David C. Wyld
Jan Zizka (Eds)

Computer Science & Information Technology

Fifth International Conference on Advanced Information Technologies and Applications (ICAITA 2016)
Dubai, UAE, November 12~13, 2016

AIRCC Publishing Corporation

Volume Editors

David C. Wyld,
Southeastern Louisiana University, USA
E-mail: David.Wyld@selu.edu

Jan Zizka,
Mendel University in Brno, Czech Republic
E-mail: zizka.jan@gmail.com

ISSN: 2231 - 5403
ISBN: 978-1-921987-58-8
DOI : 10.5121/csit.2016.61301 - 10.5121/csit.2016.61308

This work is subject to copyright. All rights are reserved, whether whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the International Copyright Law and permission for use must always be obtained from Academy & Industry Research Collaboration Center. Violations are liable to prosecution under the International Copyright Law.

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

The Fifth International Conference on Advanced Information Technologies and Applications (ICAITA 2016) was held in Dubai, UAE, during November 12~13, 2016. The Fifth International Conference on Data Mining & Knowledge Management Process (CDKP 2016), The Second International Conference on Control, Modeling and Computing (CMC 2016), The Second International Conference on Software Engineering (SOFT 2016) and The Fifth International Conference on Soft Computing, Artificial Intelligence and Applications (SAI 2016) was collocated with the ICAITA-2016. The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The ICAITA-2016, CDKP-2016, CMC-2016, SOFT-2016, SAI-2016 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Committees led to an exciting, rich and a high quality technical conference program, which featured high-impact presentations for all attendees to enjoy, appreciate and expand their expertise in the latest developments in computer network and communications research.

In closing, ICAITA-2016, CDKP-2016, CMC-2016, SOFT-2016, SAI-2016 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the ICAITA-2016, CDKP-2016, CMC-2016, SOFT-2016, SAI-2016.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

David C. Wyld
Jan Zizka

Organization

General Chair

Natarajan Meghanathan	Jackson State University, USA
Dhinaharan Nagamalai	Wireilla Net Solutions, Australia

Program Committee Members

Abd El-Aziz Ahmed	Cairo University, Egypt
Abdallah Rhattoy	Moulay Ismail University, Morocco
Abdolreza Hatamlou	Islamic Azad University, Iran
Adnan H. Ali	Institute of Technolgy, Iraq
Ali El-Zaart	Beirut Arab University, Lebanon
Amitabha Sinha	West Bengal University of Technology, India
Ammar M.Odeh	University of Bridgeport, USA
Ashraf A. Shahin	Cairo University, Egypt
Ashraf Darwish	Helwan University, Egypt
Atif Farid Mohammad	University of North Carolina,Charlotte
Ayad Salhieh	Australian College of Kuwait, Kuwait
Barbaros Preveze	Cankaya University, Turkey
Brajesh Kumar	Indian Space Research Organization, India
Cherif Foudil	LESIA Laboratory Biskra University, Algeria
Dedi Iskandar Inan	State University of Papua, Indonesia
Dias N. G. J. ,	University of Kelaniya, Sri Lanka
Dongchen Li	Peking University, China
Ederval Pablo Ferreira	Federal Institute of Espirito Santo (IFES), Brazil
Edgardo Palza	University of Quebec, Canada
Fazlul Haque K. M.,	Daffodil International University, Bangladesh
Guoqing Xiao	Hunan University, China
Hazlina Haron	Universiti Utara Malaysia, Malaysia
Hoda Mokhtar	Cairo University, Egypt
Hoshang Kolivand	Universiti Teknologi Malaysia, Malaysia
Hossein Jadidoleslamy	MUT University, Iran
Ing. Habil. Natasa Zivic	University of Siegen, Germany
Isa Maleki	Islamic Azad University, Iran
Israashaker Alani	Ministry of Science and Technology, Iraq
Jesus Pardillo Vela	University of Alicante, Spain
Jose Raniery	University of Sao Paulo, Brazil
Kanti Prasad	University of Massachusetts Lowell, USA
Keneilwe Zuva	University of Botswana, Botswana
Kosai Raoof	Universite du Maine, France
Kyaz majrashi	Institute of Public Administration, Saudi Arabia
Lei Wu	University of Houston, USA
Lian Liu	University of Kentucky, USA
Mahdi Mazinani	Islamic Azad University, Iran

Marina Marjanovic	Jakovljevic Singidunum University, Serbia
Meachikh	University of Tlemcen, Algeria
Mehrdad Jalali	Mashhad Azad University, Iran
Melih Kirlidog	Marmara University, Turkey
Mezati Messaoud	University of Ouargla, Algeria
Mohammed AbouBakr Elashiri	Beni Suef University, Egypt
Mohammad Zunnun Khan	Integral University, India
Narasimha I.V.,	University of Houston, USA
Natasa Zivic	University of Siegen, Germany
Noureddine Hassini	University of Oran, Algeria
Rafah M.Almuttairi	University of Babylon, Iraq
Rahil Hosseini	Islamic Azad University, Iran
Rangiha Mohammad	City University London, UK
Rastgarpour M	Science and Research University, Iran
Reda Mohamed Hamou	Dr Tahar Moulay University of Saida, Algeria
Revathi V	Adhiyamaan College of Engineering, India
Rim Haddad	Innov'com Laboratory, Tunisia
Ritambhra Korpal	University of Pune, India
Saad M. Darwish	Alexandria University, Egypt
Salah Al-Majeed	University of Essex , UK
Samadhiya	National Chiao Tung University, Taiwan
Satria Mandala	Universiti Teknologi Malaysia, Malaysia
Seyed Ziaeddin Alborzi	Universite de Lorraine, France
Shahaboddin Shamshirband	University of Malaya (UM), Malaysia
Shahid Siddiqui	Integral University, India
Sharvari Govilkar	University of Mumbai, India
Stefano Berretti	University of Florence, Italy
Subarna Shakya	Tribhuvan University, Nepal
Trisiladevi C.Nagavi	S.J.College of Engineering, India
Venkata Raghavendra	Adama University, Ethiopia
Waqas Nawaz	Innopolis University, Russia
Yahya Slimani	University of Manouba, Tunisia
Yousef Abusetta	Al-Jabal Al-Gharbi University, Libya
Yung-Gi, Wu	Chung Yuan Christian University, Taiwan
Yusmadi	Universiti Putra Malaysia, Malaysia

Technically Sponsored by

Information Technology Management Community (ITMC)

Database Management Systems Community (DBMSC)

Soft Computing Community (SCC)

Organized By

Academy & Industry Research Collaboration Center (AIRCC)

TABLE OF CONTENTS

The Fifth International Conference on Advanced Information Technologies and Applications (ICAITA 2016)

The Fifth International Conference on Data Mining & Knowledge Management Process (CDKP 2016)

The Second International Conference on Control, Modeling and Computing (CMC 2016)

The Second International Conference on Software Engineering (SOFT 2016)

The Fifth International Conference on Soft Computing, Artificial Intelligence and Applications (SAI 2016)

ALTERNATIVES TO BETWEENNESS CENTRALITY: A MEASURE OF CORRELATION COEFFICIENT

Xiaojia He[1] and Natarajan Meghanathan[2]

[1]University of Georgia, GA, USA,
[2]Jackson State University, MS, USA
[2]natarajan.meghanathan@jsums.edu

ABSTRACT

In this paper, we measure and analyze the correlation of betweenness centrality (BWC) to five centrality measures, including eigenvector centrality (EVC), degree centrality (DEG), clustering coefficient centrality (CCC), farness centrality (FRC), and closeness centrality (CLC). We simulate the evolution of random networks and small-world networks to test the correlation between BWC and the five measures. Additionally, nine real-world networks are also involved in our present study to further examine the correlation. We find that DEG is highly correlated to BWC on most cases and can serve as alternative to computationally-expensive BWC. Moreover, EVC, CLC and FRC are also good candidates to replace BWC on random networks. Although it is not a perfect correlation for all the real-world networks, there still exists a relatively good correlation between BWC and other three measures (CLC, FRC and EVC) on some networks. Our findings in this paper can help us understand how BWC correlates to other centrality measures and when to decide a good alternative to BWC.

KEYWORDS

Betweenness Centrality, Random Networks, Small-World Networks, Real-World Networks, Correlation Coefficient

1. INTRODUCTION

Over the past decades, a large number of centrality measures have been introduced and developed to quantify the significance and importance of the nodes in various networks. Betweenness centrality (BWC) is one of the most widely used measures, first developed in the 1970s by Freeman [1] and Anthonisse [2], independently. BWC is a measure of the degree to which a node functions as the mediation node by calculating the fraction score of all shortest paths (geodesic) between other pairs of nodes that go through it. It is expected that the network would be disconnected if one or two nodes with high BWC were removed. Thus one can expect that a node with high BWC does not belong to one of the dense groups, but connects them. For the rest of the paper, the terms 'node' and 'vertex', 'link' and 'edge', 'network' and 'graph' are used interchangeably. They mean the same.

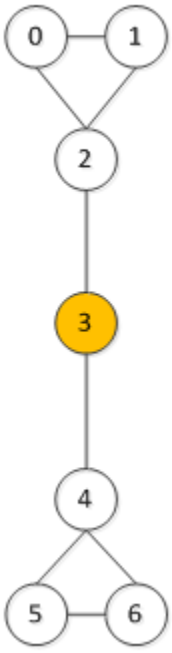
David C. Wyld et al. (Eds) : ICAITA, CDKP, CMC, SOFT, SAI - 2016
pp. 01– 10, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.61301

BWC has been widely applied to a large number of complex network analyses. For instance, it has been proposed as an indicator of the "interdisciplinary" nature of scientific journals [3]. In general, BWC of the nodes in a network increases with connectivity as a power law with an exponent η [4]. Thus, it is known to be computationally time consuming to obtain exact BWC: $O(nm)$ time for unweighted graphs and $O(nm + n^2 \log n)$ time for weighted graphs, where n is the number of vertices and m is the number of edges in the network [5][6][14]. In this paper, we focus on analyzing the correlation between BWC and five well-known centrality measures, including eigenvector centrality (EVC), degree centrality (DEG), clustering coefficient centrality (CCC), farness centrality (FRC), and closeness centrality (CLC). Random network, small-world network, and several real-world networks are involved in this paper.

2. COMPUTATION OF BETWEENNESS CENTRALITY

The computation of BWC in this paper follows the algorithm by Brandes (2001) [5]. If the number of shortest paths between two nodes i and j that pass through node k as the intermediate node is denoted as g_{ij}^{k} and the total number of geodesic between the two nodes i and j is denoted as g_{ij}, then the BWC for node k is defined as

$$BWC(k) = \sum_{i} \sum_{j} \frac{g_{ijk}}{g_{ij}} (i \neq j \neq k)$$



Betweenness of node 2:
Node 3→ 0: 1
Node 3→ 1: 1
Node 4→ 0:1
Node 4→1:1
Node 5→0:1
Node 5→1:1
Node 6→0:1
Node 6→1:1
Betweenness Centrality of node 2:8

Betweenness of node 3:
Node 4→ 0: 1
Node 4→ 1: 1
Node 4→ 2:1
Node 5→0:1
Node 5→1:1
Node 5→2:1
Node 6→0:1
Node 6→1:1
Node 6→2:1
Betweenness Centrality of node 3:9

Betweenness of node 4:
Node 5→0:1
Node 5→1:1
Node 5→2:1
Node 5→3:1
Node 6→0:1
Node 6→1:1
Node 6→2:1
Node 6→3:1
Betweenness Centrality of node 4:8

Betweenness Centrality of node 0: 0
Betweenness Centrality of node 1: 0
Betweenness Centrality of node 5: 0
Betweenness Centrality of node 6: 0

Figure 1: Representative Example to Compute the Betweenness Centrality of the Vertices in a Network

The representative BWC calculation is illustrated in Figure 1. On the basis of the algorithm proposed by Brandes (2001) [5], breadth-first search is involved in the computation. It is clear

that BWC is different to degree-based ranking as shown in Figure 1. Nodes 3 and 4 have highest degree in this present network; however, node 3 has highest BWC. Nodes 0, 1, 5, and 6 each has a degree of 2, but with a BWC of 0.

3. CORRELATION ANALYSIS

3.1. Analysis on Random Networks

Firstly, random networks were simulated to investigate all the six centrality measures including BWC, EVC, DEG, CCC, FRC, and CLC. In this section, networks with 100 nodes were simulated. Particularly, the probability of linkage between nodes is varied from 0.05 to 0.9 to evaluate above mentioned centrality measures. The probability of linkage is increased from 0.05 to 0.1 by 0.01; from 0.1 to 0.9 by 0.1. Representative random networks are shown in Figure 2 with a ranking factor of BWC. Correlation between BWC and other five measures, including DEG, EVC, CCC, FRC, and CLC, was then determined. Average correlation coefficient value was calculated based on 100 trials.

