unified formulation of Lyapunov control theory for closed quantum systems. The basic idea is: for a quantum system which does not satisfy the convergent conditions i), ii) or/and iii), we introduce an implicit Lyapunov function into the quantum control design method based on the imaginary mechanical quantity in Liouville equation, in order to make the designed control system satisfy three convergent conditions. The Lyapunov quantum control theory proposed here can be used in both degenerate cases and ideal quantum systems, which can transfer the state from an arbitrary initial state to an arbitrary target state. The “arbitrary” here means eigenstate, superposition state or mixed state. The proposed control method in this paper is a unique formation of quantum Lyapunov-based control method, which has important significance.

The rest of the paper is structured as follows: Section II is the Lyapunov-based quantum control theory, in which implicit Lyapunov functions are introduced, as well as the procedure of control designs of η_k , $\gamma_k(t)$, $v_k(t)$ and $P_{\eta\gamma}$ in detail. Section III is the numerical simulation, and Section IV is the conclusion.

II. LYAPUNOV-BASED QUANTUM CONTROL THEORY

A. Implicit Lyapunov Functions

Consider the N -level closed quantum systems governed by the following quantum Liouville equation which may be in degeneration cases:

$$i\dot{\rho}(t) = [H_0 + \sum_{k=1}^r H_k u_k(t), \rho(t)] \quad (1)$$

where $\rho(t)$ is the density operator; H_0 is the internal Hamiltonian; H_k , $k=1,2,\dots,r$, are control Hamiltonian; and $u_k(t)$, ($k=1,\dots,r$) are scalar and real total control laws.

The way to solve the degeneration problems is to introduce the implicit Lyapunov functions as the control disturbances such that the system with additional control disturbances may

satisfy those convergence conditions. A completely unified designing method of control laws is proposed here. The control laws are composed of three parts:

$$u_k(t) = \gamma_k(t) + v_k(t) + \eta_k \quad (2)$$

in which $\gamma_k(t)$ are designed to make the system (1) satisfy the convergence conditions i) and ii); $v_k(t)$ are the control laws designed to transfer any initial state to the invariant set; η_k are used to make the target state commute with the internal Hamiltonian H_0 , i.e., $[\rho_f, H_0] = 0$, so as to make the control system be able to converge to the desired target state.

The Lyapunov function is constructed as:

$$V(\rho) = \text{tr}(P_{\eta\gamma} \rho) \quad (3)$$

where $P_{\eta\gamma} = f(\eta_1, \dots, \eta_r, \gamma_1(t), \dots, \gamma_r(t))$ are functional of η_k and $\gamma_k(t)$, ($k=1,2,\dots,r$) and positive definite.

Eq. (3) is called the implicit Lyapunov function based on the average value of an imaginary mechanical quantity. The function of (1) is used to design control laws (2), in which η_k will be designed in the case $[\rho_f, H_0] \neq 0$, which does not satisfy the condition iii). $\gamma_k(t)$ will be designed in the cases when condition i) or/and ii) are not satisfied. $v_k(t)$ is used to design the control laws of transferring the state from an arbitrary initial state to an arbitrary target state.

Next, we'll give in detail the design procedures and the explanations of how these three parts play roles in control laws.

B. Control Design of η_k

In the procedure of designing η_k , first, check whether the target state ρ_f commutes with the internal Hamiltonian H_0 , and one can know what type the target state is. If the target state does not commute with the internal Hamiltonian, this results in $[\rho_f, H_0] = D$, $D \neq 0$. The supposition state and some non-diagonal mixed state are in such cases. Then a set of appropriate constant values η_k need to be introduced into the control laws. Then, $H_{0\eta} = H_0 + \sum_{k=1}^r H_k \eta_k$, ($k=1,2,\dots,r$) will be considered to be the new internal Hamiltonian. Last, design η_k in order to make

$$[\rho_f, H_{0\eta}] = 0, H_{0\eta} = H_0 + \sum_{k=1}^r H_k \eta_k \quad (4)$$

hold.

For the special case when the target state commute with the internal Hamiltonian H_0 , that is, $[\rho_f, H_0] = 0$, one can set

$\eta_k = 0$, which is the quantum system with the target state being eigenstates or some mixed states which commute with H_0 . After introducing and designing the constant values η_k , the target state ρ_f will become commute with $H_{0\eta}$ in (4).

C. Control Design of $\gamma_k(t)$

There are two objectives of designing $\gamma_k(t)$, one is to make $H_{0\eta\gamma} = H_{0\eta} + \sum_{k=1}^r H_k \gamma_k(t)$ such that $H_{0\eta\gamma}$ is strongly regular. Denote eigenstate of $H_{0\eta\gamma}$ as $|\phi_{1\eta,\gamma}\rangle, \dots, |\phi_{N\eta,\gamma}\rangle$. The control Hamiltonian with $H_{0\eta\gamma}$ is $H_{k\eta\gamma}$: $H_{k\eta\gamma} = U_{\eta\gamma}^\dagger H_k U_{\eta\gamma}$, in which $U_{\eta\gamma} = (|\phi_{1\eta,\gamma}\rangle, \dots, |\phi_{N\eta,\gamma}\rangle)$. Another objective is to make the $H_{k\eta\gamma}$ be full connected. To achieve these two objectives, $\gamma_k(t)$ can be designed as

$$\gamma_k(t) = \gamma(t) = \begin{cases} F(s), & k = k_1, \dots, k_m \\ 0, & k \neq k_1, \dots, k_m (1 \leq k_1, \dots, k_m \leq r) \end{cases} \quad (5)$$

in which F is the function of s , and satisfies $F(0) = 0$, $F(s) > 0$, and $F'(s) > 0$, which means $F(s)$ is a monotonic increasing function.

Usually, the simplest $F(s)$ can be constructed as: $s = V(\rho) - V(\rho_f)$, where $C > 0$, and $C \in R$. Combining with (3) $\gamma_k(t)$ can be designed as

$$\gamma_k(t) = \gamma(t) = C \cdot (\text{tr}(P_{\eta\gamma} \rho) - \text{tr}(P_{\eta\gamma} \rho_f)) \quad (6)$$

D. Control Design of $v_k(t)$

The role of control laws $v_k(t)$ is to ensure $\dot{V}(t) \leq 0$. $v_k(t)$, $k = 1, \dots, r$ are designed to ensure the first time derivative of Lyapunov function (3) is not greater than zero, from which we can obtain:

$$v_k(t) = K_k f_k \left(\text{itr}([P_{\eta\gamma}, H_{k\eta\gamma}] \rho) \right) \quad (7)$$

where K_k are constants and $K_k > 0$, $k = 1, \dots, r$, $H_{k\eta\gamma} = U_{\eta\gamma}^\dagger H_k U_{\eta\gamma}$, $U_{\eta\gamma} = (|\phi_{1\eta,\gamma}\rangle, \dots, |\phi_{N\eta,\gamma}\rangle)$ and $y_k = f_k(x_k)$, ($k = 1, 2, \dots, r$) are monotonic increasing functions which are through the coordinate origin of the plane $x_k - y_k$.

In fact, LaSalle invariant principle can only guarantee the control system to converge to the invariant set, but not guarantee to converge to the target state. In order to make the control system converge to the target state, we still need to deal with another problem: the number of critical states in the

invariant set, i.e., the number of the states which satisfy $\dot{V}(t) = 0$. For a closed quantum system, only when the target state commutes with the internal Hamiltonian, the number of critical states in the invariant set is at most $N!$. There are un-numerical critical states in the invariant set when the target does not commute with the internal Hamiltonian. This problem can be solved in two ways: one is to make the un-numerical critical states in the invariant set become numerical ones by introducing a set of constant values η_k into the control laws; another is to make the target state be the minimum value of the Lyapunov function (3) by designing the imaginary mechanical quantity.

The control laws (2) designed by (6) and (7) can only guarantee $\dot{V}(t) \leq 0$. In order to ensure $\dot{V}(t) < 0$, we provide another condition

$$V(\rho_f) < V(\rho_{other}) \quad (8)$$

which means the value of the Lyapunov function at the target state is less than the values of Lyapunov function at all other states.

The role of $P_{\eta\gamma}$ in (6) is to make the control system converge to the target state ρ_f . In order to do so, on one hand, we need to design $P_{\eta\gamma}$ to make the condition (8) hold, where ρ_{other} represents any other critical states in the invariant set except the target state. On the other hand, the condition $[P_{\eta\gamma}, H_{0\eta\gamma}] = 0$ must hold, which means that $P_{\eta\gamma}$ and $H_{0\eta\gamma}$ have the same eigenstates $|\phi_{1\eta,\gamma}\rangle, \dots, |\phi_{N\eta,\gamma}\rangle$. We design the eigenvalues of $P_{\eta\gamma}$ to be constant, denoted by P_1, P_2, \dots, P_N , and design $P_{\eta\gamma}$ as

$$P_{\eta\gamma} = \sum_{j=1}^N P_j |\phi_{j\eta,\gamma}\rangle \langle \phi_{j\eta,\gamma}| \quad (9)$$

In order to make (8) hold, we design P_j as follows: In

$(\rho_{f\eta})_{ii} < (\rho_{f\eta})_{jj}, 1 \leq i, j \leq N$, design $P_i > P_j$;

if $(\rho_{f\eta})_{ii} = (\rho_{f\eta})_{jj}, 1 \leq i, j \leq N$, design $P_i \neq P_j$; else if $(\rho_{f\eta})_{ii} > (\rho_{f\eta})_{jj}, 1 \leq i, j \leq N$, design $P_i < P_j$,

then $V(\rho_{f\eta}) < V(\rho_{other})$ holds, where $(\rho_{f\eta})_{ii}$ is the (i, i) -th element of $\rho_{f\eta} = U_{\eta}^\dagger \rho_f U_{\eta}$; $U_{\eta} = (|\phi_{1,\eta}\rangle, \dots, |\phi_{N,\eta}\rangle)$; $|\phi_{1,\eta}\rangle, \dots, |\phi_{N,\eta}\rangle$ are the eigenstates of $H_{0\eta} = H_0 + \sum_{k=1}^r H_k \eta_k$.