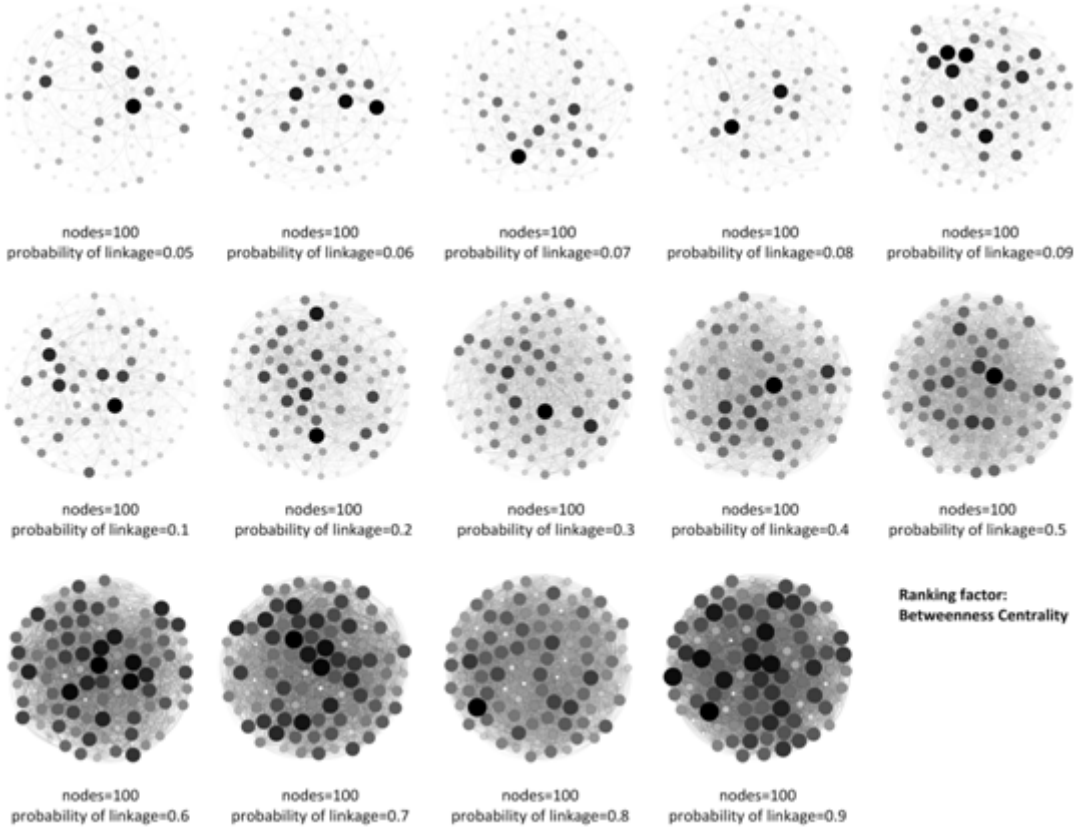


Figure 2: Simulation of Random Networks with Various Probability of Linkage Values
[Ranking Factor is Betweenness Centrality]

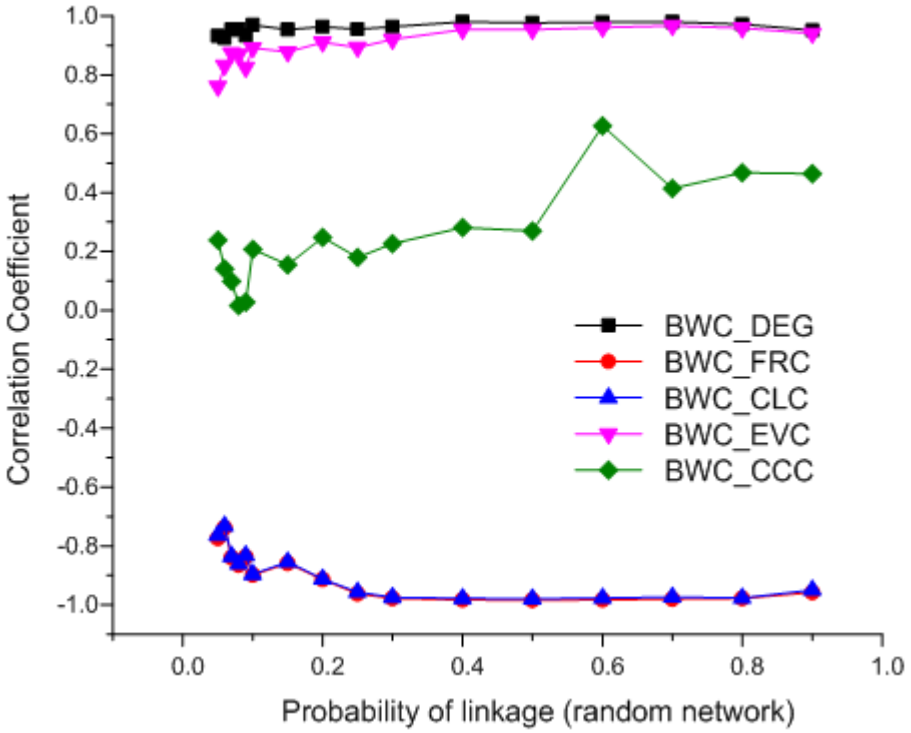


Figure 3: Correlation Coefficient between BWC and the other Five Centrality Measures: DEG, EVC, CCC, FRC and CLC on Random Networks with Various Probability of Linkage Values

As shown in Figure 3, BWC is highly correlated to all measures except CCC. Our data suggests a strong correlation between BWC and DEG, ranging from 0.9316 to 0.9513. The highest correlation of BWC to FRC, CLC, and EVC reaches -0.9576, -0.9495, and 0.94, respectively. The negative correlation indicates that an increase in one variable reliably predicts a decrease in the other one. A high value in negative correlation still suggests high correlation. It is pretty sure that we can select DEG, FRC, CLC, EVC as alternatives to BWC in random networks.

3.2. Analysis on Small-World Networks

We investigated on small-world networks evolved from regular network. Similar to random network simulation, 100 nodes with a k-regular value (initial number of links per node) of 10 are set for small-world network simulation. In this section, the probability of rewiring was varied from 0.01 to 0.09 with increment of 0.01; and from 0.1 to 0.9 with increment of 0.1. Representative small-world networks are shown in Figure 4 with a ranking factor of BWC. Correlation between BWC and the other five measures, including DEG, EVC, CCC, FRC, and CLC, was then calculated. Average correlation coefficient value was calculated based on 100 trials.

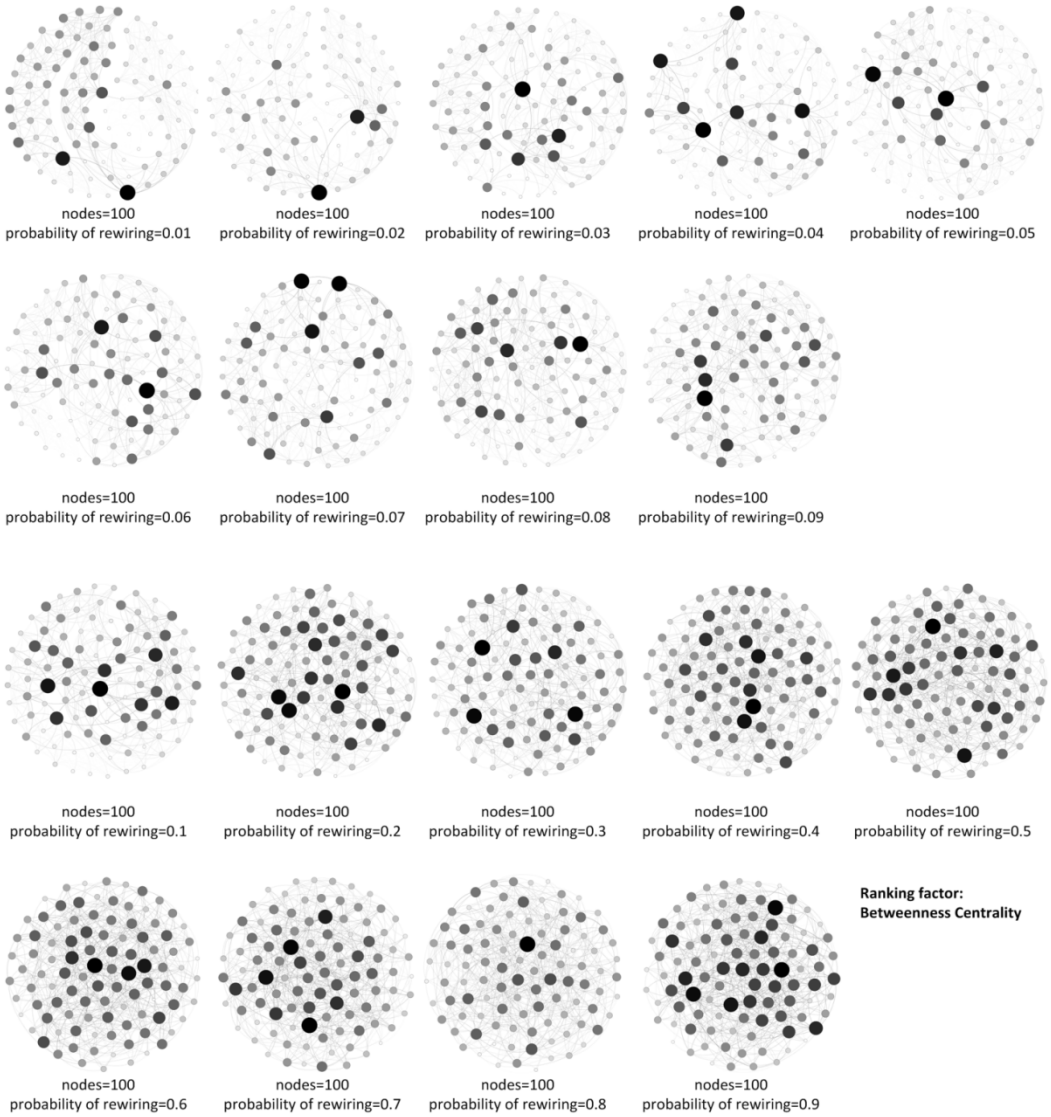


Figure 4: Simulation of Small-World Networks with Various Probability of Rewiring Values [Ranking Factor is Betweenness Centrality]

For small-world networks, there is a strong correlation between BWC and the other centrality metrics, except EVC, at a probability of rewiring lower than 0.2. The correlation coefficient was larger than 0.51 when the probability of rewiring reaches 0.2 for DEG, FRC, CLC, and CCC. The highest correlation coefficient of BWC to DEG, FRC, and CLC reaches to 0.5325, -0.7499, and -0.7348 at probability of rewiring of 0.08. The correlation between BWC and CCC decreases from 0.8131 to 0.0683 along with the increase of probability of rewiring.

In a previous work, a transformation between small-world network and random network was revealed [15]. It was found that simulated network from a regular network would be small-world network when the probability of rewiring is from 0.01 to 0.1; however, it changes to random network when the probability of rewiring is between 0.1 and 1.0. In this study, we also observed a clear turning point at probability of rewiring of 0.1 as shown in Figure 5. Overall, we could

preferably use CCC as alternative to BWC at probability of rewiring lower than 0.07. At a critical probability of rewiring lower than 0.2, we still could use DEG, FRC, CLC, and CCC as alternatives.

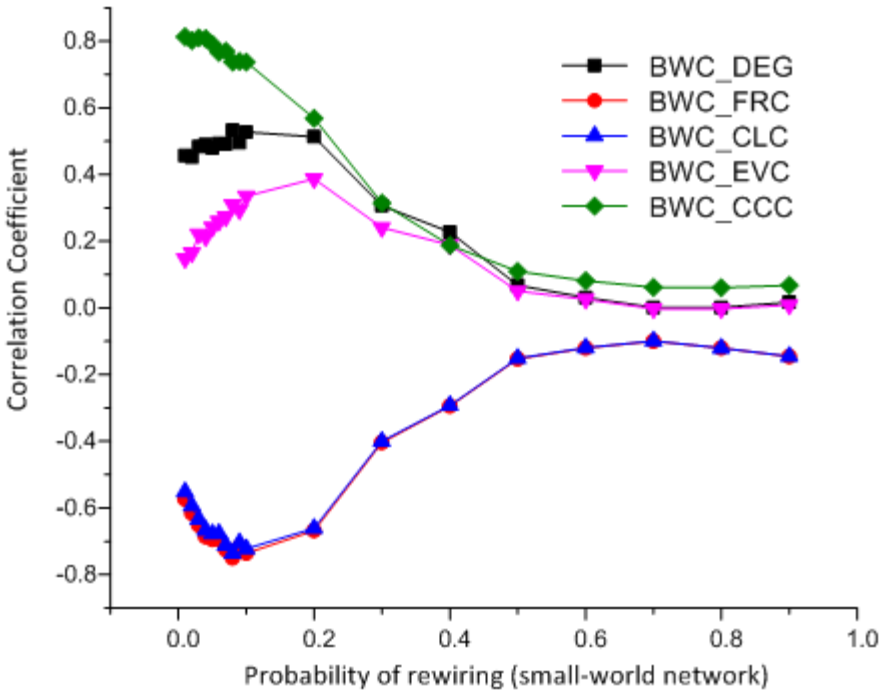


Figure 5: Correlation Coefficient between BWC and the other Five Centrality Measures, including DEG, EVC, CCC, FRC and CLC, on Small-World Networks with Various Probability of Rewiring Values

3.3. Analysis on Real-World Networks

In order to evaluate the feasibility of applying the above mentioned candidate centrality metrics to replace BWC practically, multiple real-world networks were also studied. Analysis on real-world networks is crucial to understanding how BWC relates to other measures in real world. In this study, nine real-world networks (see Figure 6) were analyzed. These are: Dolphins social network (Dolphins), Word adjacency network of common adjectives and nouns in the novel David Copperfield by Charles Dickens (WordAdj), Celegensmetabolic network representing the metabolic network of C. elegans (Celegm), Celegensneural network representing the neural network of C. elegans (Celegn), American football games network between Division IA colleges during regular season Fall 2000 (Football), Karate Social network of friendships between 34 members of a karate club at a US university in the 1970 (Karate), LesMis Coappearance network of characters in the novel Les Miserables (LesMis), the 1997 US Airports network (AirNet), and Political books network (BookNet). Average correlation between BWC and other five measures, including DEG, EVC, CCC, FRC, and CLC, was determined on 100 trials.