For the above deduction, refer to the proof of Theorem 2 in [24].

Based on LaSalle's invariance principle, the convergence of the control system with above control laws designed by $u_k(t) = \gamma_k(t) + v_k(t) + \eta_k$ in (2), we proven the following

theorem: Consider the control system (1) and the constructed Lyapunov function (3), under the action of control laws (2), in which $\gamma_k(t)$ is designed by (6); $v_k(t)$ is designed by (7); η_k is used to make (4) hold, $P_{\eta\gamma}$ is designed as (9), which can make the control system satisfy:

- i) $\omega_{l,m,\eta\gamma} \neq \omega_{i,j,\eta\gamma}$, $(l,m) \neq (i,j)$,
 $i, j, l, m \in \{1, 2, \dots, N\}$, $\omega_{l,m,\gamma} = \lambda_{l,\eta\gamma} - \lambda_{m,\eta\gamma}$, where $\lambda_{l,\eta\gamma}$ is the l -th eigenvalue of $H_{0\eta\gamma} = H_0 + \sum_{k=1}^r H_k(\eta_k + \gamma_k(t))$ corresponding to the eigenstate $|\phi_{l,\eta\gamma}\rangle$;
- ii) $\forall j \neq l$, for $k = 1, \dots, r$, there exists at least a $(H_{k\eta\gamma})_{jl} \neq 0$, where $(H_{k\eta\gamma})_{jl}$ is the (j,l) -th element of $H_{k\eta\gamma} = U_{\eta\gamma}^\dagger H_k U_{\eta\gamma}$ with $U_{\eta\gamma} = (|\phi_{1,\eta\gamma}\rangle, \dots, |\phi_{N,\eta\gamma}\rangle)$;
- iii) $[P_{\eta\gamma}, H_{0\eta\gamma}] = 0$; For any $l \neq j, (1 \leq l, j \leq N)$, $(P_{\eta\gamma})_{ll} \neq (P_{\eta\gamma})_{jj}$ holds, where $(P_{\eta\gamma})_{ll}$ is the (l,l) -th element of $P_{\eta\gamma}$, and the control system will converge toward the invariant set E :

$$E = \left\{ \rho_\gamma(t_0) \mid (U_{\eta\gamma}^\dagger \rho_\gamma(t_0) U_{\eta\gamma})_{ij} = 0, \gamma = \gamma(\rho_\gamma(t_0)), t_0 \in \mathbb{R} \right\} \quad (10)$$

The proof of the theorem is similar to the proof in [24], and we will not repeat it here.

By designing the control laws proposed in this paper, a quantum system in degenerate case can become an ideal quantum system, which satisfies three convergent conditions of state transfer. In fact, the proposed control designed method in this paper is also suitable for the state transfer of ideal quantum systems, so up to now we establish a complete Lyapunov - based closed quantum control theory, which is not only suitable for quantum systems in non-degenerate, but also suitable for the quantum systems in degenerate cases.

III. NUMERICAL SIMULATION

In this section, we perform an experiment to design a specific controller to transfer a state to a superposition state by using the implicit Lyapunov control based on the average value of an imaginary mechanical quantity.

Consider a 3-level quantum system, whose internal Hamiltonian is non-strong regular, and the control Hamiltonians are not full connected:

$$H_0 = \begin{bmatrix} 0.3 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.9 \end{bmatrix}, H_1 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (11)$$

In the numerical simulation experiment, the initial state ρ_0 is a mixed state which does not commute with the

internal Hamiltonian and the target ρ_f is a mixed state which commutes with the internal Hamiltonian:

$$\rho_0 = \begin{bmatrix} 0.1 & 0.1 & 0.04 \\ 0.1 & 0.5 & 0.08 \\ 0.04 & 0.08 & 0.4 \end{bmatrix} \quad (12)$$

$$\rho_f = \text{diag}\{0.5687, 0.3562, 0.075\}$$

According to the design ideas proposed in this paper, the control laws are $u_1(t) = \gamma_1(t) + v_1(t)$, in which $v_1(t)$ is designed as:

$$v_1(t) = K_1 \left(\text{itr}([P_{\gamma_1}, H_1] \rho) \right) \quad (13)$$

in which K_1 is the gain of $v_1(t)$, and $K_1 > 0$.

The implicit function $\gamma_1(t)$ is designed as:

$$\gamma_1(\rho) = M_1 \cdot (\text{tr}(P_{\gamma_1} \rho) - \text{tr}(P_{\gamma_1} \rho_f)) \quad (14)$$

where M_1 is the gain of $\gamma_1(t)$, and $M_1 > 0$.

According to the design method of the imaginary mechanical quantity in (9), design the eigenvalues of P_{γ_1} are:

$$P_1 < P_2 < P_3, \quad P_{\gamma_1} = \sum_{j=1}^3 P_j |\phi_{j,\gamma_1}\rangle \quad (15)$$

where $|\phi_{j,\gamma_1}\rangle$ is the eigenstates of $H_0 + \sum_{k=1}^r H_k \gamma_k(t)$.

In the simulation experiment, the simulation step is set to be 0.01 a.u., and control duration is 300 a.u.. The parameters used in experiment are: $M_1 = 0.1$, $K_1 = 0.34$, $P_1 = 0.01$, $P_2 = 2$ and $P_3 = 2.9$. The results of numerical simulating experiments are shown in Fig. 3 and Fig. 4. Fig. 3 represents the evolution curves of density metrics, in which ρ_{ii} is the diagonal elements of ρ . Fig. 4 shows the control curves of the $\gamma_1(t)$, $v_1(t)$ and $u_1(t)$.

From Fig. 3 and Fig. 4, one can see that at the time 300 a.u., $\rho_{11} = 0.56811$, $\rho_{22} = 0.35215$, $\rho_{33} = 0.07973$, and transfer probability is 99.53%, which verifies the effectiveness of the proposed method in this paper.

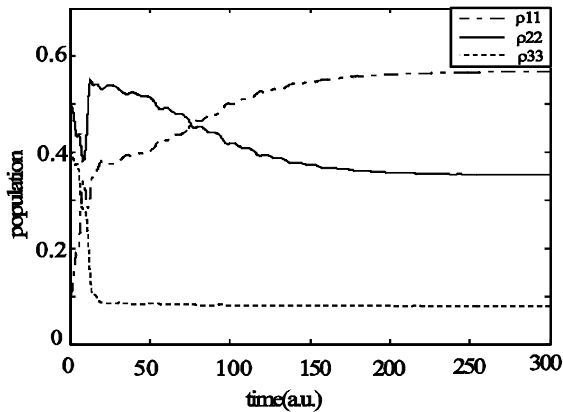


Figure 3 Evolution curves of density metrics ρ_{ii} .

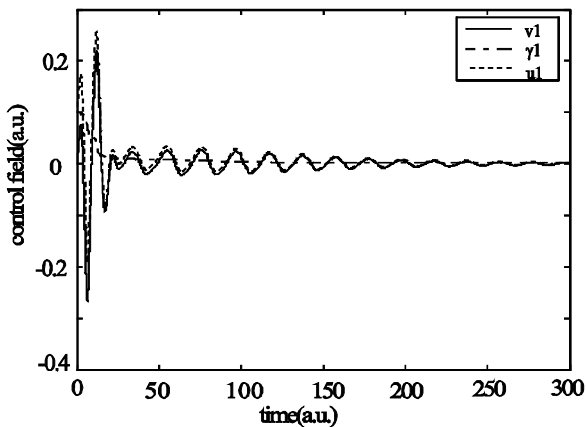


Figure 4 Control fields of the control system.

IV. CONCLUSION

This paper proposed a complete design procedure of control laws for closed quantum systems in degenerate cases. The proposed control design method is also suitable for ideal quantum systems. Based on the Lyapunov-based control theory of quantum systems proposed in this paper, the state transfer task of closed quantum systems from arbitrary initial state to arbitrary final state can be completed, and the Lyapunov-based control theory of closed quantum systems has been established.

ACKNOWLEDGMENT

This work was supported partly by the National Science Foundation of China under Grant No. 61573330, and the Hefei University Foundation for Talents No. 0394840903.

REFERENCES

[1] C. Altafini, "Coherent control of open quantum dynamical systems", *Physical Review A*, 70, pp. 062321-1-062321-8, 2004
 [2] W. S. Warren, H. Rabitz, and M. Dahleh, "Coherent control of quantum dynamics: The dream is alive", *Science*, 259, pp. 1581-1589, 1993

[3] S. Cong and Y. Lou, "Coherent control of Spin 1/2 quantum systems using phases", *Control Theory and Applications*, vol. 25, Feb. 2008, 187-192
 [4] W. C. Cao, X. S. Liu, and H. B. Bai, *Science in China Series G: Physics, Mechanics & Astronomy*, vol. 51, Jan. 2008, pp. 29-37.
 [5] S. Cong, J. Wen, and X. Zou, "Comparison of time optimal control for two level quantum systems", *Journal of Systems Engineering and Electronics*, vol. 15, Jan. 2014, pp. 95-103.
 [6] G. S. Uhrig, "Keeping a Quantum Bit Alive by Optimized π -Pulse Sequences", *Physical Review Letters*, vol. 98, pp. 100504-1 - 100504-4. 2007
 [7] L. Viola, E. Knill and S. Lloyd, "Dynamical Decoupling of Open Quantum Systems", *Physical Review Letters*, vol. 82, Dec. 1999, pp. 2417-2421.
 [8] S. Cong, L. Chan, and J. Liu, "An Optimized Dynamical Decoupling Strategy To Suppress Decoherence", *International Journal of Quantum Information*, vol. 9, pp. 1599-1615, 2011
 [9] D. Dong and I. R. Petersen, "Sliding mode control of quantum systems", *New Journal of Physics*, vol. 11, pp. 105033-1-105033-18, 2009
 [10] R. J. Matthew, "Control Theory: From Classical to Quantum Optimal, Stochastic, and Robust Control", *Notes for Quantum Control Summer School*, Caltech, August, 2005.
 [11] D. D'Alessandro, "The optimal control problem on SO(4) and its applications to quantum control", *IEEE Transaction on Automatic Control*, vol. 47, Jan. 2002, pp. 87-92.
 [12] N. Khaneja, R. Brockett, and S. J. Glaser, "Time optimal control in spin systems", *Physical Review A*, vol. 63, March 2001, pp. 032308-1-032308-13.
 [13] J. P. Palao and R. Kosloff, "Optimal control theory for unitary transformations", *Phys. Rev. A*, 2003, 68, 062308.
 [14] S. Grivopoulos, *Optimal control of quantum systems*. Santa Barbara: University of California, 2005.
 [15] S. Grivopoulos and B. Bamieh, "Lyapunov-based control of quantum systems", *Proceedings of the 42nd IEEE Conference on Decision and Control*, 2003, Maui, USA: IEEE, pp. 434-438.
 [16] S. Kuang and S. Cong, "Lyapunov control methods of closed quantum systems", *Automatica*, vol. 44, Jan. 2008, pp. 98-108.
 [17] M. Mirrahimi(a), P. Rouchon, and G. Turinici. "Lyapunov control of bilinear Schrödinger equations". *Automatica*, vol. 41, pp. 1987-1994. 2005
 [18] X. Wang and S. Schirmer, "Analysis of Effectiveness of Lyapunov Control for Non-generic Quantum States", *IEEE Transactions on Automatic Control*, vol. 55, June 2010, pp. 1406-1411.
 [19] C. Altafini. "Feedback stabilization of isospectral control systems on complex flag manifolds: application to quantum ensembles". *IEEE Transactions on Automatic Control*, vol 52, Nov. 2007, pp. 2019-2028.
 [20] A. C. Doherty and K. Jacobs, "Feedback control of quantum systems using continuous state estimation", *Physical Review A*, 60, pp. 2700-2711, 1999
 [21] S. Lloyd and J. J. E. Slotine. "Quantum feedback with weak measurements", *Phys. Rev. A*, vol. 62, pp. 012307-012311, 2000
 [22] S. W. Zhao, H. Lin, J. T. Sun, and Z. G. Xue. "An implicit Lyapunov control for finite-dimensional closed quantum systems". *International Journal of Robust and Nonlinear control*, vol. 22, Issue 11, pp. 1212-1228., 2012
 [23] F. Meng, S. Cong, and S. Kuang, "Implicit Lyapunov control of multi-control Hamiltonian systems based on state distance", *Proceedings of the 10th World Congress on Intelligent Control and Automation (WCICA)*, pp. 5127-5132, 2012
 [24] S. Cong, F. Meng, and S. Kuang, "Quantum Lyapunov control based on the average value of an imaginary mechanical quantity", *Proceedings of the 19th World Congress The International Federation of Automatic Control (IFAC)*, Cape Town, South Africa, pp. 9991-9997, 2014

Performance Evaluation of DSDV and AODV Protocols for Green Corridor Management in a Metropolitan City

Sandeep Sagar Kariyappa Shivappa, Asha Hanumantharaya, Narendra Kumar Gurumurthy
 Department of Electronics and Communication Engineering, University Visvesvaraya College of Engineering,
 Bangalore University, Bangalore, India
 email:sandeepsagarks@gmail.com, ashahveda@gmail.com, gnarenk@yahoo.com

Abstract—Air pollution is a serious health concern in densely populated metropolitan cities. Vehicle traffic congestion is a major challenge, especially in medical emergency cases, such as transportation of life saving drugs, accident victims, or transplant organs. In this paper, we propose a signal free corridor based air pollution management for ambulances by maintaining proper speed in which ambulances do not encounter any signals, decreasing the vehicle density and air pollution at junctions. Performance evaluation of Destination Sequenced Distance Vector Protocol (DSDV) and Ad-Hoc On Demand Distance Vector Protocol (AODV) for the vehicular environment in multiple scenarios has been simulated and the results are encouraging.

Keywords- AODV; DSDV; Signal free route; NS-2.

I. INTRODUCTION

Air pollution is mainly caused by fuel-wood and biomass burning, fuel adulteration, vehicle emission and traffic congestion. It is concentrated in densely populated metropolitan cities. Some of the reasons why pollution is concentrated in metropolitan areas include (1) High density of vehicles, (2) Increased number of traffic signals, (3) Increase in the number of diesel vehicles, (4) Increased number of two wheelers, (5) Decreased carpooling, (6) Reduced public transportation usage.

Bangaluru is a metropolitan city having a population of 150,000,000 [3]. As per recent statistics, the vehicle population in Bengaluru is 5,800,000, and 90% of registered vehicles are motor bikes and cars, with 1300 new vehicles getting added every day. Only 40,000,000 people use public transport services like BMTC (Bangalore Metropolitan Transport Corporation). The rest of the population uses their own vehicles. Hence, travel speed has dropped to 15 Km/h.

A. Vehicle Statistics in Bengaluru [3]

We include below some vehicle statistics in Bengaluru.

- 1) Two wheelers – 3,841,140
- 2) Light motor vehicles-1,141,460
- 3) Heavy truck vehicles-108,850
- 4) Auto rickshaws-149,950
- 5) Heavy goods vehicles-73,470
- 6) Floating vehicles-200,000
- 7) other vehicles-244,890

Total number of vehicles is 5,759,760.

The average speed of vehicles on many metropolitan city roads is less than 15 kilometers per hour during peak hours [3]. At such speeds, vehicles in India emit air pollutants 4 to 8 times higher than pollutants emitted when there is less

traffic congestion. Indian vehicles also consume increased carbon footprint fuel per trip than in the case where traffic congestion is less. The more severe the traffic congestion, the longer the vehicles stay at junctions, causing greater air pollution [1].

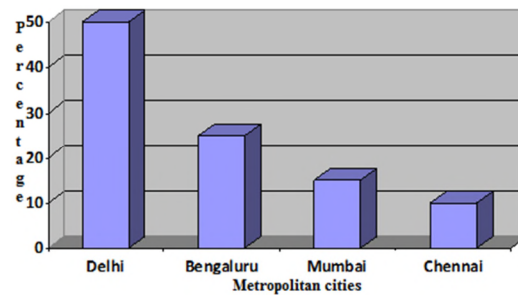


Figure 1. Air pollution levels in different metropolitan cities of India [3].

Bangalore has reported that, one in every two children of the city are suffering from bronchial related diseases, ranking second in the list of cities with highest air pollution levels in India (see Fig. 1). The annual average values of air pollutants in Bengaluru show a linear increase year after year (see Fig. 2). Ideas of road widening, one way streets, strict traffic rules, efficient navigation system can reduce congestion. Vehicles must not wait for a long time at traffic signals and, hence, signal free routes are recommended.

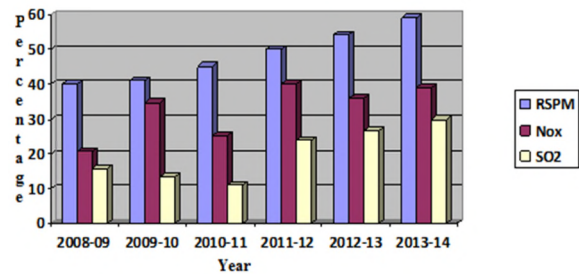


Figure 2. Annual average values of air pollutants in Bengaluru [3].

II. SIGNAL FREE ROUTE

Signal free route is a congestion free route between the source and the destination. In vehicles transport, this means no encountering of any signals of traffic by maintaining optimum speed and coordinating between vehicles on the road. There is an exponential increase in the number of vehicles on the road every day. Traffic congestion reduces average traffic speed during peak hours. At low speeds,

scientific studies reveal, vehicles burn fuel inefficiently and pollute more per trip [2]. Similar to the signal free route for ambulances while transporting transplant organs, signal free routes can be designed for the Bangalore long distance routes using the DSDV and AODV protocols. These protocols show the shortest possible path between the source and the destination. To overcome the hurdles of traffic signals, proper speed has to be maintained by the vehicles. Signal free route for ambulances at the time of emergencies can also be designed using Vehicle Ad-Hoc Networks (VANET). Sensing traffic congestion on the route of an ambulance, the dispatch center communicates to vehicles on the route and thereby vehicles communicate with each other for maintaining proper speed on signal free route and giving way for the ambulances in case of emergencies.

III. VANETS

VANETs are a form of Mobile Ad-Hoc Networks (MANETs) with the mobile nodes being vehicles and it is mainly aimed at providing safety and hassle-free transportation for passengers. In these networks, vehicles are equipped with communication equipment that allows them to communicate with each other i.e., Vehicle to Vehicle (V2V) communication and also to exchange messages with a roadside network infrastructure i.e., Vehicle to Infrastructure (V2I) communication [2]. It provides an efficient way to use vehicular networking. VANETs are applied to control the high density of vehicles during congestion in metropolitan cities. They find their applications widely in electronic toll collection, monitoring, collision warning, road signal alarms, parking lot payment, and so on. Specific areas of research in VANETs include routing, broadcasting, quality of service and security [1].