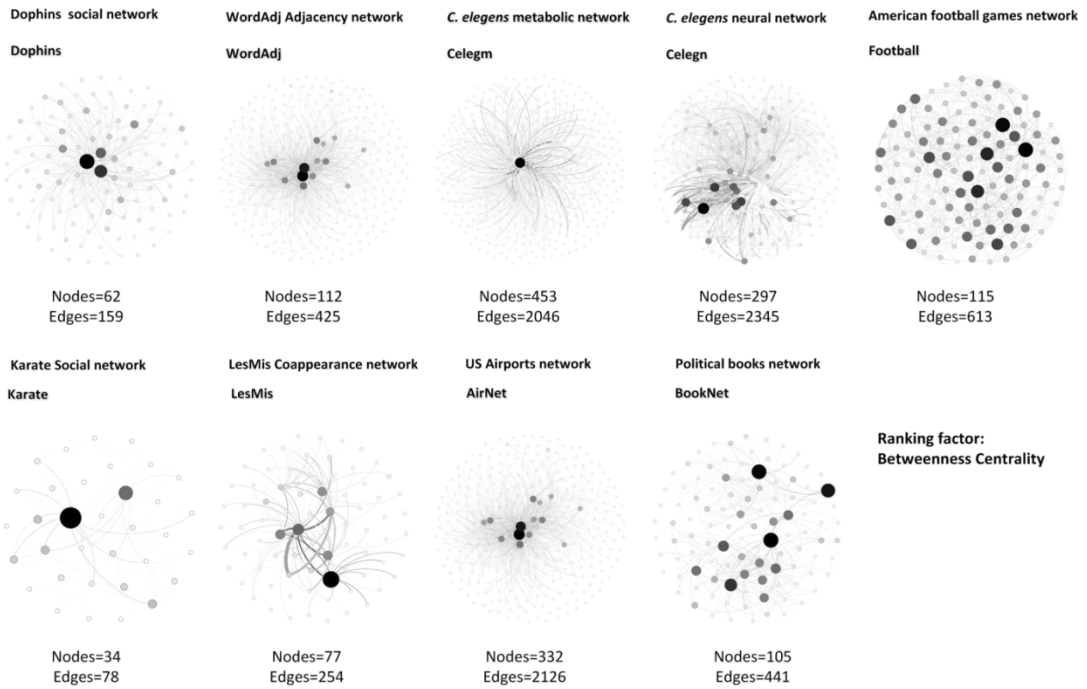


Figure 6: Distribution of the Nodes in Real-World Networks
[Ranking Factor: Betweenness Centrality]

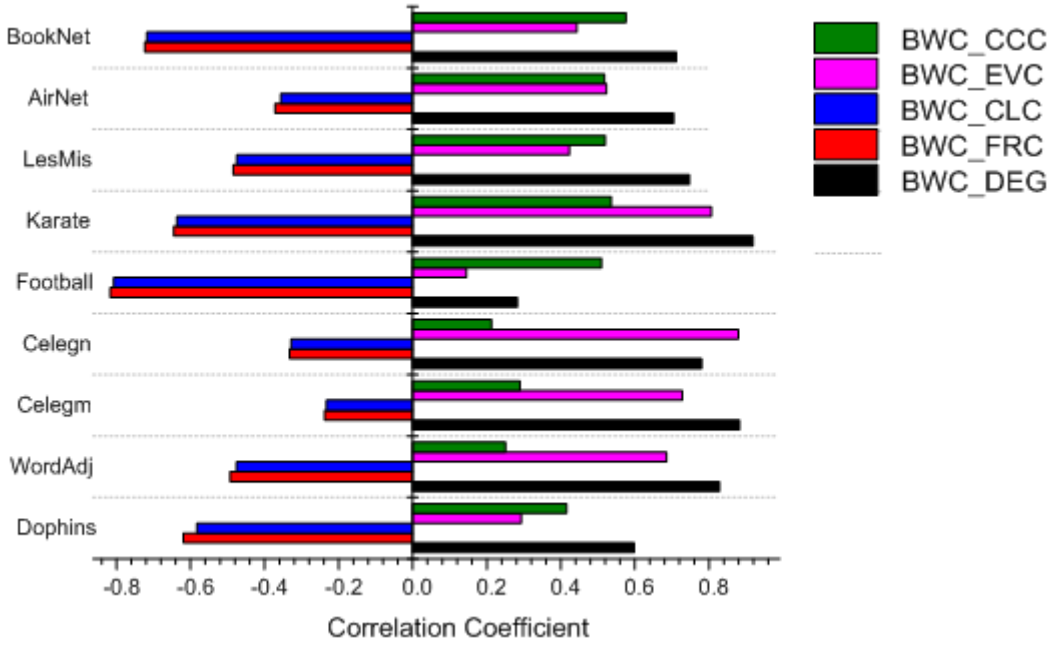


Figure 7: Correlation Coefficient between BWC and the other Five Centrality Measures, including DEG, EVC, CCC, FRC and CLC on Real-World Networks

Unlike the random and small-world networks, the correlation of BWC to CLC and FRC is relatively low with a correlation coefficient value less than 0.6 for five tested networks out of nine (WordAdj, Celegm, Celegn, LesMis, and AirNet). Similarly, correlation coefficient between BWC and EVC is also relatively low with a value less than 0.6 for five tested networks including Dolphins, Football, LesMis, AirNet, and BookNet. Particularly, the correlation coefficient between BWC and EVC on Football network only shows a value of 0.14. It is also similar to BWC and DEG with a correlation coefficient of 0.28 on Football network. Notably, the correlation coefficient between BWC and CCC is lower than 0.6 on all tested networks. It is noteworthy that BWC correlates well with DEG on all but Football network. On Football network, BWC has a high correlation with FRC and CLC

4. RELATED WORK

Recently, Meghanathan (2016) proposed a hybrid centrality metric (takes both the degree and the shortest paths into account) called the local clustering coefficient-based degree centrality (LCCDC) [10]. The local clustering coefficient (LCC) of a vertex is a measure of the probability that any two neighbors of the vertex are connected. If a vertex has a larger LCC value, then the neighbors of the vertex can directly communicate with each other rather than going through the particular vertex. If the neighbors of a vertex do not need go through the vertex for shortest path communication, then it is more likely that the rest of the vertices in the network would not need to go through the vertex for shortest path communication. If a vertex has a smaller LCC, then the neighbors of the vertex are more likely to go through the vertex for shortest path communication between themselves (as there is more likely not a direct edge between the two neighbors, because of the low LCC for the vertex). More specifically, if a vertex has a low LCC and a high degree, then several of the neighbors of the vertex (and as a result, several of the two-hop, three-hop neighbors and so on) are more likely to go through the vertex for shortest path communication. Such vertices are expected to have a higher BWC. The LCCDC metric captures such high BWC vertices (with a strongly positive correlation) and could be used to rank the vertices in a graph in lieu of the BWC. Since the strongly positive correlation between BWC and LCCDC has been already studied in [10], in this paper, we explore any of the other well-known centrality metrics exhibit a strong correlation with BWC.

There are some other algorithms proposed to further develop the application of BWC. For instance, the random-walk betweenness measure calculated for all vertices in a network in worst-case time $O((m+n)n^2)$ using matrix methods [8]. Others such as bounded-distance betweenness [9], distance-scaled betweenness [9], edge betweenness [11] and group betweenness [12] are also introduced. Alternatively, an approximation computation of BWC of a given vertex with an adaptive sampling technique is discussed in the paper by Bader et al (2007) [7]. Nevertheless, the computation cost of these betweenness measures is still high. It is more feasible if we could find another centrality measure with low computation cost that is highly correlated to BWC. It was shown that the BWC is related to the degree in social networks [13] and scale-free network [14]. However, there still lacks substantial support on the alternatives to BWC.

BWC measures the interrelationships among vertices. The results of our simulation studies suggest that BWC is highly correlated to DEG on most tested networks. Leydesdorff (2007) [3] also observed high correlation between BWC and DEG with a correlation coefficient value of 0.724 on Journal Citing Social Networks. Recently, Pozzi et al (2013) [16] observed a strong correlation of the centrality indices between unweighted BWC and DEG calculated on Planar

Maximally Filtered Graphs (PMFG) with a value of 0.97. There is also a moderate correlation between BWC and CLC papered with a value of 0.54 [3]. CLC refers to the relatedness among a set of vertices, providing a global measure of relationships among all vertices. A good correlation between BWC and CLC is valuable when it comes to a connection between global and local view.

5. CONCLUSIONS

In this paper, we analyzed the six commonly studied centrality measures on random networks, small-world networks, and multiple real-world networks. A clear correlation of BWC to DEG is shown on most tested networks. It is safe to conclude that there is a strong correlation between BWC and DEG. In addition, FRC, CLC and EVC can also serve as alternatives to BWC in random network. For small-world networks, DEG, FRC, CLC and CCC could be preferably used as alternative to BWC at probability of rewiring lower than 0.2. Unlike the random and small-world networks, BWC is relatively less correlated to CLC and FRC on five tested real-world networks out of nine. DEG still is one of the best alternatives to BWC on real-world networks. In conclusion, we have found the computationally-cheap DEG as a good candidate to replace computationally-expensive BWC on most occasions.

REFERENCES

[1] L. C. Freeman, "A Set of Measures of Centrality based on Betweenness," Sociometry, vol. 40, no. 1, pp. 35-41, March 1977.

[2] J. M. Anthonisse, "The Rush in a Directed Graph," Stichting Mathematisch Centrum. Mathemaatische Besliskunde (BN 9/71), pp. 1-10, 1971.

[3] L. Leydesdorff, "Betweenness Centrality as an Indicator of the Interdisciplinarity of Scientific Journals," Journal of the American Society for Information Science and Technology, vol. 58, no. 9, pp. 1303-1319, 2007.

[4] M. Barthelemy, "Betweenness Centrality in Large Complex Networks," European Physical Journal B, vol. 38, pp. 163-168, 2004.

[5] U. Brandes, "A Faster Algorithm for Betweenness Centrality," The Journal of Mathematical Sociology, vol. 25, no. 2, pp. 163-177, 2001.

[6] M. E. J. Newman, "Scientific Collaboration Networks. II. Shortest Paths, Weighted Networks, and Centrality," Physical Review E, vol. 64, 016132, June 2001.

[7] D. Bader, S. Kintali, K. Madduri and M. Mihail, "Approximating Betweenness Centrality," Algorithms and Models for the Web-Graph, Lecture Notes in Computer Science, vol. 4863, pp. 124-137, 2007.

[8] M. E. J. Newman, "A Measure of Betweenness Centrality based on Random Walks," Social Networks, vol. 27, pp. 39-54, 2005.

[9] S. P. Borgatti and M. G. Everett, "A Graph-Theoretic Perspective on Centrality," Social Networks, vol. 28, no. 4, pp. 466-484, October 2006.

[10] N. Meghanathan, "A computationally lightweight and localized centrality metric in lieu of betweenness centrality for complex network analysis," Vietnam Journal of Computer Science, pp. 1-16, June 2016.

[11] M. E. J. Newman and M. Girvan, "Finding and Evaluating Community Structure in Networks," Physical Review E, vol. 69, 026113, February 2004.

[12] M. G. Everett and S. P. Borgatti, "The Centrality of Groups and Classes," The Journal of Mathematical Sociology, vol. 23, no. 9, pp. 181-201, 1999.

[13] K-I. Goh, E. Oh, B. Kahng and D. Kim, "Betweenness Centrality Correlation in Social Networks," Physical Review E, vol. 67, 017101, January 2003.

[14] K-I. Goh, B. Kahng and D. Kim, "Universal Behavior of Load Distribution in Scale-Free Networks," Physical Review Letters, vol. 87, 278701, December 2001.

[15] N. Meghanathan, "Distribution of Maximal Clique Size under the Watts-Strogatz Model of Evolution of Complex Networks," International Journal in Foundations of Computer Science and Technology, vol. 5, no. 3, pp. 1-12, May 2015.

[16] F. Pozzi, T. D. Matteo and T. Aste, "Spread of Risk across Financial Markets: Better to Invest in te Peripheries," Scientific Reports, vol. 3, no. 1665, 2013.

THE IMPACT OF EXISTING SOUTH AFRICAN ICT POLICIES AND REGULATORY LAWS ON CLOUD COMPUTING: A LITERATURE REVIEW

Mpho Mohlameane[1] and Nkqubela Ruxwana[2]

[1]The Da Vinci Institute, 16 Park Avenue, Modderfontein, South Africa
mpho.mohlameane@gmail.com
[2]Tshwane University of Technology, 2 Aubrey Matlakala St,
Soshanguve, Pretorita, South Africa
ruxwananl@tut.ac.za

ABSTRACT

Cloud computing promises good opportunities for economies around the world, as it can help reduce capital expenditure and administration costs, and improve resource utilization. However there are challenges regarding the adoption of cloud computing, key amongst those are security and privacy, reliability and liability, access and usage restriction. Some of these challenges lead to a need for cloud computing policy so that they can be addressed. The purpose of this paper is twofold. First is to discuss challenges that prompt a need for cloud computing policy. Secondly, is to look at South African ICT policies and regulatory laws in relation to the emergence of cloud computing.

Since this is literature review paper, the data was collected mainly through literature reviews. The findings reveals that indeed cloud computing raises policy challenges that needs to be addressed by policy makers. A lack of policy that addresses cloud computing challenges can negatively have an impact on areas such as security and privacy, competition, intellectual property and liability, consumer protection, cross border and juridical challenges.

KEYWORDS

Cloud Computing, Policy, Law, Regulation.

1. INTRODUCTION

Technology evolution has helped economies around the world to become more competitive. The emergence of Information and Communication technology (ICT) has helped businesses around the world to streamline and improve their business processes in order to respond quickly to customer needs. Of those technology evolution, cloud computing happens to be a technology that when used, it can help reduces capital expenditure and furthermore improve competitiveness globally.

David C. Wyld et al. (Eds) : ICAITA, CDKP, CMC, SOFT, SAI - 2016
pp. 11– 24, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.61302

However, with technology evolution such as cloud computing, there are some noticeable challenges which might hinder the adoption rate and therefore deny companies economy of scale. Key amongst those challenges includes security and privacy, reliability and liability, access and usage restriction. Therefore such challenges necessitate a need for cloud computing policy that when adopted, it can help improve public confidence in the adoption and use of cloud computing and furthermore help improve competitiveness.

In this paper we set the scene by first describing the overview of cloud computing, more specifically defining cloud computing and also a brief discussion of cloud deployment models and services. We then proceed by discussing key challenges that leads to a need for cloud computing policy. We furthermore move to the core aspect of this paper which is the discussion of key South African laws and regulation in relation to cloud computing.

The next section describe the overview of cloud computing, starting with the definition of cloud.