VANETs are important components of Intelligent Transportation Systems (ITS) to set up communication systems in the form of location tracking using Global Positioning System (GPS), wireless networks. The need for ITS arises from the real time problems of traffic congestion, to provide a comfortable, hassle-free transportation system to the vehicles. Congestion is a major challenge for the efficiency of the transportation system, travel time, speed and fuel consumption. It has its application to retrieve information about instant accidents and traffic messages and helps in designing and developing a better traffic signaling system. It also offers multimedia applications such as Internet connectivity and multimedia access [2]. The better the management of traffic congestion on roads, the better the air pollution management.

Some of the characteristics of VANETs include [10]:

- 1) Highly dynamic topology,
- 2) Frequently disconnected network,
- 3) Mobility modeling and prediction,
- 4) Communication environment,
- 5) Hard delay constraints.

A. IEEE 802.11p and IEEE 1609 Standards

IEEE 802.11p standard [4] describes the extensions added to 802.11 that had significant overheads used in

different multiple vehicular scenarios for Media Access Control (MAC) operations. 802.11p protocol operates in 5.8/5.9 GHz with a guard band from 5.850-5.855 GHz, to support ITS applications of On Board Unit (OBU) to Road Side Unit (RSU) communication. RSU to RSU communications are developed that describe the security, management and physical access in Wireless Access in Vehicular Environments (WAVE) communication. At the Transmission Control Protocol-Wi-Fi Wireless Short Message Protocol (TCP-WI WSMP) layer, Unicasting (for non-safety applications) and Broadcasting (for safety applications) are specified providing limitations in the case of low bandwidth occupancy and low power usage [1].

The IEEE 1609 family of standards for WAVE [4] describes the architecture, communication model, management structure, security mechanisms and physical access for high speed (about 27 Mb/s) and short range (about 1000m) low latency wireless communications in the vehicular environment. The primary architectural components defined by the 1609 standards are the OBU, RSU and WAVE interface. IEEE 1609.0 defines the usage of WAVE environment by applications, the respective management activities are defined by IEEE p1609.1, the security protocols are defined by IEEE p 1609.2, and the network-layer protocols are defined by IEEE p 1609.3. Additionally, the standard also provides extensions to the physical channel access defined in IEEE 802.11 to support the WAVE standards in IEEE p1609.4.

The IEEE 1609 family standards find various applications in design, specification, implementation and testing of WAVE devices used in transportation, automotive and traffic management. Network, hardware and application designers of ITS use these standards as they define the communications architecture for Dedicated Short Range Communication (DSRC) based V2V and V2I interactions, and they pose as the basis for the low-latency interface design of on-board and roadside devices. ITS application designers use the standards to deploy the basis for interface definitions between system components and as a framework for application architecture. The architecture, interfaces and messages defined in the IEEE 1609 family of standards for WAVE support the operation of secure wireless communications between vehicles and infrastructure, as well as between vehicles [9].

B. Dijkstra's Algorithm

The Dijkstra's Algorithm is a graph based search algorithm that gives the shortest path solution for the single source graph with positive edge costs by producing the shortest path tree. This routing algorithm finds the path of lowest cost between the source node and every other node, thereby resulting in the shortest path between source and destination. It has wide applications in network routing protocols such as Open Shortest Path First (OSPF) and is also helpful in finding the shortest distance to the destined city from a source city [2].

C. Ad-Hoc On Demand Distance Vector Protocol

The AODV algorithm is a reactive protocol whose route finding is carried out by Route Discovery Cycle (RDC) maintaining active routing. It is a descendant of DSDV protocol providing unicast and multi-cast communication and adding an extra feature of sequence number to prevent loop formation. AODV is noted for its quick adaptation under dynamic link conditions, consuming less network bandwidth, which can be scalable for large network [5].

D. Destination Sequenced Distance Vector Protocol

DSDV is a Bellman-Ford algorithm based protocol that provides solutions to routing loop problems by adding sequence numbers. DSDV provides the best performance in networks with moderate mobility and few nodes. Requiring regular update of its routing tables becomes the major setback for its performance in dynamic environments and large network [6].

E. NS-2

Network Simulator-2 (NS-2) is a well-accepted tool for network simulation, as its architecture is suitable for extensions and interfacing with other simulation modules. It can be implemented using IEEE 802.11 protocol. It is used in the simulation of routing protocols and it is highly used in ad-hoc networking research. NS-2 supports popular network protocols, offering simulation results for wired and wireless networks alike. AODV and DSDV routing protocols are used for simulation [1].

IV. PERFORMANCE EVALUATION OF GREEN CORRIDOR MANAGEMENT

Vehicles equipped with communication devices can interact effectively by exchanging information on congestion with the dispatch center and other surrounding vehicles, making it feasible for the vehicles to travel in coordination and giving way to ambulances in a signal free route. Congestion Detection Algorithms (CDA) are used to find the high traffic density areas with low vehicular speeds allowing the drivers of ambulances to reroute to any other shortest path available. RSU are deployed to transmit data during obstacles in order to restrict the exchange of data during its need from the centralized location. The information updates on congestion can be given to drivers in the form of text messages to the OBU.

Ever since the Bangalore International Airport came into existence, a number of road development projects in all parts of the city have been deployed to make the travel to the airport convenient and faster. In parallel, there is an exponential increase in the number of vehicles every day, which in turn increases congestion and air-pollution, reducing average traffic speed and causing inefficient usage of fuel by vehicles.

Bangalore has evidenced the instances of green corridor with the help of Bangalore traffic police during major life saving emergency cases. Green corridor for live heart transport from Bangalore to Chennai reduced the travel time

to 2 hours from 7 hours in September 2014 [5]. Deriving the idea from this, signal free routes can be designed for the Bangalore long distance routes using the DSDV and AODV protocols depicting the less possible congestion and less air polluted path between the source and destination. Vehicles that maintain proper speed stay in the signal free route, experiencing less traffic congestion by increasing the average travel speed by at least 2 times. The air pollutants density in the vicinity of traffic signals and travel time are decreased considerably.

A. Implementation

The required aspects for the performance evaluation are:

- I. Availability of sufficient VANET modules on the routes to detect congestion.
- II. Existence of dispatch centre that holds the updated data, based on traffic congestion.
- III. Presence of RSU helps in signal propagation during obstacles restricting propagation.
- IV. Vehicles in signal free route maintain speed in the given range, as shown in Table II.

B. Case Study

1) For Public Transport:

The vehicle of interest starts from the source at Dr. Ambedkar Institute of Technology College (Dr. AIT) to reach the destination University Visvesvaraya College of Engineering (UVCE) and can go by 4 routes and achieve the shortest path with fewer signals in route. For the least air pollution, the vehicle has to travel through the route which has the least number of traffic signals that contribute to the shortest path.

1. Route 1 has 10 traffic signals.
2. Route 2 has 11 traffic signals.
3. Route 3 has 11 traffic signals.
4. Route 4 has 9 traffic signals.

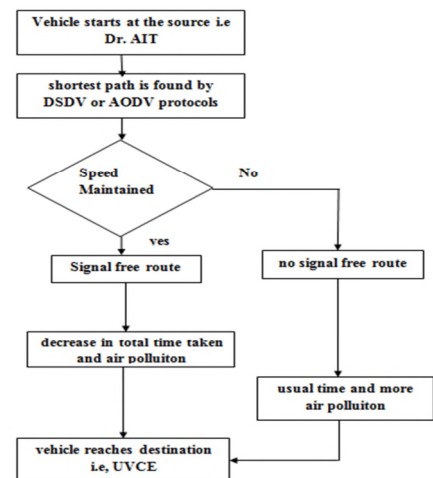


Figure 3. Flowchart for Case 1.

Since route 4 has the least number of traffic signals, the vehicle of interest travels along this route. The idea of signal free route can be implemented by taking speed, congestion, direction of traffic, duration, number of traffic signals on the route and emergency into considerations. The shortest path is found by DSDV or AODV protocol considering traffic signals as nodes traversing through the minimum number of traffic signals, and, consequently, minimal air pollution. In turn, it also increases the average speed and decreasing the time taken to reach the destination by 50% compared to the existing traffic system. The air pollutants spread over the route instead of getting concentrated in traffic junctions, as depicted in Fig. 3.