2. OVERVIEW OF CLOUD COMPUTING

2.1. Definition of Cloud computing

Cloud computing can be briefly defined as computing over the internet, whereby services such as data storage, application software are accessible over the internet [1]. The European Community for Software and Services [2] defines cloud computing as "the delivery of computational resources from a location other than your current one"

The National Institute of Standard and Technology (NIST) gave a detailed definition as they define cloud computing as " a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service-provider interaction" [3].

The above definitions characterise cloud computing as convenient and on-demand service, shared resource pooling, rapid elasticity and broad network access [12].

2.2. Cloud Computing Deployment Models

By cloud deployment models, we refer to cloud targeted deployment models such as public, private, hybrid and community.

- **Public Cloud** – refers to computing services that are publicly accessible over the internet by subscribed cloud consumers. Examples of public cloud services includes but not limited to Google Email, Google Drive, Dropbox, Amazon Web Services (AWS), etc. [4] [5].

- **Private Clouds** – unlike public clouds whereby the hosting infrastructure and software application is hosted by third party and there is multi-tenancy of different cloud customers, private clouds are privately owned and accessible privately. In most cases enterprises that have deployed private clouds choose not to publish some of the private and confidential data to public cloud, rather manage such sensitive data internally [4] [5].

- **Hybrid Clouds** – is made up of two or more clouds with the characteristics of both private and public cloud [7], [8], [9]. NIST [3] corroborates the above by defining hybrid cloud as "a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds).

- **Community Cloud** – Institutions with shared mission and goals, computing needs, policies and security requirements can form what is termed "community cloud" as a means to share hosting infrastructure (hardware, network resources, etc.) amongst community members [4], [5], [10].

2.3. Cloud Computing Service Models

Cloud services models refers to computing services that are offered on the cloud. These includes services such as Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Software as a Service (SaaS) [11].

- **Platform as a Service** are services that creates a platform for software application developers to develop, deploy and manage their applications over the internet. Examples of PaaS include but no limited to Amazon Elastic Beanstalk, Microsoft Windows Azure, Google App Engine, and Apache Stratos.

- **Infrastructure as a Service** are services that provides customers with hardware and computing infrastructure that is typically hosted and managed by cloud computing service provider [13]. Examples includes but not limited to Amazon EC2, Google Compute Engine and Rackspace.

Software as a Service is the delivery of software solutions in a form of a service whereby the software is hosted by the service provider and is accessible over the internet. You don't necessary have to install the software locally on your computing device, rather you access it over the internet. The advantage therefore is the licensing model which is subscription based [16]. Examples includes Microsoft 365 office, Google Apps and Salesforce.

3. CHALLENGES THAT LEAD TO A NEED FOR CLOUD COMPUTING POLICY

There are several challenges that lead to a need for cloud computing policy that when addressed, can help improve public confidence and improve the adoption rate. These challenges are discussed below starting with key security and privacy.

3.1. Security and Data Privacy

Cloud services are accessible over the internet and are made up of cloud services such as email services (Gmail, Yahoo mail, etc.), social network (such as Facebook, etc.), storage services such as Dropbox, Google Drive, VMware storage services, etc. These cloud services stores huge

amount of personal and sensitive information at data centers around the world. Due to the fact that these data is accessible over the internet, this raises privacy and security risks.

Cloud computing like other networked systems, is more prone to traditional threats such as data confidentiality, privacy and authentication. This is because cloud computing resources are accessed over the internet and can be exposed to such security threats. Other threats includes but no limited to – abuse and nefarious use, insecure interfaces and APIs, malicious insiders, shared technology issues, data loss or leakage, account and service hijacking.

Conversely, there are security issues which are more specific to the cloud environment and which are exacerbated by the multi-tenancy and distributed nature of the cloud. These include confidentiality, service availability, access control, identity management and privacy amongst the list [14], [15].

3.2. Access to the Cloud and Necessary ICT Infrastructure

Cloud computing due to its cost model, has the potential of enabling Small and Medium Enterprises (SMEs) to have access to the latest software services that will enable them to compete with large enterprises. Conversely, most of non-profit organisations make use of technological tools to advance their social and community developmental course. Since such organisations survive on donor funding, it makes sense to create access points to cloud services at a reduced cost.

Furthermore many countries have broadband challenges as most of it is not strong enough to enable business and general public to make use of cloud computing services, especially in rural places. Moreover, the cost of internet access and data usage tend to be expensive in most of countries, especially in South Africa.

3.3. Competition and Antitrust

The International Telecommunication Union (ITU) attributes the following as cloud computing competition concerns, namely: - data portability, lack of industry standards, public procurement practices and restrictive license condition [16]. Cloud computing deployment models consist of Software as a Service (SaaS), Infrastructure as a Service (IaaS) and Platform as a Service (Pass). ITU assert that “vendor lock-in” can occur within any of these three segments and this can become an impediment regarding data or application movements [16].

Conversely, Sluijs, Larouche and Sauter [17] identified three potential concerns which may constitute anticompetitive behaviour, namely 1) Interoperability and portability concerns, 2) Vertical integration and discrimination and 3) Internalization of market by a group of Internet Service Providers (ISP) to deny cloud computing providers to provide ubiquitous services. Vertical integration and discrimination could occur if the ISP decides to vertically integrate within the cloud computing market, i.e. deciding to provide cloud computing services as part of their differentiation strategy [17].

3.4. Wiretapping and Electronic Surveillance

In 2013, there were revelations made by former United States of Americas National Security Agency (NSA) contractor Edward Snowden regarding the NSA's programs that compromises the privacy of data stored in the cloud, especially the data that stored by USA cloud computing providers such as Microsoft and Google. The revelations by former NSA contractor has sparked public outrage in the cloud computing community. Such public outrage can negatively affect USA cloud providers doing business in other countries. This necessitate a need for transparency in electronic surveillance activities conducted by government agencies in order to ensure that the privacy of citizen is not violated.

3.5. Intellectual Property and Liability

As with on premise software, cloud computing services providers also rely on copyrights and patent laws and other forms of protections for their intellectual property (IP). We have recently witnessed companies such as Apple, Google and Microsoft filling patent and IP court papers seeking recourse due to the violation of their IP. Therefore to promote the development of cloud computing, there is a need for strong policy initiative that will comprehensively address issues of IP and copyright in order to protect cloud computing companies against any infringement of their IP and copyrights.

3.6. Consumer Protection and Vendor Lock-in

Issues of cloud reliability and liability necessitate a need for cloud computing policy regime. Cloud users expect cloud services to be reliable at all time, however this becomes a challenge. For instance cloud provider blackouts can render some of the mission critical applications unavailable.

Van Belle & Hinde [18] argues that Cloud computing risks and challenges from a South African perspective also include cloud lock-in. Cloud/Vendor lock–in is largely used by cloud vendors to technically ensure that consumers heavily depend on cloud vendor services, thus making it difficult for the consumer to easily switch from one vendor to another. Therefore there is a need to mitigate this cloud lock-in in order to ensure that consumers are protected from this and can easily switch from one vendor to another.

3.7. Cybercrime

The cloud computing market is growing rapidly and there is huge amount of data that is being stored on the cloud, and these vaults of data attracts cybercriminals. As cloud computing market grows, it will be very challenging to protect such data from cyber-attacks. This necessitate a need for a policy that will ensure effective law enforcement to deal with such acts of crimes.

3.8. Cross Border and Jurisdictional Challenges

Most cloud computing vendors have data centers around the world and most cloud consumers do not know the location where their information is stored. For instance, a cloud user staying in Nigeria can access his/her data that is stored in Australia. A question worth asking is which laws and regulations apply? The location of cloud provider's data center can have an influence the way

in which users are legally protected. There is a need for policy initiative that will address cross border cloud computing regulatory issues and therefore improve public confidence.

4. SOUTH AFRICAN POLICIES IN RELATION TO CLOUD COMPUTING: ANALYSIS AND DISCUSSION

The recent BSA Global Cloud Computing Scorecard survey done by the Software Alliance [19] revealed a significant improvement of South Africa regarding privacy laws and regulations. This is due to the recent privacy laws such as Protection of Personal Information Act (POPI). While cloud computing is gradually gaining its maturity and the usage increases, new challenges will eventually start to emerge.

For the purpose of this paper, three key regulatory laws are reviewed and discussed, namely POPI, Electronic Communication and Transaction Act, as well as The South African Competition Act. We purposefully selected these three regulatory laws as they are more inclined to the fundamental purpose of this study, as well as to analyse, discuss and interpret under the auspices of the identified challenges that leads to cloud computing policy, which were discussed in the previous section of this paper. Of those identified regulatory laws and regulations, we first start discussing and analysing POPI Act

4.1. Protection of Personal Information Act (POPI)

With the increase of criminal acts such as identity theft and the intrusion of personal privacy, governments all over the world have increasingly become concerned regarding these criminal acts as well as the way in which personal information is collected, handled and processed by companies. South Africa is not an exception to these global concerns of privacy issues, hence there is now a new act called the Protection of Personal Information Act (POPI) which was enacted on 26 November 2013.

The premise of POPI is to regulate the way in which organisations collect and process personal information for business purpose. This is done with the aim of ensuring that individual right to privacy is safeguarded. Therefore gone are the days whereby organisations could obtain individual personal information and process it without the consent of the data subject. The data subject is described as the person to whom the personal information relates [20]. POPI scope is very wide as it applies to almost every activity that one might do with the personal information, including the handling and processing of employees personal information.

The purpose of POPI is to uphold and enhance privacy of personal information as well as prescribing to international standards with regard to data privacy and protection [21]. POPI is influenced by other privacy regulations in countries such as the United Kingdom (UK), Canada, Australia and the European Union [21]. POPI act consists of 8 conditions regarding the lawful processing of personal information act, namely - accountability, processing limitation, purpose specification, further processing limitation, information quality, openness, security safeguards and data subject specification. These 8 principles are discussed below, starting with accountability condition.

Some of the advantages that can be drawn from the POPI Act is that it strives to ensure privacy, safety and confidentiality of data subject personal information. It strives to ensure that data is

processed in a lawful manner and with accountability. Furthermore, it ensures that the data subject participates in the process of processing of personal information. Moreover, it hold the responsible party and the operator accountable with regard to the unlawful processing of personal information. However it is worth mentioning that there are some noticeable challenges regarding POPI act.

POPI act is one of the comprehensive legislation act and compliance by companies is going to be a cumbersome process. The act allows companies to have in place privacy and security policies, however most of the enterprises do not have such policies [21], [22]. According to the survey conducted by ITWeb/Delloite, over 55% of companies participated in the survey did not have security or privacy policies in place [22]. Thus a lack of readiness in ensuring POPI compliance remains a concerns.

Furthermore, there are challenges with chapter 9 of the act which deals with cross-border data transfer. As discussed above, the act allows data to be transferred to other countries provided the data subject gives a concern and the laws and of the foreign country regarding privacy policies are similar to that of the republic. By having a restriction on information flow can affect international trade. From economic point of view, the unrestricted flow of information can ensure the development of global supply chain. While the restriction of information flow from privacy and security point of view makes sense, this has to be minimized to a level that it does not negatively impact global trade.

Another argument the researcher would like to raise regarding restrictive cross-border information flow is that in relation to competition. Most multinational companies like Google, Amazon and Microsoft offers competitive prices for their cloud computing services. Some of these multinationals stores personal information of data subject to countries which might not have privacy laws similar to that of South Africa, and as a result, this might have a competitive disadvantage for their South African market segment.

For instance, you might end up in a situation whereby there are local CSP which has local data centers but charges exorbitant prices compared to multinationals, because they comply with the act. It would be beneficial to maintain a balance between the POPI act and competitiveness, in order to ensure that there is economic growth through competitiveness and furthermore there is privacy laws that do not hinder such competitiveness and economic growth.

4.2. Electronic Communication and Transaction Act (ECTA 2002)

The Electronic Communication and Transaction Act (ETCA) was promulgated in 2002 in order to provide a regulatory framework to address electronic communications and other transactions conducted on the internet. With the increase of internet usage and e-commerce industry in South Africa, the ECTA was purposefully designed and implemented to:

- Facilitate and regulate the use of electronic communications and transactions.
- Necessitate the development of South African national e-strategy by the ministry of communication

- Promote access and use of electronic communication and transaction by Small, Micro and Medium Enterprises (SMMEs).

- Provide a regulatory framework to guard against abuse and malicious use of information systems.

- Promote the use of e-government services.

The ECTA of 2002 is a detailed regulatory act and consist of fourteen chapters outlining the government regulatory position regarding the use of electronic communication and online transaction. Furthermore the act address key factors such as and not limited to - facilitating electronic transactions, registration of cryptography providers, accreditation of authentication services providers, consumer and personal information protection in relation to electronic communication and transaction, liability of service providers and cybercrime.

One of the advantages of ECTA of 2002 is that some of the provisions in the Act are in line with international standards [23]. Another advantage of this Act is the provision for the protection of South African (SA) consumers, even though some critics suggest that the Act does so extensively to a level that it can create some trade challenges with international suppliers doing business with SA consumers [24].

The Act affords consumers with cooling-off period, to cancel without reason and penalty any credit agreement for the supply of goods within 7 days after the date of receipt of goods or services and within 7 days after the conclusion of the agreement. Therefore this serves as an advantage to the consumer as it afford him/her the opportunity to terminate the contract if he/she is not satisfied with the goods or services provided.

Cybercrime has become a global problem and the issue of cybercrime is not only affecting South Africa but other African countries as well [25];[26]. The increase in cybercrime in some countries in Africa is exacerbated by a rapid increase in the use of Information and communication technologies such smartphones, increased bandwidth, adoption of mobile money applications, etc. [27]; [28]; Kritzinger & von Solms, 2012). However the provision of cybercrime in the ACTA of 2002 serves as an advantage as it aims to address issues of cybercrime as well as making proclaiming cybercrime as a criminal offence punishable by jail term or fine.

There are notable challenges regarding this ECTA of 2002. De Villers [23] argued that some terms in the act are not clear, therefore there is a need for clarification of terms in order to increase legal certainty. Another challenge with this act is issues around trans-border jurisdiction and enforcement of judgment [23]; [29]. For instance, a South African consumer cannot institute a legal action against a supplier in a foreign country as the foreign court might refuse to recognize or enforce the ECTA of 2002.

Furthermore another challenges is around the use of cryptography as a form of data security, however with the emergence of cloud computing, the question is whether cryptography is compatible and sufficient enough in ensuring security online, the safety, integrity and authenticity of the data processed in the cloud environment. There seems to be are security challenges regarding the use of cryptographic in the cloud environment [30]; [31];[32]. Due to the diverse

layers of cloud computing, it is complex to manage cryptographic keys. Van Dijk and Juels [33] argued that the use of cryptography alone in enforcing privacy is not sufficient enough.

Another challenge with this act is that it was promulgated a decade ago and there are concerns regarding its applicability with the advancement of technology [34], such as cloud computing, mobile commerce, smartphones, phablets, and so forth. The advancement in technology necessitate a need for policy makers to review and adjust policies. Failure to do this will render some policies inefficient and raise gaps which can then cause problems.

4.3. The South African Competition Act

The South African government has a competition policy which was drafted in the early years of our democracy. The then new administration of the founding father of our democracy, the late Tata Nelson Mandela, brought new policy reforms. Tata is isiXhosa word that means "Father" and it is this word that many South Africans use to refer to the late President Nelson Mandela.
The policy reforms back then were necessary as part of a country's comprehensive program for economic, social and political transformation and as well as the re-integration of South African economy in a global economy after years of exclusion and isolation under the apartheid regime [35]. In October 1998, the South African competition act was legislated and became an act regulating competition in South Africa. The purpose of the South African competition act is to promote and sustain completion in South Africa, by:

- *The promotion of the efficiency, adaptability and development of the economy.*
- *Providing consumers with competitive prices and product choices.*
- *The promotion of employment and the advancement of the social and economic welfare of South Africans.*
- *The expansion of opportunities for South African participation in world markets and recognizing the role of foreign competition in the Republic.*
- *Ensuring that small and medium-sized enterprises have an equitable opportunity to participate in the economy.*
- *The promotion of a greater spread of ownership, in particular to increase the ownership stakes of historically disadvantaged persons.*[36]

From the objectives above, it can be argued that the aim of South African competition act is twofold or rather has a dual objective [37], firstly to promote and sustain competition in South African, and secondly to achieve economic transformation in South Africa and address "the historical economic structure and encourage broad-based economic growth" [38].

One of the advantages of the South African Competition Act is the control of anti-competitive conduct and the promotion of competition within the domestic market [40]. Due to the fact that South African emerged from a highly segregated society and with the economy used to be dominated by large conglomerates [40], it was therefore necessary to design a policy that is inclusive and promotes fair competition. The promotion of competition within the domestic

market includes the participation of previously disadvantaged people (mainly African natives) so that they can also participate in the economy of South Africa.

Furthermore another advantage of the Act is its dual objective, as this competition law is perceived to have exceed its scope of typical competition law as compared to laws of developed countries, as it aims to find a balance between the promotion of competition and development [39]. In addition, another advantage of this Act is that it takes into consideration aspects that are unique to South Africa's development, including but not limited to public interest issues such as empowerment, employment and SMEs [35].

Furthermore the establishment of three institutions (such as the competition commission, competition tribunal and competition appeal court) ensures that there are processes in place with regard to addressing anti-competitive conduct and as well as ensuring a fair trial amongst the parties which might be involved in competitive disputes.

Legh, Staples and Masamba [37] argued that while progress has been made by the competition commission in promoting and furthering the policy and its objectives, there are however some challenges. One of the challenges as noted by Legh, Staples and Masamba [37] is the competition commission authorities' failure to follow due process in enforcing the competition Act.

Furthermore it is of a view of a researcher that the South African Competition Act in its current form is more likely to cause issues with the emergence of cloud computing. This is corroborated by Luciano and Walden [41] as they have noted that competition law challenges are likely to emerge as cloud computing gains its maturity, however they are some areas currently which competition issues have started to emerge. These includes open standards and public procurement, interoperability and public procurement, as well as data portability and data protection.

5. CONCLUSIONS

From the discussion above, it is evident that cloud computing raises policy challenges that needs to be addressed by policy makers. A lack of policy that addresses cloud computing challenges can negatively have an impact on areas such as security and privacy, competition, intellectual property and liability, consumer protection, cross border and juridical challenges. Moreover, other categories identified by Yoo [42] which can have policy implications on the cloud includes Industry structure, Data Centers, Server-related technologies, Access networks and Regulations.

Some of existing policies and regulation reviewed, were enacted decades ago, thus raising a gap which needs to be covered in order to accommodate new technological developments. What is interesting is that such challenges are not only experienced by South Africa, as there are other countries who are stuck with legislation and policy documents that have been enacted decades ago. For instance, some privacy policy acts such as the United States Stored Communication Acts, was enacted decades ago and therefore not sufficient to handle regulatory issues that are being raised by the emergence of cloud computing.

The gap between policy and new technological developments is becoming so significant to a level that some argue that policies regarding information security should be rethought. Thus Cloud computing is very broad term and it is most likely to raise some policy questions which will need

to be addressed. From the discussion above and from a South African perspective, current policies and regulatory frameworks are not conducive enough to address all cloud challenges which impacts adoption of cloud computing services.

It is evident from the discussion above that there is a need for a cloud computing policy framework that adheres to international standards and that will address some of the pressing challenges emanating from the adoption and use of cloud computing. Moreover, there is a need for a well thought cloud computing policy that does not overregulate but addresses cloud challenges with the aim of improving public user confidence in the cloud and rapid adoption rate computing. Conversely, this policy should be interdisciplinary and should include factors such as competitiveness, finance, technology and law [43].

REFERENCES

[1] J. Kaur, A. Sehrawat, and N. Bishnoi. 2014. Survey Paper on Basics of Cloud Computing and Data Security. International Journal of Computer Science Trends and Technology, 2 (3), May 16-19.

[2] European Community for Software Services. 2010. "White Paper of Software and Service Architectures, Infrastructure and Engineering – Action Paper on the Area for the Future EU Competitiveness, "Volume 2: Background information, Version 1.3, [Online]. Available from http://www.euecss.eu/contents/documentation/volume%20two_ECSS%20White%20Paper.pdf [Accessed: 24/03/2011]

[3] National Institute of Standards and Technology (NIST). 2011. "The NIST Definition of cloud computing," Available from: http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf [Accessed: 20/03/2012].

[4] M. Carroll, A. Van der Merwe, and P. Kotze, "Secure cloud computing: Benefits, risks and controls," Proceedings of the 2011 Information Security for South Africa (ISSA 2011) Conference, August 15-17, 2011.

[5] K. Junck and M. Rahman, "Cloud Computing avoids downfall of application service providers" in International Journal of Information Technology Convergence and Services, vol. 1, pp. 1-20, 2011.

[6] S. Ramgovind, M. Ellof, and E.Smith, "The management of security in Cloud computing," Information Security for South Asia (ISSA) [Online]. Available from: http://icsa.cs.up.ac.za/issa/2010/Proceedings/Full/27_Paper.pdf [Accessed: 15/04/2012].

[7] S. Garg and R. Buyya, Green Cloud computing and Environmental Sustainability. The University of Melbourne, 2011.

[8] G. Conway and E. Curry, "Managing Cloud Computing: A Life Cycle Approach," 2nd International Conference on Cloud Computing and Services Science, Porto, 2012.

[9] S. Chowhan and R.Saxena. Customer Relationship Management from the Business Strategy Perspective with the Application of Cloud Computing. In the Proceedings of DYNAA, 2011. vol. 2, no.1.

[10] L. Wenhao, "A community cloud oriented workflow system framework and its scheduling strategy," 2010 IEEE 2nd Symposium on Web Society (SWS), pp.316-325, 2010.

[11] S. Qureshi and M. Kamal. 2011. Role of Cloud Computing Interventions for Micro-Enterprise Growth: Implications for Global Development. In Proceedings of the Fourth Annual SIG GlobDev Workshop: Fourth Annual SIG GlobDev Workshop. Shanghai: Routledge, ISBN 978-0-9826068-2-7.

[12] S. Shahane, R. Kulkarni. 2014. Cloud Auditing: An Approach for Betterment of Data Integrity. International Journal of Soft Computing and Engineering, 3(6), Jan: 107-112.

[13] S. Bhardwaj, L. Jain, and S. Jain, "Cloud computing: A study of infrastructure as a service (IAAS)", International Journal of engineering and information Technology, vol. 2, pp. 60-63, 2010.

[14] M. Mujinga, B. Chipangura. Cloud computing concerns in developing economies. In Paper presented at the 9th Australian Information Security Management Conference, December 5-7, 2011.

[15] M. Ahmed, M. Hossain. 2014. Cloud Computing and Security Issues in the Cloud. International Journal of Network Security and Its Applications, 4(1), Jan: 25-36.

[16] International Telecommunication Union. Demystifying Regulation in the Cloud: Opportunities and Challenges for Cloud Computing. GSR12 Discussion Paper. October 2012.

[17] J. Sluijs, P. Larouche and W. Sauter. Cloud Computing in the EU Policy Sphere: Interoperability, Vertical Integration and the Internal Market, 3 (2012) JIPITEC 12, para 12.

[18] C.Hinde, JP Van Belle. 2012. Cloud Computing in South African SMMEs: Risks and Rewards for Playing at Altitude. International Research Journal of Computer Science Engineering and Applications, 1(1) October.

[19] BSA Software alliance. "2016 Global Cloud Computing Scorecard" [Online]. Available from: http://cloudscorecard.bsa.org/2016/pdf/BSA_2016_Global_Cloud_Scorecard.pdf [Accessed 20/06/2016] M. Patton. Qualitative evaluation and research methods. 3rd ed (Thousand Oaks, CA: Sage Publications, Inc., 2001).

[20] The Department of Justice. "Protection of Personal Information Act" [Online]. Available from: http://www.justice.gov.za/legislation/acts/2013-004.pdf [Accessed 10/06/2016].

[21] PwC. 2011. The protection of personal information bill: The journey to implementation.

[22] D. Kafouris. "Getting with the POPI programme" [Online]. Available from: http://www.itweb.co.za/index.php?option=com_content&view=article&id=69449 [Accessed: 15/02/2016].

[23] M. De Villers, 2004. Consumer Protection under the Electronic Communications and Transactions Act 25 of 2002. Dissertation. Johannesburg: University of Johannesburg.

[24] Geredal, S.L. 2006, 'The Electronic Communications and Transactions Act' in Thornton, L (ed.) Telecommunications Law in South Africa

[25] J. Warner. 2011. Understanding Cyber-Crime in Ghana: A View from Below. 5(1), Jan-July: 736-749.

[26] F.W. Onifade and K.J. Adebayo. 2011. VDetector: Attacking the Attacker towards Combating Phishing and Identity Theft on the Internet. Journal of Information Technology Impact. 11(2): 133-144.

[27] I.Z. Dhlamini Cyber Security Awareness Initiatives in South Africa: A Synergy Approach [Online]. Available from http://researchspace.csir.co.za/dspace/bitstream/10204/5941/1/Dlamini_2012.pdf [Accessed 14/04/2015].

[28] A. Harris, S. Goodman and P. Traynor. 2013. Privacy and Security Concerns Associated with Mobile Money Applications in Africa. Washington Journal of Law, Technology and & Arts. 8(3): 246-264.

[29] C. Erasmus, 2011. Consumer Protection in International Electronic Contracts. Mini-Dissertation. Potchefstroom: North-West University.

[30] S. Dara. 2013. Cryptography Challenges for Computational Privacy in Public Clouds. ACM, p.23- 33.

[31] N.S. Chauhan, A. Saxena, and J.V.R. Murthy. October 2013. An Approach to Measure Security of Cloud Hosted Application. In Proc. of IEEE CCEM 2013. Bangalore, India

[32] M. Chandramouli, R. Iorga and S. Chokhani,"Cryptographic key management issues and challenges in cloud services," NIST Report 7956, pp.1-31, 2013

[33] Marten van Dijk and Ari Juels. On the impossibility of cryptography alone for privacy preserving cloud computing. In Proceedings of the 5th USENIX conference on Hot topics in security, HotSec'10, pages 1–8, Berkeley, CA, USA, 2010. USENIX Association

[34] Z. Jobodwana. 2009. E-Commerce and Mobile Commerce in South Africa: Regulatory Challenges. Journal of International Commercial Law and Technology. 4(4): 287-298

[35] T. Hartzenberg, Competition Policy and Practice in South Africa: Promoting Competition for Development Symposium on Competition Law and Policy in Developing Countries , 26 Nw. J. Int'l L. & Bus. 667 (2005-2006)

[36] South African Competition Commission. "Competition act no. 89 of 1998 [Online]. Available from http://www.compcom.co.za/wp-content/uploads/2014/09/pocket-act-august-20141.pdf [Accessed: 14/07/2015]

[37] R. Legh, J. Staples and M. Masamba. Competition Law Sibergramme [Online]. Available from: http://www.bowman.co.za/FileBrowser/ArticleDocuments/CompetitionLaw_SG_3_of_2012-2.pdf [Accessed: 23/04/2016]

[38] Organisation for Economic Co-operation and Development (OECD). Public Interest Considerations in Merger Control [Online]. Available from: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP/WP3/WD(2016)13&docLanguage=En [Accessed: 05/08/2016]

[39] F. Kronthaler and J.Stephan. 2006. Progressivity and Flexibility in Developing an Effective Competition Regime: Using Experiences of Poland, Ukraine, and South Africa for developing countries. Discussion Paper. Available from: http://nbn-resolving.de/urn:nbn:de:gbv:3:2-4380 [Accessed 08/07/2016]

[40] A. Török. Competition policy reform in South Africa - Towards the mainstream CP model for 'transition' economies in the Third World? [Online]. Available from: http://www.iwh-halle.de/projects/competition_policy/sa_country_study.pdf [Accessed 06/07/2016].