TABLE I: TABLE TO SHOW THE DISTANCE BETWEEN EACH SIGNALS, SPEED AND TIME IT TAKES TO REACH WITH THE USUAL TRAFFIC SYSTEM

From	To	Distance	Speed	Time
Dr.AIT	Nagarabhavi Circle	2	20	6
Nagarabhavi Circle	Chandralayout	1.2	14	5
Chandralayout	Attiguppe	0.9	18	3
Attiguppe	Underpass	3	20	9
Underpass	Rajajinagar 5th Block	1.6	16	6
Rajajinagar 5th Block	More	1.4	14	6
More	Okalipuram	0.75	15	3
Okalipuram	Majestic	1.4	8.4	10
Majestic	UVCE	3.2	17	11

TABLE II: TABLE TO SHOW THE DISTANCE BETWEEN EACH SIGNALS, SPEED AND TIME IT TAKES TO REACH WITH SIGNAL FREE ROUTE

From	To	Distance	Min Speed	Max Speed	Time
Dr.AIT	Nagarabhavi Circle	2	60	80	2
Nagarabhavi Circle	Chandralayout	1.2	36	72	2
Chandralayout	Attiguppe	0.9	27	36	2
Attiguppe	Underpass	3	45	72	4
Underpass	Rajajinagar 5th Block	1.6	48	64	2
Rajajinagar 5th Block	More	1.4	42	56	2
More	Okalipuram	0.75	22.5	30	2
Okalipuram	Majestic	1.4	28	33.6	3
Majestic	UVCE	3.2	32	38.4	6

2) For Ambulance:

The signal free route for ambulances at the time of emergency and live organ transportation is designed using VANETs, control unit and routing protocols. Considering speed, congestion, direction of traffic, duration, number of traffic signals in the route, number of ambulances in the same traffic junction, as major aspects, a green corridor is designed for ambulances. The VANETs in the path ahead of the ambulance periodically predict the possible congestion by considering the speed of surrounding vehicles and

transmits the information to the surrounding nodes. In the case of vehicles traveling at a greater speed, no congestion message is transmitted but, real time congestion message is transmitted to the surrounding nodes in the case of vehicles traveling at a lower speed than that of other vehicles (nodes), enabling the ambulance to find alternate shorter paths to the nearest hospital. The unavailability of alternate route justifies the need for a signal free corridor, during which all traffic signals in the route are controlled by the ambulance, unblocking the way by signaling green. As the ambulance approaches the traffic signal within a distance of 500 m, the control of the nearest traffic signal in the route is taken over by the ambulance automatically until it is given safe passage; the ambulance is equipped with GPS and Radio Frequency (RF) module. The longitude and latitude of the traffic congested location is tracked using the GPS and the same information is updated in the database of the dispatch center. The technique of controlling traffic signal lights in signal free route consists of two units which synchronize with each other and help achieve a hassle-free travel for ambulances, Fig. 6.

The GPS installed in the ambulance unit senses the positional coordinates of an ambulance and the controller is equipped with embedded system and encoder. It predicts the direction of the ambulance using the information from the GPS receiver and transmits the same information to the decoded signal which causes the embedded system to work in emergency mode and signal green in the lane of the ambulance and red in other lanes until the ambulance crosses the junction. As soon as the ambulance crosses the junction, the embedded system starts to work in normal mode, resuming the previous conditions.

Priority will be assigned in the case of two or more ambulances in the same traffic junction, based on the distance between each ambulance and the traffic junction and the received signal strength from other ambulances. The closer the ambulance, the better the received signal strength and the higher its priority. Thereby, an efficient signal free corridor for life saving emergency cases is created.

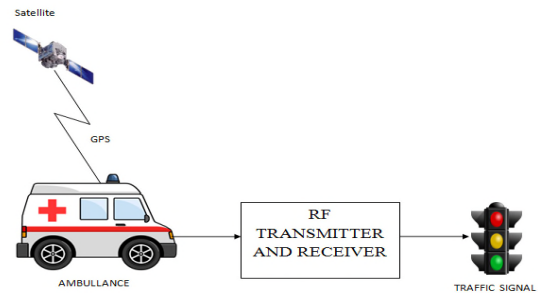


Figure 4. Proposed traffic signal monitoring system.

V. SIMULATION AND RESULTS

Experimental Analysis: To simulate the multiple scenarios of the wireless network, a TCL program is written

in NS-2. The simulation related parameters are given in Table III.

TABLE III: SIMULATION RELATED PARAMETERS

Channel Type	Wireless
Radio-Propagation Model	Two-Ray Ground
Network Interface Type	Wireless Phy
MAC Type	Mac/802.11
Interface Queue Type	Queue/Drop Tail/Pri Queue
Antenna Model	Antenna/Omni
Link Layer Type	LL
Max packet in ifq	50

Case 1) The nodes serve as traffic signals at respective places alternating red, yellow and green color. Mobile nodes with black, brown, cyan, serve as vehicles. The node numbered 41 is the vehicle of interest and moves from source to destination, traversing a minimum number of traffic signals in the signal free route and maintaining the speed in proper range. The routing protocol used is DSDV resulting in a minimum number of packet loss, Fig. 5.

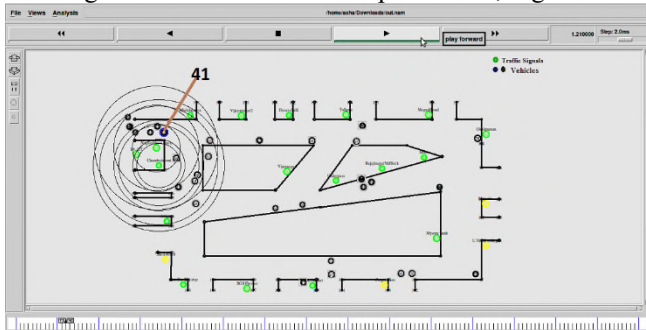


Figure 5. Signal free route using DSDV protocol.

The nodes representing traffic signals at respective places alternate red, yellow and green color. Mobile nodes with black, brown, cyan, act as vehicles. The node numbered 41 is the vehicle of interest, moves from source to destination traversing a minimum number of traffic signals in signal free route maintaining the speed in proper range. The routing protocol used is AODV resulting in an increased number of packet loss compared to DSDV, Fig. 6.

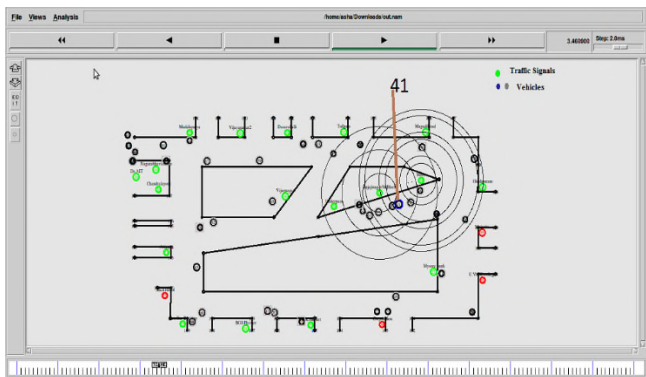


Figure 6. Signal free route using AODV protocol

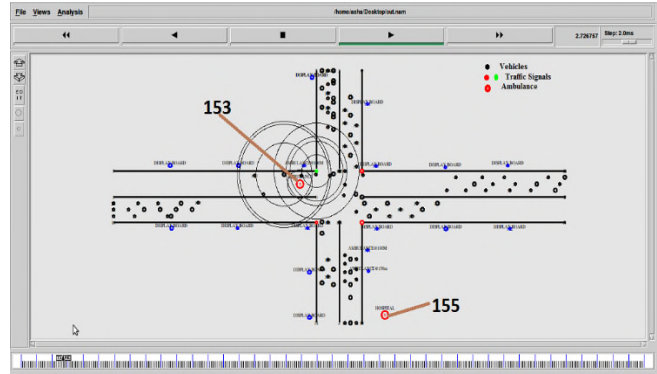


Figure 7. Signal free route in case of single ambulance.

Case 2) In case of an ambulance, considering a traffic junction with 4 lanes, blue nodes represent display board, 4 traffic signals, each for respective lanes alternating red, yellow, green color, red color. The node numbered 153 is an ambulance and the node numbered 55 represents a hospital. Knowing the positional co-ordinates of the ambulance, the display board displays the distance at road sides, vehicles ahead of the ambulance give way for it, and the RF transmitter in ambulance controls the traffic signal, Fig. 7.

In case of two or more ambulances, the nearest ambulance to the traffic junction and the highest signal strength received by the RF receiver is given highest priority. The nodes numbered 111 and 153 are ambulances destined to reach node 155, the hospital. Ambulance 111, being closer to the traffic junction than 153, impacts to signal green for 111 lane first and finally to 153, Fig. 8.

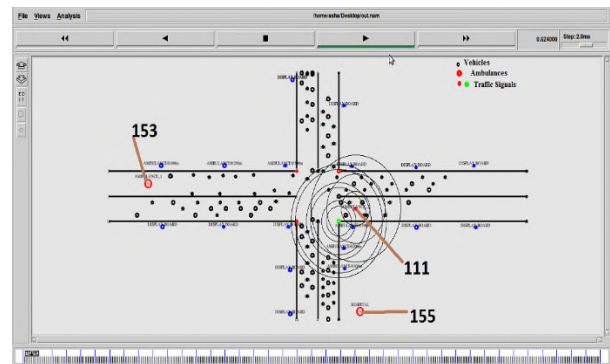


Figure 8. Signal free route in case of two or more ambulances.

The idea of signal free route for air pollution management has been simulated using both DSDV and AODV protocols for various scenarios. The average speed of vehicles on the road has doubled compared to the existing traffic system (Fig. 9) and the average travel time has decreased to half the usual travel time; the usual travel time and speed is as mentioned in Table I. Fig. 10 depicts decreasing the air pollution at traffic junction areas at least by 2 times.

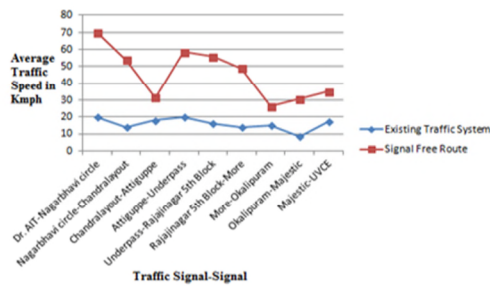


Figure 9. Comparison of existing traffic system and signal free route with respect to average traffic speed.