[41] I.Walden and L. Luciano. Ensuring Competition in the Clouds: The Role of Competition Law? (April 7, 2011) [Online]. Available from: SSRN: http://ssrn.com/abstract=1840547 [Accessed: 25/12/2015]

[42] C.S. Yoo. 2011. Cloud Computing: Architectural and policy implications, Review of Industrial Organization 38(4): 405–421.

[43] A. D. Helvacioglu-Kuyucu. 2011. Exploring Policy-Formulation for SMEs in Cloud Computing: The Case of Turkey. IBIMA Business Review. https://www.itu.int/ITU-D/treg/Events/Seminars/GSR/GSR12/documents/GSR12_Cloud_Walden_5.pdf [Accessed 12/01/2016]

AUTHORS

Mpho J. Mohlameane was born in Pretoria on the 19th of September 1982. Mohlameane currently holds National Diploma in Software Development (2004), Bachelor Degree in Software Development (2010) and Master's Degree in Business Information systems (2012). All the qualifications were obtained from Tshwane University of Technology, Pretoria, South Africa. Mohlameane's major field of study is in information systems. He is current employed at State Information Technology Agency (SITA) as a Senior Software Developer.

Nkqubela L. Ruxwana was born in Eastern Cape on the 25th of December 1983. Dr. Ruxwana currently holds National Diploma in Engineering Computer Systems, Bachelor Degree in Computer Systems, and Masters in Information Systems from Tshwane University of Technology, Pretoria, South Africa. Dr. Ruxwana also holds Master's Degree in Business Leadership (University of South Africa) and a PhD in Information Technology from Nelson Mandela Metropolitan University (NMMU), South Africa. Currently he is a Professor and a Research Supervisor at Tshwane University of Technology, with extensive industry and research experience in the domains such as business analysis, strategy, project management, business intelligence, enterprise architecture, cloud computing, ICT4D, and health informatics.

FORMAL MODELING AND VERIFICATION OF MULTI-AGENTS SYSTEM USING WELL-FORMED NETS

Meriem Taibi[1] and Malika Ioualalen[1]

[1]LSI - USTHB - BP 32, El-Alia, Bab-Ezzouar, 16111 - Alger, Algerie
taibi,ioualalen@lsi-usthb.dz

ABSTRACT

Multi-agent systems are asynchronous and distributed computer systems. These characteristics make them also a discrete-event dynamic system. It is, therefore, important to analyze the behavior of such systems to ensure that they terminate correctly and satisfy other important properties. This paper presents a formal modeling and analysis of MAS, based on Well-formed Nets, in order to ensure the absence of any undesired or unexpected behavior. To validate our contribution, we consider the timetable problem, which is a multi-agent resource allocation problem.

KEYWORDS

Multi-agent system, Well-formed Nets, Model Checking.

1. INTRODUCTION

The paradigm of multi-agent systems (MAS) [1] offers an original way of modeling complex system. Therefore, multi-agent systems have been used in several areas, such as telecommunications, finance, Internet, energy, health, embedded systems ... etc. When designing MAS, it is often hard to guarantee the system specifications that have been designed, actually fulfil the requirements, i.e., whether it satisfies the design requirements, especially for critical applications. Tests and simulations have contributed for a long time to validate such systems. However, these techniques allow to investigate only one part of the global behavior. Thus, they differ from the formal verification techniques, which ensure that a property is verified by all possible system executions [2]. Consequently, it becomes crucial to have rigorous methods of formal specification and verification to ensure the safe development of agent based systems. These systems can be critical with no risk of error for some properties, such as security, integrity and robustness. Model checking techniques are widely used in analyzing MASs due to their completeness and automaton [3].

We have already presented in previous works an e-commerce multi-agent system modeling using Colored Petri Nets [4] [5], where some general properties verification was performed using CPN Tools [6] .Unfortunately CPN Tools does not allow verification of specific properties. In addition it suffers from the so called state explosion problem: the number of states in the model grows

David C. Wyld et al. (Eds) : ICAITA, CDKP, CMC, SOFT, SAI - 2016
pp. 25– 38, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.61303

exponentially. In this paper, we present an efficient formal approach for modeling and verifying multi-agent systems, based on Well-formed Nets (WN) and model checking verification using GreatSPN [7] and SPOT [8] tools. The main advantage of Well Formed Nets is the notion of symbolic reachability graph that is composed of symbolic states. A symbolic state is a state representing several concrete states in the state space of the system described by the Petri net. So, much larger state spaces can be represented. Indeed, we present the Agent by class of color and his actions by transition associated by one or by several conditions.

We study in particular the interaction protocols. Interaction protocols enable agents to reach a solution in a quicker way. The agents know the messages they can receive in a given state, the message they can send and the rules that guide their choice in case of non-determinism. The agents thus go faster towards solution. As case study, we take FIPA contract net protocol applied to timetable problem. The timetabling problem is a resource allocation problem. It aims at finding an appropriate timetable for a set of courses to be scheduled within limited resources such as professors, student groups and class time.

There are generally two types of constraints in timetabling: hard and soft constraints. Hard constraints are those that must be satisfied and cannot be violated. For example, a professor can't give two courses at the same time to two different groups. Soft constraints are those that are preferably satisfied, but may be relaxed if necessary in order to meet hard constraints. For reasons of simplification we are interested by the verification of the hard properties which are expressed by the temporal logic used in this work. We begin by defining some atomic propositions that will help us to translate timetabling properties into LTL formula.

This paper is organized as follows. Related work are presented in Section 2. Section 3 details the analysed system and the proposed models. Section 4 describes the verification of the desirable properties and experimental results. Finally, we conclude our paper by giving some perspectives in Section 5.

2. RELATED WORK

Petri Nets (PN) have been successfully used in several areas for the modeling and analysis of distributed systems [9]. Several studies have been proposed to model MAS with Petri Nets. Balague [10], proposed a model for a promotional game of viral marketing on the Internet. She used Stochastic Petri Nets for modeling a multi-agent wish list. Gazdare [11] used Colored Petri Nets (CPN) as a formal method to model a transport MAS with containers, then, simulated and solved the storage problem. Lyu [12] used a Stochastic Petri Net (SPN) model to assess survivability and fault tolerance of mobile agents systems. They use the model for design and evaluation of their proposed agent architecture through simulation.

EL Fallah-Seghrouchni [13], Boukredera [14] and Khosravifar [15] proposed to use the CPN formalism to model interaction protocols. These Petri Net-based approaches provided a MAS's specification to facilitate applications design and implementation. However, they did not address the verification problem of the proposed models. The advantages of having a Petri net model were not exploited. The work presented by Hsieh in [16] proposed a new model called a collaborative Petri net and addressed the question of deadlock and undesirable state avoidance under the contract net protocol. Other Petri Net extensions were proposed in more recent works. [17] defined nested predicate transition nets to analyze multi-agent system and a set of translation

rules that translate the multi-agent model to an executable PROMELA model [18]. Marzougui in [19] proposed an Agent Petri Net to model interactions between agents. The transformation of the obtained model, in an ordinary Petri Net, is also required to analyze the behavioral properties of the system.

Recently, model checking techniques are widely used in analyzing MASs due to their completeness and automaton. So several model checker are proposed for modelling and verifying critical properties of MASs, e.g., MCMAS [20], MCK[21], SPIN [2] and NuSMV [20]. These approaches, however, still have some limitations. Specifically, MCMAS and MCK mainly focus on concurrent systems without stochastic behaviors, which limits their application in unreliable environments or agents with random behaviors. In our work, we use a formal model, based on Well formed Nets, a class of high level Petri Nets, allowing qualitative analysis together with performance evaluation. This special class of high level Petri nets, allows to express symmetrical behaviours, which generates more compact state space.

3. MODELING MULTI-AGENT INTERACTIONS USING WN

An agent is an active and autonomous entity, it perceives its environment and interacts with other agents to achieve its goal.

The communication between agents can be structured by the use of protocols, structured descriptions of possible interactions between two or more agents. Protocols are a formalization of processes, which allow the organization of recurring tasks.

Several protocols have been proposed (see for instance the proposal of FIPA[1]).

3.1 FIPA contract net protocol

The contract net protocol [22] is an elementary protocol that facilitates task allocation between a group of agents' roles. In this protocol, there are two different types of roles, an Initiator and a Participant. The finite automata in Fig.1 and Fig. 2 model the different states and transitions of these roles. The Interaction Protocol is composed of a sequence of four main steps, illustrated by the sequence diagrams shown in Fig. 3. The agents must go through the following loop of steps to negotiate each contract.

1. The Initiator sends a "call for proposal" (CFP).

2. Participants who receive the announcement can answer by either a *Proposal* or *reject*.

3. Initiator receives and evaluates proposals, it sends a *Contract* to participant agents, whose proposals are accepted, and *Refuse* to other agents.

4. At the end of interaction, the participant sends to the initiator agent, an *Inform message* to confirm the action achieving, or a *failure message* in a failure case.

[1]FIPA: The Foundation for Intelligent Physical Agents

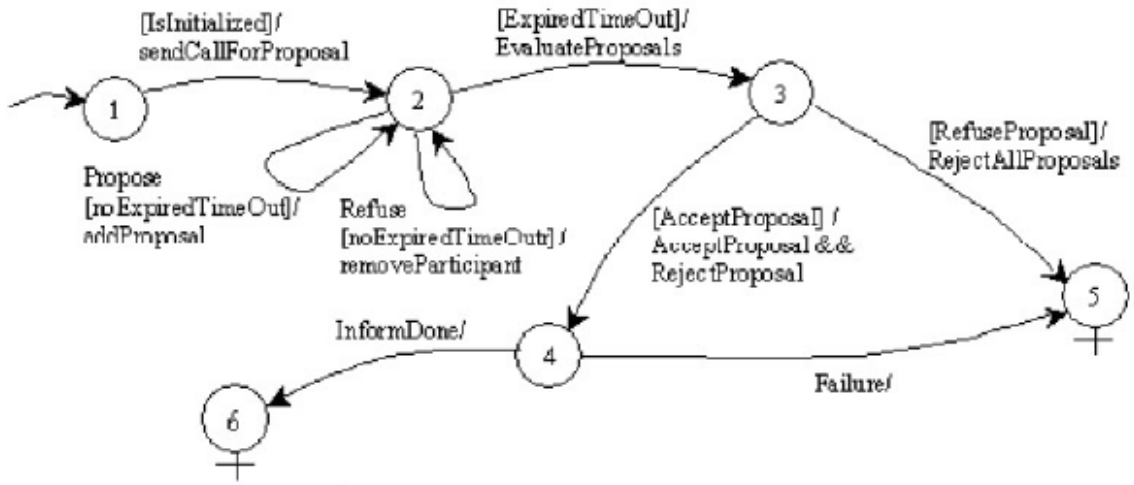


Fig. 1. Initiator automaton

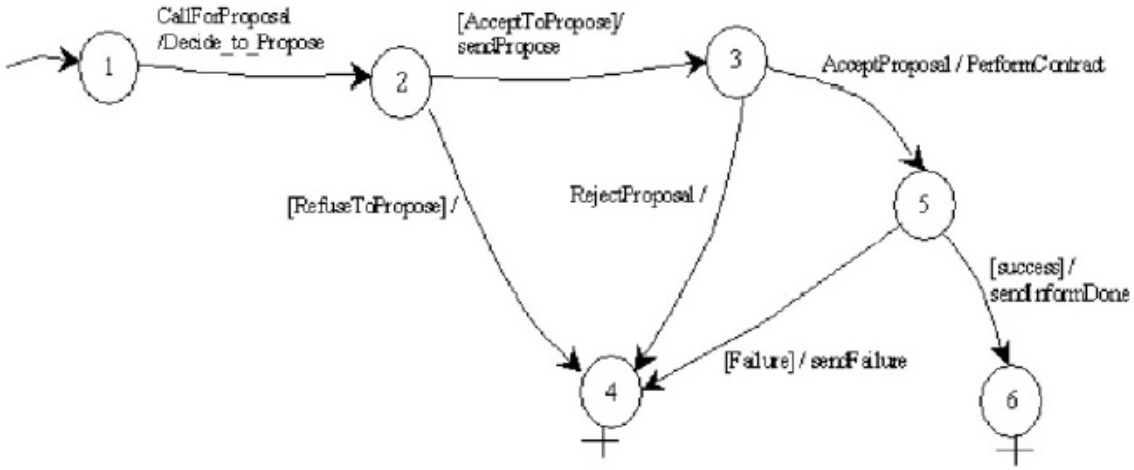


Fig. 2. Participant automaton

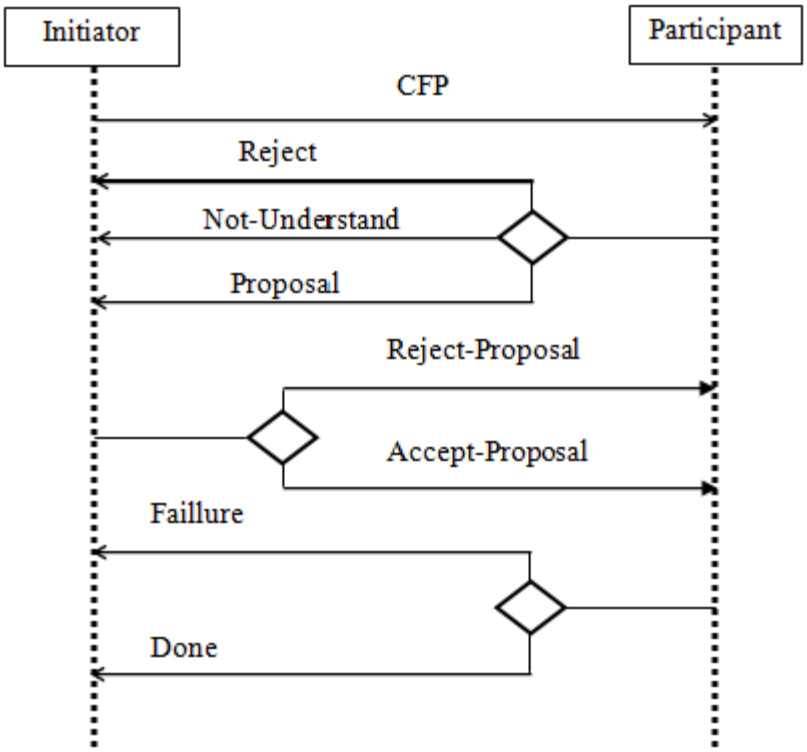


Fig. 3. FIPA Contract-Net Sequence Diagram

3.2 The Well formed Petri Nets (WN)

As mentioned in the introduction, our method is based on the Well-Formed Petri nets. Petri Nets are state based models which are well known for being able to model complex systems with concurrency and conflicts, even in the stochastic context. Moreover, WN model can also take advantage of behaviourial symmetries of systems' entities, if there are such symmetries. Finally, WNs are a well studied class of high level colored Petri nets and benefit from a large set of analysis algorithms and tools

A Well-formed Net [23] is a colored Petri net, where places and transitions are provided with a structured type of tokens. In this model, tokens are grouped into basic classes called color classes. These classes are brought together to form a color domain, which is associated to places and transitions. Colors of a place label its tokens, whereas colors of a transition define possible firings of the transition. Thus, an initial marking of a place is defined as a multiset (bag) of colored tokens. A color function is attached to each arc: its role is to define for, a given color of the associated transition, the number of colored tokens to add or to remove from the attached place.