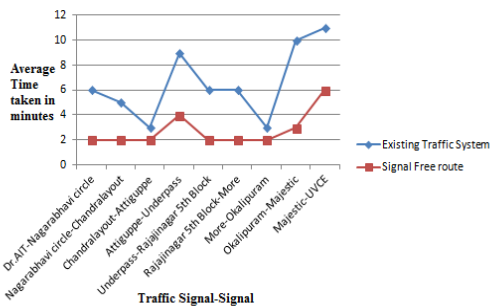


Figure 10. Comparison of existing traffic system and signal free route with respect to average time taken to reach next signal.

VI. ADVANTAGES AND DISADVANTAGES

Some of the advantages of the system include: a decrease in air pollution and a decrease in the average time taken to reach the destination by increasing the average traffic speed. It is an efficient method for ambulances and other emergency vehicles.

Among the shortcomings of the system, we can list: the probability of increase in road accidents due to increase in average traffic speed and uncoordinated vehicles on road.

VII. CONCLUSION

This paper aims at managing air pollution in metropolitan cities by creating signal free route reducing congestion at traffic junctions. Vehicles move in coordination with each other on a signal free route and maintain a proper speed, increasing the average speed of vehicles on the road. The vehicles with average speed more than 30kmph burn fuel efficiently which results in reduction of air pollution by at least 2 times. The DSDV and AODV

protocols predict the least congested traffic route. The signal free route for ambulances at the time of emergencies is designed using VANETs and automatic control of traffic signal lights, resulting in a hassle-free transportation for ambulances. The idea of signal free route can be extended to other emergency vehicles like Vehicles with Z cross security, cash van, fire engines, etc., decreasing the average travel time. In the future, the idea of signal free route can be extended to other emergency vehicles like VIP, cash van, fire engines, etc., decreasing the average travel time. In order to decrease injuries from road accidents, a rule of compulsory wearing of seat belts for all passengers in the vehicles becomes necessary.

REFERENCES

- [1] Aparajitha Murali, K. Bhanupriya, B. Smitha Shekar, and G. Narendra Kumar, "Performance Evaluation of IEEE 802.11p for Vehicular Traffic Congestion Control", 11th International Conference on ITS Telecommunications, pp. 732-737, 2011.
- [2] B. Smitha Shekar, G. Narendra Kumar, H. V. Usha Rani, C. K. Divyashree, Gayatri George and Aparajitha Murali, "GPS Based Shortest Path for Ambulances using VANETs", International Conference on Wireless Networks (ICWN 2012) IPCSIT volume 49, pp. 190-196, 2012.
- [3] Website: Bangalore Traffic Police, www.bangalorettrafficpolice.gov.in.
- [4] Christoph Schroth, Florian Dtzer, Timo Kosch, Benedikt Ostermaier, and Markus Strassberger, "Simulating the Traffic Effects of vehicle-to-vehicle messaging Systems," 5th International Conference on ITS Telecommunications, 2005.
- [5] Ian D. Chakeres, and Elizabeth M. Belding-Royer, "AODV Routing Protocol Implementation Design", 24th International Conference on Distributed Computing Systems Workshops (ICDCSW'04), March 2004.
- [6] Prof B. N. Jagdale, Prof. Pragati Patil, Prof P. Lahane, and Prof D. Javale, "Analysis and Comparison of Distance Vector, DSDV and AODV Protocol of MANET," International Journal and Distributed and Parallel Systems (IJDPS) volume 3, No. 2, pp. 121-131, March 2012.
- [7] K. Sangeetha, P. Archana, M. Ramya, and P. Ramya, "Automatic Ambulance Rescue with Intelligent Traffic Light System," Coimbatore, India. IOSR Journal of Engineering, volume 04, Issue 02, pp. 53-57, February 2014.
- [8] E. Geetha, V. Viswanadha, and G. Kavitha, "Design of an Intelligent Auto Traffic Signal Controller with Emergency Override," International Journal of Engineering Science and Innovative Technology (IJESIT), volume 3, Issue 4, pp. 670-675, July 2014.
- [9] Sebastian Griang, Petri Mahonen, and Janne Riihijarvi, "Design of Performance Evaluation of IEEE 1609 WAVE and IEEE 802.11p for Vehicular Communications," Second International Conference on Ubiquitous and Future Networks (ICUFN), pp. 344-348, 2010.
- [10] Aamir Hassan, "VANET Simulation," <http://www.diva-portal.org/smash/get/diva2:239905/FULLTEXT01.pdf>, Master's Thesis in Electrical Engineering, School of Information Science, Computer and Electrical Engineering, HalmstadUniversity,2009.

Neural Signal Processing and Motion Capture as a Feedback Mechanism to Improve Interceptive Human Movement

Devanka Pathak

Centre for Creative Computing
Bath Spa University, Corsham Court
Corsham, UK, SN13 0BZ
Email: devanka.pathak14@bathspa.ac.uk

Meiyu Shi

Tourism Institute
Beijing Union University
Beijing, 100101, China
Email: shimeiyu72@163.com

Tin-Kai Chen

Department of Animation and Game Design
Shu-Te University
Kaohsiung City, 82445, Taiwan
Email: tchen1@stu.edu.tw

Hongji Yang

Centre for Creative Computing
Bath Spa University, Corsham Court
Corsham, UK, SN13 0BZ
Email: h.yang@bathspa.ac.uk

Abstract—In this paper, we present a framework to explore the role of motion capture and neural information processing in a coordinated execution of movements in the sporting context. We discuss the perception-cognition-action coupling from a motor function consideration. For this, we present a generic experimental design for brain source connectivity estimation. We show the visualisation of the brain connectivity using a sample Electroencephalography (EEG) data-set. We propose to extrapolate the application of similar design to study sporting movements such as cricket batting. We present the case for the use of portable and mobile EEG sensors to study such a low latency decision-making task. Finally, we describe a preliminary framework on how to use and validate the efficacy of neurofeedback in coaching skilled human movement. Taking a multi-modal approach, we included motion capture data to study the skilled movement. From this, we present the wrist movement variation in a shadow batting task by a novice batsman.

Keywords—Interceptive; Movement; Neural; Neurofeedback; Motion-capture.

I. INTRODUCTION

Humans could perform complex movements, e.g., in sports, dance, and other skilled activities. Although the actions manifest in physical dynamics, specific internal mental models precede most of these movements. The human brain together with the peripheral Central Nervous Systems (CNS) dictates the quality of motion. The quality, in turn, depends on training and feedback, especially in the case of skilful execution of movement patterns, e.g., in a sporting context.

Brain-Computer Interfaces (BCI) are systems where the signals from the brain are used to control a computation platform directly. This paper presents the theoretical background and validation of the computational model that explores the role of neural circuitry in interceptive action execution as a BCI feedback system. We also present a video analytics method

for studying the pattern of movement involved in a defensive cricket shot - the so-called forward defence.

Fundamental models of skilled actions in the human movements are well known [1]. To get a full picture of skilled movement execution and the sensory dynamics of the human agent, both internal and external influences should be considered. The internal models refer to the neural processes that govern the CNS in preparation and execution of the concerned movement patterns, while the external processes are the physical manifestation by the subject performing the same movement tasks.

In cricket, a batsman has to move the bat to the right place at the right time to intercept a fast moving ball; the mental models influence the outcome. Indeed, motor control follows an internal forward model [2]. In the case of such an interceptive movement, the batsman, the bat, and a travelling ball form a closed loop feedback-feedforward system. Feedback from internal and external agents helps the subject to evaluate the past performance, while the internal feedforward models help anticipate the unknown variables before task execution.

Feedback plays a critical role in human motor activities [3]. For example, an improper feedback would induce inefficiencies in the movement mechanisms, and that would cause the motor activity to suffer. Self-adjusting instructions in an automatic system are equivalent to influencing the part of the brain that generates a particular motor behaviour. Hence, finding the source localisation as described in existing literature, could help the training process to achieve the desired mental state [4] [5]. This research takes a multimodal approach. The sub-sections below provide a brief introduction on various modalities.

A. The Neurofeedback Approach

Neurofeedback refers to the method of identifying the brain regions that get triggered during a functional task execution and then to use the information to provide feedback to the subject via visual or auditory cues. For this, different sensor data are collected, e.g., Electroencephalography (EEG), Magnetoencephalography (MEG), and functional Magnetic Resonance Imaging (fMRI), etc. In the real world sporting context, a portable EEG device is ideal. We present connectivity analysis on a sample dataset. EEG allows to carry out high temporal resolution studies, which implies that we can see what happens in the brain when the subject performs a task in near real-time.

B. The Motion Capture Approach

The spatial and temporal components of the movement (position, velocity, and acceleration) carry biomechanics signature of action and can be used to compare the quality of movement variation in a single subject or across subjects [6]. Active marker systems, such as the Optotrak [7], allow to place markers on the subject and to observe individual parts of the movement. We present a preliminary movement analysis from Optotrak Motion Capture System (MOCAP).

C. Background on Motion Capture and Brainwave Data Analysis

The motion capturing system, as used in the experiment presented in this paper, uses markers on different body parts and allows to find granular variations in different body regions undergoing movement. Similarly, brainwave sensors allow capturing functional correlates of different wave-bands generated during a task execution. The firing of neural circuits gives rise to electrical activities in the brain. The sensors (EEG) placed on the scalp can detect and measure the electric component of the electromagnetic waves from the electrical dipoles in these circuits. Similarly, MEG measures the magnetic component in the signals. To use a neurofeedback paradigm in training movement, it is necessary to find the location of the sources related to a particular functional activity. Source localisation from the detected signals forms the Inverse problem. These brainwave signals fall into different groups based on their frequency ranges that correspond to different functional mental states. Typical frequency ranges dominant in EEG are **alpha**, **beta**, **delta**, and **theta**. Fig. 1 shows different groups and their associated functional correlates.