A color domain is a Cartesian product of color classes. A total order, expressed by a successor function, can be defined on a color class. The Cartesian product defining a color domain can be empty (for example, in the case of a place containing neutral tokens). It can also contain repetition of a class (modelling internal synchronization of this class). A color class, grouping colors of same nature (eg. processes, resources), can be divided into static sub-classes, where a sub-class contains colors with identical behaviours, even in terms of performance.

A color function is built from standard operations (linear combination, composition, etc) of basic functions. The projection (denoted by X or X_i^j in figures) selects an element of a tuple; it is represented by a typed variable or by X if no confusion is possible. The synchronization/diffusion (denoted by Si or Si, k) returns the set of all colors of a class (Si) or a sub-class (Si, k). The successor function is defined for ordered classes only and returns the color following a given color.

A transition or an arc function can be guarded by an expression which is a linear combination of atomic predicates. An atomic predicate expresses the equality of two variables, or restricts the color domain of a variable to a static subclass. A predicate is evaluated on colors of a transition firing.

The structured definition of a WN allows us to exploit automatically system symmetries, by compacting its reachability graph, leading to a Symbolic Reachability Graph (SRG). An SRG is composed of symbolic markings, where each symbolic marking represents a set of ordinary (colored) markings having equivalent behaviours. Several qualitative properties can be checked on the SRG (reachability of a marking, deadlock freeness, etc.)

Formally, a well-formed Petri Net N is a tuple $(P, T, C, cd, Pre, Post, Inh, Guard, Pri, M_0)$ with [23]:

- P,T: the finite sets of places and transitions,
- $C = C_i / i \in I = 1, , n$: the set of basic color classes; C_i is possibly partitioned into n_i static sub-classes: $C_i = \bigcup_{j=1}^{n_i}$,
- $cd : P \bigcup T \rightarrow Bag(I).cd(r) = C_1^{e_1} \times C_2^{e_2} \times \cdots \times C_n^{e_n}$ is the color domain of a node r; $e_i \in IN$ is the number of occurrences of C_i in the color domain of r, where $Bag(I)$ is the set of multisets (bags) on I.
- $Pre, Post, Inh$: the input, output and inhibition standard color functions from $C(t)$ to $Bag(C(p))$.
- $Guard(t) : C(t) \rightarrow true, false$ is a standard predicate associated with the transition t. By default, $Guard(t)$ is the constant function of value $True$.
- $Pri : T \rightarrow IN$ the priority function. By default, we assume $\forall t \in T, Pri(t) = 0$;
- $M0 : M0(p) \in Bag(C(p))$ is the initial marking of p.

3.3 Case study: Timetabling management benchmark

The Timetabling management helps users (teachers and student groups) to set their course sessions. Each user has an agent assistant who manages his schedule. The problem is described by:

– Teacher agent: The goal of such agent is to ensure teacher load, by fixing hours and courses he must have with all classes. The teacher agent has to instantiate and activate the initiator role of FIPA Contract Net, to determine every hour of its timetable, and hence his behavior can be described by the following rules:

  - If the teaching load is not empty, then we activate the Initiator role.
  - If the teaching load is empty, we suspend the agent teacher.

– Student-group agent: The student-group agents need to define their sessions with teachers. For their part, the student-groups must answer the teachers requests, so they instantiate and activate the participant role in the FIPA Contract Net protocol.

Our modeling, using WN, is essentially based on the description given in the section 3.1. Well formed Nets offer powerful expression as Colored Petri Net and also reduction characteristics to construct symbolic state space. The model is composed of transitions, places and arcs, indeed we present an agent by a token and its actions by transitions:

Structural representation

– Places represent agents states (before and after sending or receiving operations). Places contain also the different exchanged messages.
– Transitions allow to express sending and receiving messages or some other processing actions.
– Tokens express the different roles and the various exchanged messages.
– Incoming arcs labels specify data required for firing the associated transition.
– Outgoing arcs labels specify data produced by a firing transitions.

Token coloration To differentiate tokens, we use four color classes:

– The color class Te defines the set of teachers agents,
– The color class Cl defines the set of classes or student-groups,
– The class Hr defines the different time-slots (hours). To simplify the management, we consider that the time-slots have the same duration (2 hours for example).
– The color class Cr allows to definite the different courses, offered by teachers or asked by students.

We associate to each place the color domain hat can mark it, and with each transition the color domain for which it is fired.

Description of models The global MAS model is shown in Fig. 4. Initially classes and teachers agents are inactive in *Classes* and *CFP* places respectively. Each teacher agent sends a call for proposal message to all classes agents by the *Broadcast* transition. One call is composed of sender (x), receiver (y) and proposed course (a).

Once the call is received by class agent, two cases can be distinguished:

- The proposed course is different from the requested one, in this case the offer of the teacher agent is rejected via the transition *Send-Reject*.
- The proposed course is asked by agent class, in this case, a *Proposal* is sent to teacher through the transition *Send-Prop*.

the proposal is treated according to two other cases:

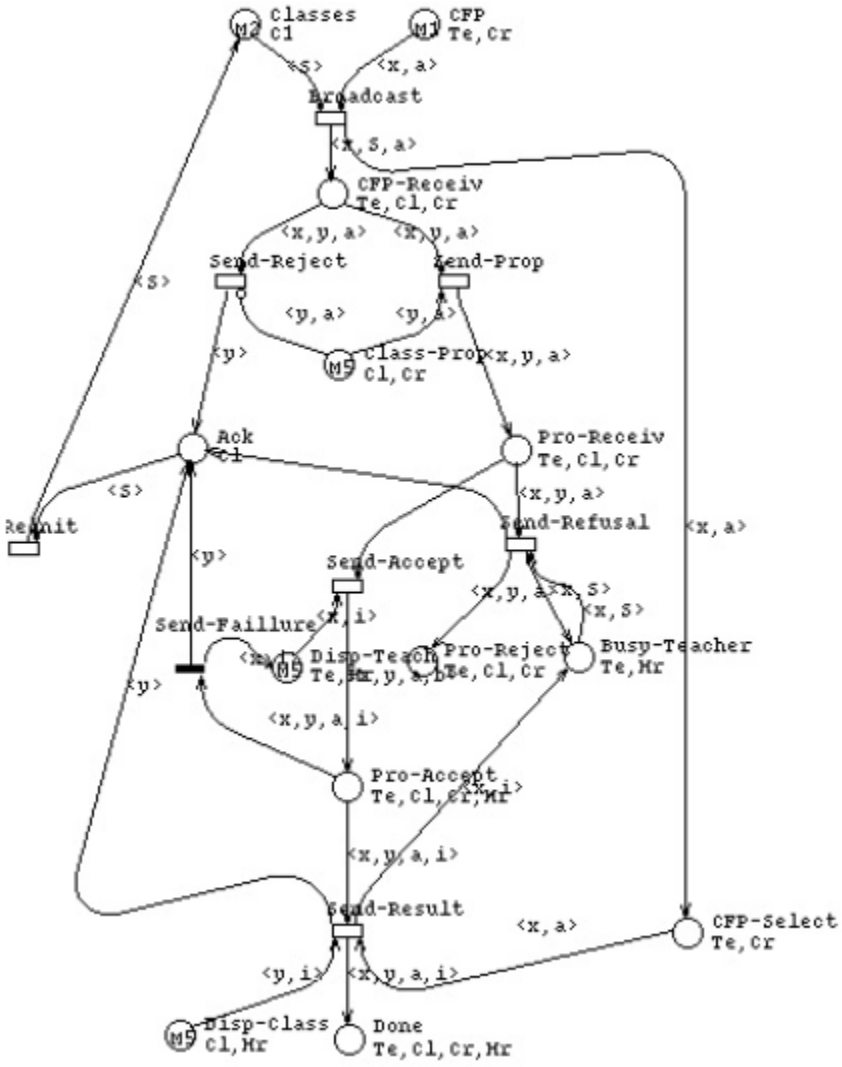


Fig. 4. The MAS model

- Teacher has no availability, in this case a *Refusal* is sent (*Send-Refusal* transition).
- The teacher answers by an acceptance and proposes a time-slot (i) to class agent ((*Send-Accept* transition)). If the student-group has the same availability, the interaction will end by a success and an affectation is created in the place (*Done*). Otherwise the interaction ends by a failure (*Send-Failure* transition).

4. EXPERIMENTAL RESULTS

In this section, we check first some properties and then we analyse the obtained numerical results.

4.1 Properties verification

As discussed earlier, we use SPOT model checker to verify some properties, We begin by defining some atomic propositions that will help us to translate timetabling properties into LTL.

- P_1: the teacher t_1 is assigned to the group g_1 for the time slot h.
- P_2: the teacher t_1 is assigned to the group g_2 for the time slot$h$.
- P_3: the teacher t_2 is assigned to the group g_1 for the time slot h.
- P_4: the teacher t_1 receives more then proposal for his call for the course a.
- P_5: the teacher t_1 accepts one proposal for the course a.

Then the properties can be written down as follows:
Absence of conflict A situation of conflict is detected when the same group is assigned to two different teachers in the same time-slot or when the same teacher is assigned to two different groups in the same hour. These two properties are expressed respectively by the following expression, F_1 and F_2:

$$F_1 : G(!(P_1 \wedge P_3))$$
$$F_2 : G(!(P_1 \wedge P_2))$$

This means that the place *Done* is never marked with two tokens as (t_1, g_1, c_1, h) and (t_2, g_1, c_2, h) (resp. (t_1, g_1, c_1, h) and (t_1, g_2, c_2, h)) in the same time.

Absence of deadlock: A deadlock is detected when there are some offers for the same CFP and none of them will be accepted. The absence of deadlock is expressed as follows:

$$F_3 : G((P_4 \Rightarrow P_5))$$

This property may simply be expressed as: if the place $Pro - Receiv$ is marked by at least two tokens, then the place $Pro - Accept$ will also be marked.

After expressing properties, their verification is split as follows:

(a) Computation of the state space,
(b) Translation of the negated formula (the properties are expressed in LTL),
(c) Synchronized product of the former two objects,
(d) Emptiness check of the resulting product (This step allows to calculate a counter example when the property is not verified). The print screen in Fig.5 displays the empty set obtained after verification.

Moreover, we declare a set of initial states. The system starts at the waiting state for protocol initialization.

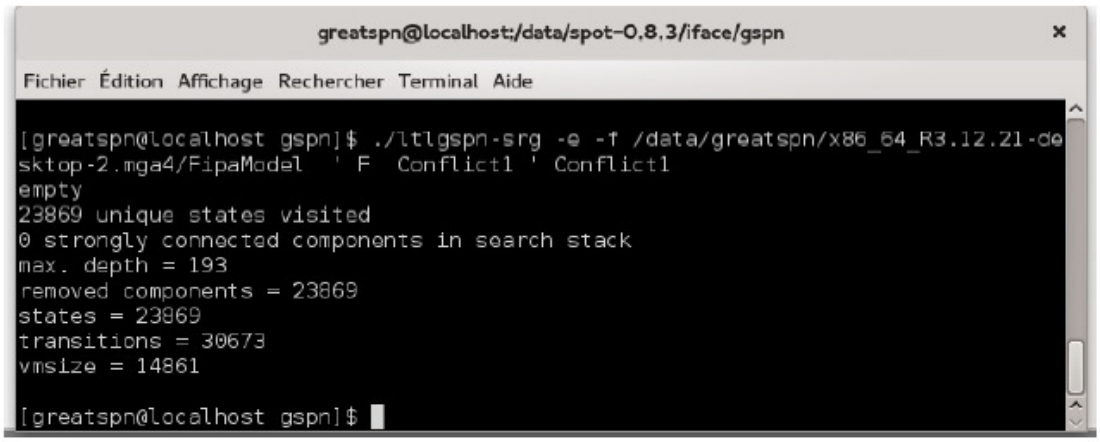


Fig. 5. Property verification

4.2 State Space Analysis Results

Contrary to the former works, in our modeling Agents are represented by tokens, WN allowed not only the use of colors classes to differentiate agents' roles but also to regroup similar behavior. The main advantage of Well Formed Nets is the notion of symbolic reachability graph that is composed of symbolic states. A symbolic state is a state representing several ordinary states in the state space of the system described by the Petri net. So, much larger state spaces can be represented (the gain factor can be up to exponential).

The experimental results of our example are presented in Table 1. The table shows the number of ordinary and symbolic states according to agents number . The gain factor was able to exceed 10000 states grouped in one symbolic state (the gain factor is calculated as: $(Ord\_number - Sym\_number) \div Ord\_number$). A considerable improvement is also noticed comparing these results with thus obtained in [5], where the same protocol was modeled with hierarchical Colored Petri net and the state space generation became very slow from 5 agents, as shown in Table 2. .