The organisation of the rest of the paper is as follows. In Section II we present the methodologies followed in collecting the sample data [11]. In Section III we describe the connectivity metrics followed by the experimental design of the Motion Capture system in Section IV. Section V concludes with a summary of the benefits, advantages and limitations of our approach and describes the future direction for this work.

II. EXPERIMENTAL DESIGN METHOD: NEUROIMAGING

We present the experimental design for neuroimaging example dataset below.

A. Connectivity Analysis Protocol in EEG and MEG

The brain source localisation could be used to visually display the neural network connectivity for functional task performance [8]. The brain network connectivity analysis using EEG and MEG for a high temporal resolution extend the neurofeedback modality. Combining EEG and MEG will make it possible to distinguish the mechanisms that are the event-related from that evoked potential. Thus allowing precise identification of the brain areas during a successful and failed execution of the batting task described in Section IV.. Hence, the connectivity patterns during the successful performance of a cricket shot could be used to provide feedback in future performance. It is possible to gamify the feedback-feedforward loop by designing a rewarding and penalising the subject in a scoring scale. The gamification part will be explored further in future work. The neuroimaging is then to be combined with MOCAP data analysis to provide feedback on, e.g., ideal hand and wrist movement in a defensive cricket batting stroke.

B. The Sample Dataset

The MGH/HMS/MIT Athinoula A. Martinos Center for Biomedical Imaging at Massachusetts General Hospital (MGH), Harvard Medical School (HST), and Massachusetts Institute of Technology(MIT) acquired and made available the example dataset captured with the Neuromag Vectorview system. EEG data from a 60-channel electrode cap was obtained simultaneously with the MEG. The raw data refers to the continuous time series, the Epochs imply the collection of time-locked trials and averaged data over trials, e.g., storing Left Auditory, and Right Visual in a single file is the averaged data known as Evoked. Details of the data collection protocol are available in the literature. The methodology follows one subject's brainwave recording associated with triggered finger movement [4]. An occasional appearance of a smiley face was the stimulus at the centre of the subject's visual field. The instruction to the subject was to press a key with the right index finger as soon as possible after the face appeared. Tab-I lists the trigger codes [9].

TABLE I
TRIGGER CODES FOR THE SAMPLE DATA SET.

Name	Code	#Contents
LA	1	Response to left-ear auditory stimulus
RA	2	Response to right-ear auditory stimulus
LV	3	Response to left visual field stimulus
RV	4	Response to right visual field stimulus
Smiley	5	Response to the smiley face
Button	32	Response triggered by the button press

C. Brainwave Data Analysis Protocol

Cortical surface-based functional brain imaging involves segmentation and surface reconstruction [10]. To cortical constraint, the EEG/MEG source, the data analysis protocol uses the MRI of the subject. The computational algorithm covers the following stages and analysis [9] [11]:

- Preprocessing and denoising



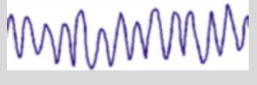


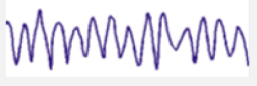
Power Spectral Density	Band	Frequency (Hz)	Correlates
	Delta	< 3	sleep
	Theta	3 – 7	memory creation
	Alpha	8 – 13	relaxation, closed eyes, intrinsic focus, reflection
	Beta	13 – 30	intense concentration, cognition
	Gamma	30 +	multi-sensory, euphoria, high focus
	Mu	8 – 12 (over sensorimotor)	suppression has been linked with empathy

Fig. 1. Comparison of EEG frequency bands and corresponding mental state activation. Please see Section C. for more details

- Source estimation
- Visualisation of sensor- and source-space data
- Time-frequency analysis
- Statistical testing
- Estimation of functional connectivity
- Applying machine learning algorithms

We validate the example relevant to the functional connectivity analysis after performing the preprocessing steps.

D. Preprocessing

Preprocessing eliminates the defective EEG channels to make sure that errors due to incorrect data do not propagate further in the pipeline. Signal Space Projection (SSP) and Independent Component Analysis (ICA) routines suppress the artefacts [12]. Fig. 2 shows the result of the covariance matrix estimates.

III. CONNECTIVITY METRICS

Dynamic statistical parametric mapping (dSPM) [13] and MNE [14], sLORETA estimates source activation from MEG and EEG data. To study the brain region connectivity, both model-based and data-driven approaches are applicable, respectively in the time and frequency domains. Connectivity

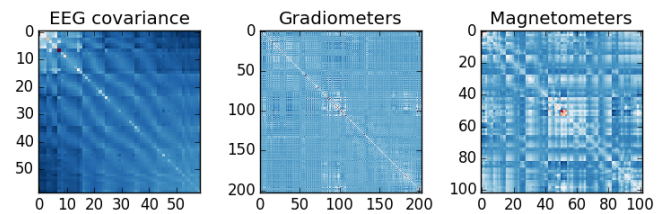


Fig. 2. Covariance matrix estimation on the raw data with Signal Space Projection

analysis provides a way to perform multivariate analysis of brain region in response to different stimuli such as auditory and visual. Another area to study is the connectivity within the brain and brain-CNS regions. The example of connectivity between a seed-gradiometer close to the visual cortex and all other gradiometers as shown in Fig. 3 uses the metric Squared Weighted Phase Lag Index [15].

Fig. 4 shows the connectivity computed between 4 labels across the spectrum between 5 and 40 Hz.

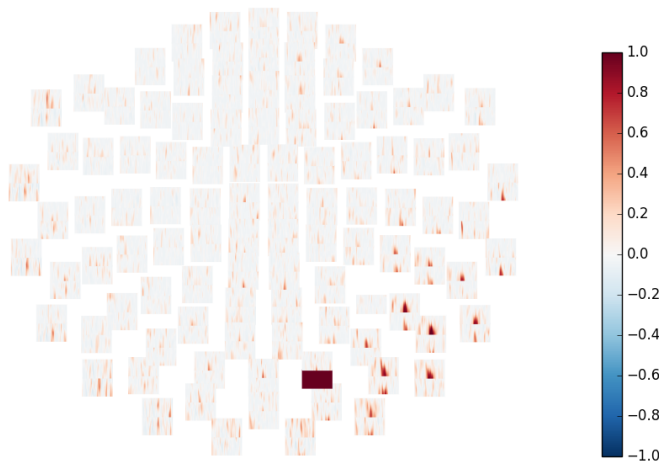


Fig. 3. Connectivity map of a seed gradiometer using Squared Weighted Phase Lag Index

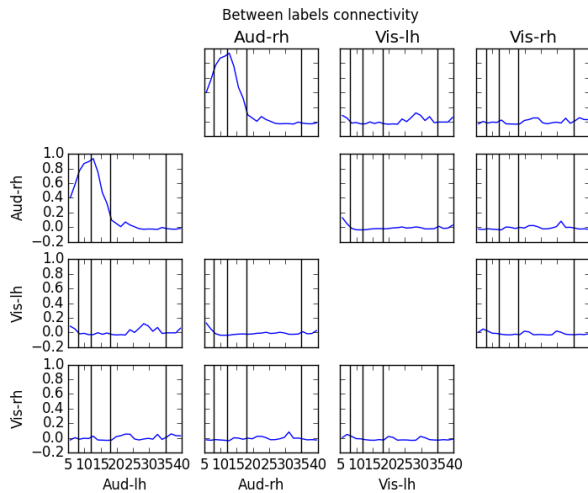


Fig. 4. Inverse Connectivity Spectrum

A. EEG BCI Motor Imagery And Real Time Feedback

In this section, we discuss some preliminary results from a motor imagery data available at PhysioNet [16]. The data collection is as per the experimental protocol described in [17].

The motor imagery could be decoded from this type of dataset by separating the signal into additive components, which have maximum differences in variance between the windows of the multivariate signal. This method is known as the Common Spatial Pattern (CSP). Work is underway on this dataset to improve the classification accuracy and to use the method in real-time data analysis similar to the imagery protocol described above.

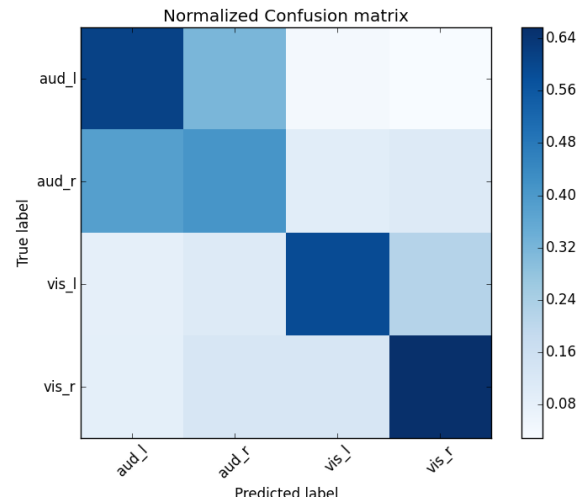


Fig. 5. Labels for rescaled channels for left and right motor imagery stimulus

Event Related Potential (ERP) is decoded with xDAWN as shown in Fig. 5 [18], [19]. For each event type, a set of spatial xDAWN filters is trained and applied on the signal. Channels are concatenated and rescaled to create feature vectors. They are fed into a Logistic Regression. The real-time feedback mechanism with a client-server setup could be used for feedback. The server is started so that future stimuli for the classification task are presented via the client. This is predicted less accurately, and an on-demand adaptation of the stimuli is issued to improve the performance of the classifier to compute various statistics in real-time. Currently, we are exploring the simulated data with a plan to extend the pipeline to include real experiments in future.

IV. EXPERIMENTAL DESIGN: MOCAP

We present the experimental design for MOCAP dataset below.

A. Method

The aim of the study is to perform an initial test for verification of the experiment design using the passive motion capture system with an inter-reliability test.

B. Participant

A novice (with no experience in cricket), a right-handed male student in the Department of Animation and Game Design in Shu-Te University volunteered. We followed the health and safety briefing, risk assessment, and obtained informed consent as per Bath Spa University and Shu-Te University standard protocols. The participant’s age is 24 years old, using his preferred right hand to perform the task.

C. Apparatus

We conducted the experiment in the motion capture laboratory in Hengshan Innovation Base in Shu-Te University, Taiwan. We initially assessed four participants in a single batch, thanks to the maximum capability of the laboratory. But for the result presented here, we included data from only one

participant. The rest of the data will be used for a comparative analysis of the movement patterns among participants in future work. OptoTrak Inc. provided the passive motion capture system shown in Fig. 6 and Fig. 7.



Fig. 6. On body marker on subjects. Only one subject’s data is presented here.

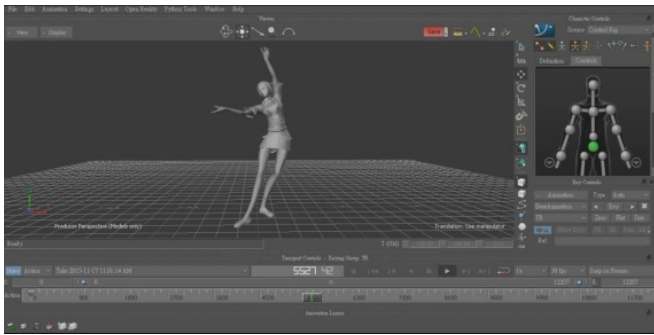


Fig. 7. Screenshot of the MOCAP software renderer.

The system has 20 pieces of high-speed IR camera (Product ID: Prime 17W), each of the cameras has 70 degree of Field of view (FOV) wide angle, offering true, edge to edge coverage across the camera’s image sensor. It has advantages of perfectly matching the imager’s resolution, 20 pcs of motion capture IR camera, with an amazing 70° FOV, 1.7 MP of resolution and a 360 FPS capture rate. Furthermore, the following equipment and tools were employed:

- A 60" LCD monitor was provided to play the video showing the movement of the swing movement.
- The data analysis was performed using MATLAB and SPSS version 13.

D. Procedure

The objective measurement involves a laboratory-based movement experiment with three repetitions. We captured the movements on 120 frames-per-second (FPS). During the trial, the participant was instructed to perform a cricket bat

swing motion based on the video playing in front of him. We performed a five-minute warm-up test before the experiment. It ensures the inherent reliability of the study.

E. Limitations

One potential limiting assumption was that the bias might be minuscule and could be normalised. Hence, we ignored any biases because the number of samples collected is large. Thus significantly reducing the bias on frames captured from the system.

F. Data Analysis

The motion data collected at the 120 (FPS) resulted in a total of $N = 1,588$ frames, implying a successful recording of approximately 1MB of data. Thus, provided enough data for doing an analysis by quantitative method. The variance of velocity is as shown in Fig. 8. The analysis of variance indicated that the mean movement time was no significant difference among three repetitive movements, $p < 0.05$. Thus, the experiment design was consistent.

G. MOCAP Results

In the case of cricket batting, for right-handed batsmen, the left hand is the leading side and vice versa for the left-handed batsman. The leading hand is the most important in controlling the bat movement. Hence, we focused on the data from the three markers located at the left-hand wrist on top, bottom and side-on positions. We calculated the position vectors from the Cartesian coordinate values at each instant during the movement as provided by the MOCAP and generated Acceleration Profile for the left-hand wrist as shown in Fig. 8.

One subject was undergoing multiple trials to produce data so that random statistical significance could be detected. In the next phase, the subject will be identifying the ball movement direction as a stimulus presented on a screen. A second task will be to predict the spin and swing direction of the ball concurrently capturing the MOCAP and EEG data.

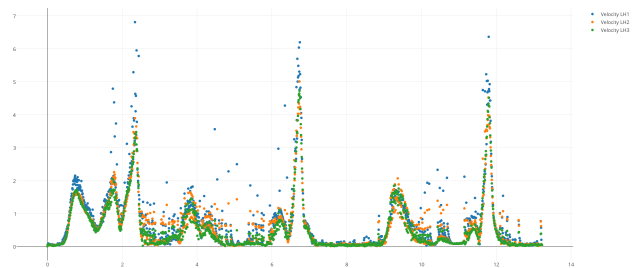


Fig. 8. Variation of velocity with time of the wrist of the leading left hand of a right handed subject performing the shadow forward defence shot

V. CONCLUSIONS, PERSPECTIVES, AND THE FUTURE DIRECTION

In athletic performances, perception and action need to be in a synchronised state. If an athlete is great at perceiving

and processing a mental model on the sport specific task but unable to perform the required action successfully, that will be of no use in successful execution of the task.

The hypothesis that the action influences perception is one of the ideas we aim to establish a multimodal methodology encompassing neurofeedback and motion capture feedback on task execution. In a case of cricket batting, this implies that a successful hitting of the target at practice could lead to the perception of increased size, and slower movement of the ball against failed execution of interceptive action would result in a perception of smaller size and faster motion of the ball. This research extends the signal processing part from the previous work [4], [8], and also looks at the experimental movement data that could be correlated to neuroimaging data in future as a gamified neurofeedback or as a neurogaming application for training athletes. To attain the level of sophistication to be used in real world situations, we need to improve and develop existing data analytics methods and combine with other movement related modalities to build a practical training framework. We establish the feasibility of connectivity measures in MEG-EEG monitoring and suggest ways measure motor performance from a combination of cognitive states and motion capture in an interceptive movement. To combine these modalities for training interceptive action is the goal and future direction of this research.

ACKNOWLEDGEMENT

This work was supported in part by Natural Science Foundation of China (Grant no. 71473018). Special thanks to the following staff members of the motion capture lab: Ji-Fu Wang, Wei-Cheng Wang, and Yen-Nan Tung. This work benefited from the informal discussion and insights on elite performance from ex-cricketer and Delhi Ranji Trophy cricket team strength and conditioning coach N. Bordoloi.

REFERENCES

- [1] E. Morsella, J. Bargh, and P. M. Gollwitzer, *Oxford Handbook of Human Action*, Oxford University Press, Oxford, 2008.
- [2] R. C. Miall and D. M. Wolpert, "Forward Models for Physiological Motor Control," *Neural Networks*, 8, 2008.
- [3] N. Wiener, *Cybernetics or control and communication in the animal and the machine*, MIT Press Paperback Edition, 1965.
- [4] D. Pathak, H. Yang, T. K. Chen, J. Fishenden, and A. Lee, "Measuring Brain Signals to Evaluate the Role of Creativity in Interceptive Human Movement," *IEEE ISCC*, Oxford, 2016.
- [5] T. Mitsuru et al. "Event related desynchronization-modulated functional electrical stimulation system for stroke rehabilitation: A feasibility study," *Journal of NeuroEngineering and Rehabilitation*, vol. 19, 2012.
- [6] J. H. Hamill and K. M. Knutzen, *Biomechanical Basis of Human Movement*, Wolters Kluwer, 2009.
- [7] R. A. States and E. Pappas, "Precision and repeatability of the Optotrak 3020 motion measurement system," *Journal of Medical Engineering & Technology*, vol. 30, no. 1, pp. 11-16, 2006.
- [8] D. Pathak, H. Yang, and T. K. Chen "Neurofeedback and Creativity in Interceptive Human Movement: A Theoretical Model for Neurocybernetics Based Kinaesthetic Multimodal Learning Agent," *IEEE IWSC*, Vienna, 2016.
- [9] A. Gramfort et al. "MEG and EEG data analysis with MNE-Python" *Frontiers in Neuroscience*, vol. 7, 2013.
- [10] A. M. Dale, B. Fischl and M. I. Sereno, "Cortical Surface-Based Analysis I: Segmentation and Surface Reconstruction," *NeuroImage*, vol. 9, 1999.

- [11] A. Gramfort et al. "MNE software for processing MEG and EEG data," *NeuroImage*, vol. 86, 2014.
- [12] M. A. Uusitalo and R. J. Ilmoniemi, "Signal-space projection method for separating MEG or EEG into components," *Med Biol Eng Comput*, 19, vol. 35, pp. 135-140, 1997.
- [13] A. M. Dale et al. "Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity," *Neuron*, vol. 26, no. 1, pp. 55-67, 2000.
- [14] O. Hauk, "Keep it simple: a case for using classical minimum norm estimation in the analysis of EEG and MEG data," *Neuroimage*, vol. 21, no. 4, pp. 1612-21, 2004.
- [15] Vinck et al. "An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias," *NeuroImage*, vol. 55, no. 4, pp. 1548-1565, Apr. 2011.
- [16] A. L. Goldberger et al. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215-e220, 2000.
- [17] G. Schalk, S. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A General-Purpose Brain-Computer Interface (BCI) System," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034-1043, 2004.
- [18] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert, "xDAWN algorithm to enhance evoked potentials: application to brain-computer interface," *IEEE Biomedical Engineering, IEEE Transactions on*, vol. 56, no. 8, pp. 2035-2043, 2009.
- [19] B. Rivet, H. Cecotti, A. Souloumiac, E. Maby, and J. Mattout, "Theoretical analysis of xDAWN algorithm: application to an efficient sensor selection in a P300 BCI," *IEEE Signal Processing Conference, 19th European*, pp. 1382-1386, 2011.