5. CONCLUSION AND FUTURE WORKS

Multi-agent systems, are a topic of research in such diverse areas like marketing, e-commerce, artificial intelligence and operational research. The high complexity of these systems forces designers to use formal methodologies associated with automated tools to analyze their behavior. In this paper, we presented a Well-formed Net model for time tabling problem, using FIPA Contract Net Protocol. This type of problem is modeled for the first time by Petri Nets. We used GreatSPN connected to SPOT model checker nets

Nb agent (Teachers,Classes)	Sym. states number	Ord. states number	Gain fact. (%)	Time (s)
2 (1,1)	19	51	62,75	0
4 (2,2)	167	1815	3231,37	0
6 (3,3)	658	36655	70582,35	5
8 (4,4)	1825	5,55E+05	1,08E+06	15
10 (5,5)	3784	7,22E+06	1,41E+07	30
16 (8,8)	23869	1,044E+10	2,05E+10	171
32 (16,16)	346485	3,08E+179	6,04E+179	2658

Table 1. Experimental results with WN

Nb agent	State Space Nodes	Time (s)
3	99	0
4	5954	5
5	39618	422

Table 2. Experimental results with CPN

to verify some properties. Furthermore, the state space analysis showed the efficiency of WN and symbolic states generation to model such systems and highlighted the scalability open issue.

We are interested in our future work by analyzing a larger-scale multi-agent systems. We will extend our modeling and analysis by introducing the temporal dimension in order to perform a quantitative analysis and compute system performances, such as average waiting time, average affectation number, etc.

REFERENCES

[1] Ferber, J.: Les systèmes multi-agents vers une intelligence collective. Inter-Editions (1995)

[2] Thierry-Mieg, Y., Baarir, S., Duret-Lutz, A., Kordon, F.: Nouvelles techniques de model-checking pour la vérification de systèmes complexes. Revue Génie Logiciel (69) (2004) 17–23

[3] Song, S., Liu, Y., Zhang, J., Sun, J.: An extensive model checking framework for multi-agent systems. In: Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems. AAMAS '14, Richland, SC, International Foundation for Autonomous Agents and Multiagent Systems (2014) 1645–1646

[4] Taibi, M., Ioualalen, M., Salmi, N.: Analyzing e-commerce multi-agent systems using hierarchical colored petri nets. In: ICT Innovations 2012 Web Proceedings ISSN 1857-7288. (2012)

[5] Taibi, M., Ioualalen, M.: Analyzing negotiation in e-commerce multi-agent system based on fipa protocol. In: ICT Innovations 2013 Web Proceedings ISSN 1857-7288. (2013)

[6] : Cpn tools. http://cpntools.org

[7] Balbo, G., Conte, G., Donatelli, S., Franceschinis, G., Marsan, A.M.: Modelling with Generalized Stochastic Petri Nets. John Wiley & Sons Ltd (Import) (November 1995)

[8] Duret-Lutz, A., Poitrenaud, D.: Spot: an extensible model checking library using transition-based generalized buchi automata. In: IN PROC. OF MASCOTS04, IEEE Computer Society (2004) 76–83

[9] Celaya, J.R., Desrochers, A.A., Graves, R.J.: Modeling and Analysis of Multi-agent Systems using Petri Nets. JCP (2009, pp. 981-996)

[10] Balague, C.: Les Systèmes multi-agents en marketing : Modélisation par les réseaux de Petri. PhD thesis, École des Hautes études Commerciales (2005)

[11] Kefi-Gazdare, M.: Optimisation Heuristique Distribuée du Problème de Stockage de Conteneurs dans un Port. PhD thesis, ECOLE CENTRALE DE LILLE (2008)

[12] Lyu, M.R., Chen, X., Wong, T.Y.: Design and evaluation of a fault-tolerant mobile-agent system. IEEE Intelligent Systems **19**(5) (Sept 2004) 32–38

[13] El Fallah Seghrouchni, A., Haddad, S., Mazouzi, H.: A formal study of interactions in multi-agent systems. International Journal of Computers and their Applications **8** (2001) 23–32

[14] Boukredera, D., Aknine, S., Maamri, R.: Modeling temporal aspects of contract net protocol using timed colored petri nets. In: STAIRS. (December 2012) 83–94

[15] Khosravifar, S.: Modeling multi agent communication activities with petri nets. International Journal of Information and Education Technology, Singapore **3**(3) (September 2013, pp. 310–3014)

[16] Hsieh, F.: Model and control holonic manufacturing systems based on fusion of contract nets and petri nets. Automatica **40**(1) (2004) 51–57

[17] Chang, L., He, X.: A methodology to analyze multi-agent systems modeled in high level petri nets. International Journal of Software Engineering and Knowledge Engineering **25**(7) (2015) 1199

[18] Holzmann, G.J.: The model checker spin. IEEE Trans. Softw. Eng. **23**(5) (May 1997) 279–295

[19] Marzougui, B., Barkaoui, K.: Interaction Protocols in Multi-Agent Systems based on Agent Petri Nets Model. International Journal of Advanced Computer Science and Applications (IJACSA) 4(7) (2013) 166–173

[20] Lomuscio, A., Qu, H., Raimondi, F.: Mcmas: A model checker for the verification of multi-agent systems. In: Proceedings of the 21st International Conference on Computer Aided Verification. CAV '09, Berlin, Heidelberg, Springer-Verlag (2009) 682–688

[21] Gammie, P., van der Meyden, R.: Mck: Model checking the logic of knowledge. In: Computer Aided Verification, 16th International Conference, CAV 2004, Boston, MA, USA, July 13-17, 2004, Proceedings. (2004) 479–483

[22] : Fipa contract net interaction protocol. http://www.fipa.org/specs/fipa00029

[23] Chiola, G., Dutheillet, C., Franceschinis, G., Haddad, S.: Stochastic well-formed colored nets and symmetric modeling applications. IEEE Trans. Computers **42**(11) (1993) 1343–1360

MODEL CHECKERS –TOOLS AND LANGUAGES FOR SYSTEM DESIGN- A SURVEY

Shubha Raj K B and Suryaprasad J

Department of Computer Science and Engineering,
PESIT-Bangalore South Campus, Bengaluru, Karnataka, India
{shubharajkb, surya}@pes.edu

ABSTRACT

For over four decades now, variants of Model Checkers are being used as an approach for formal verification of systems consisting of software, hardware or combination of both. Though various model checking tools are available like NuSMV, UPPAAL, PRISM, PAT,FDR, it is difficult to comprehend their usage for systems in different domains like telecommunication, automobile, health and entertainment. However, industry experts and researchers have showcased the use of formal verifications techniques in various domains including Networking, Security and Semiconductor design. With current generation systems becoming more complex, there is an urgent need to better understand and use appropriate methodology, language and tool for definite domain. In this paper, we have made an effort to present Model checking in detail with relevance to available tools and languages to specific domain. For novices in the field, this paper would provide knowledge of model checkers languages and tools that would be suitable for various purposes in diverse systems.

KEYWORDS

Formal Methods, Formal Verification, Model Checkers, System Modelling

1. INTRODUCTION

The cutting edge technologies in some of the critical systems like Cyber Physical Systems, Mobile Cloud Computing, Wireless Sensor Network and Mobile Crowd Sensing systems have high degree of complexity. The complexities of these systems are due to the challenges in conformance of software, hardware, network, telecommunication and mobile industry. Large software systems comprise of several million lines of source code. Additional complexities could include structural, environmental (Reactive, Ubiquitous, context-aware), application domain and communication complexity. Hence, it is difficult to understand the requirements, architecture, design, implementation and testing of these systems unless a precise engineering notion is used. One of the systematized tactics that is prevalent to achieve reliable and correct system is Formal Methods (FM) [1][2][3].

Formal methods are techniques used to model complex systems as mathematical entities [18]. They are required for specifying and verifying the system and to ensure that the system is developed "correctly". Formal methods may not be suitable for all types of applications like; problems over simple domains are usually less complex and do not warrant formal methods. Formal methods are vital whenever the cost of failure is high in business critical systems, safety critical systems and machine critical systems [9].

David C. Wyld et al. (Eds) : ICAITA, CDKP, CMC, SOFT, SAI - 2016
pp. 39– 51, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.61304

The development of safety-critical systems requires the use of formal methods for specifying and analyzing critical components and their properties. Formal methods are used at all stages of System development that is at Specification [2][6], Architecture [11][22][13], Design[24][8], Coding[15][16][17] and Testing[18][9].

Typically Formal methods are employed to

- Verify the mathematic model at any time of system development
- Find errors at early stage
- Reduce system cost, development time and effort to build the system
- Improve quality of the system like security, reliability and performance and so on

An application which primarily uses the traditional structured development techniques may use formal methods only for the purpose of documenting data dictionaries. The objectives of these applications will have different impacts on the development process and consequently will influence different choices of model checking techniques. In Formal Methods, mainly we have two main steps. One is Formal Specification and the second one is Formal Verification. Currently there are many techniques that are employed to verify the correctness of system being built like inspection, audit, testing, review, simulation, walkthrough and formal verification. Among these, the Formal Verification techniques offer a formal proof grounded on mathematical model of the system. We need suitable methodologies, tools and languages which can assist in early detection of defects in a structured manner. One such option would be to explore formal verification techniques. There are two types of formal verification; Automatic (Model Checkers) and semi-automatic (Theorem Provers). Here, we focus primarily on Model checkering Languages and Tools with a discussion on current challenges.

There are mainly three major steps in model checking process, namely:

1. System Specification
2. System Modeling
3. System Verification

The rest of the paper is structured as follows: In Section 2, we present one of the inputs of the model checking tool that is Formal specification methods. Section 3 outlines the Formal modelling methods. Section 4 summarizes the formal verification methods. Section 5 briefly discusses the types of model checking methods and their corresponding tools. Section 6 describes model checkers tools and their practical application in various domains. In Section 7, model checking languages are described in detail. Section 8 draws conclusion.

2. SYSTEM (PROPERTY/FEATURE) SPECIFICATION

To apply formal methods, first we should know the characteristics of the problem domain and the complexity of their modeling [4. Typically, we can specify the properties of the system using temporal logic and automata. Table 1 list the distinct types of Temporal Logics and their abbreviations.

Table 1. Distinct Types of Temporal Logics.

Temporal-based PSL (Property Specification Languages)		Stochastic Logics
LTL	Linear Temporal Logic	Continuous Stochastic Logic (CSL)
CTL	Computational Temporal Logic	Probabilistic Reward Computation Tree Logic (PRCTL)
CTL*	Combination of LTL and CTL	Continuous Stochastic Reward Logic (CSRL)
PLTL	Probabilistic LTL	Continuous Stochastic Logic (CSL)
PCTL	Probabilistic CTL	Continuous Stochastic Logic (CSL)
TCTL	Timed CTL	Probabilistic Reward Computation Tree Logic (PRCTL)

Temporal logic is a variant of modal logic for expressing temporal modalities and representing propositions qualified in terms of time. Depending upon the system types, we have to select the distinct types of temporal logic. LTL is more suitable for specifying sequential systems. Whenever we have to verify branching cases in some of the states, then CTL is more suitable. When systems are stochastic in nature, during that time Probabilistic CTL and Probabilistic CTL are appropriate PSL. When we need branching CTL and TCTL are more applicable. Stochastic model may be continuous or discrete. Some of the stochastic logics are CSL, PRCTL and CSRL.

3. SYSTEM MODELLING METHODOLOGIES

Inputs to Model Checking are system modelling and system specification. Once we know the system specifications from the requirement document, next step is to model the system. We can model the system using Finite Automata or a graph at the early stages of system development like at architecture or design level. During design or at architecture level, we have to model systems without implementation (code). When we have to verify the source code, then we have to extract the model from source code. Software model checkers are suitable for extracting model from the source code. There are different methods to Model system formally as specified in the Table 2. Some are text based and others are graphics based. Graph based include Petrinets, state chart and Statemate.

Table 2. Distinct Methods to model the systems

Approaches for modelling System	Characteristics	
Finite State Machine [8]	It is an abstract method having finite number of states. The modelled system can be in only one state at any particular instance in time. Here we have timed automata and hybrid automata.	
Labelled Transition System (Kripke Structure)[19]	It is a labelled transition graph that can sufficiently capture the temporal behaviour of reactive systems	
Model extraction from code (Software Model checkers)	Model extraction from some of the programming languages like C, JAVA and .NET and their corresponding tools are listed below.	
	C	CBMC (C Bounded Model Checker), BLAST (Berkeley Lazy Abstraction Software Verification Tool), CPAchecker (Configurable Program Analysis Checker), DSVerifier (Digital Systems Verifier), ESBMC, LLBMC (Low-Level Bounded Model Checker), SATABS (SAT-based Predicate Abstraction for ANSI-C)
	JAVA	JavaPathFinder, BANDERA
	.NET	MoonWalker

Process Algebra	It is a framework to model concurrent systems. Under this framework we have CSP, CCS and ACP.
Petrinets, Statechart, Statemate	Graph based formal specification languages

Table 3 provides input regarding three main semantic models according to untimed, timed and stochastic/probabilistic system category. Once we understand the basic fundamental concepts to model a given system, later it is easy to select the model checking tools. Otherwise selecting the tool is very difficult task. During the design time, if our aim is to check time critical system, then we have to use timed transition system.

Table 3. Types of Semantic model and their purposes

Types of Semantic Model	Purpose	MC	Applications
Label Transition System	• For untimed systems Finite state Machine Process Algebraic based Model Graph Transition Model	NuSMV [19], CADENCE SMC, ARC,DIVINE Edinburgh CWB, GEAR,LTSA, LTSmin	Used whenever we need only the concurrency problem without timing and stochastic features
Timed Transition System	• Timed automata • Timed process Algebra	UPPAAL,RED PAT,MRMC	When precise constraints on the timing of events are needed, timed automata are the high-level model.
Probabilistic Semantic Model	For Markov Decision Process	PRISM[21][22], MRMC, [64],CADP[24], Modest, Toolset[25] Mobius [26]	Used to model various sources of uncertainty [20].

4. SYSTEM VERIFICATION

According to ISO/IEC/IEEE 15288, system verification includes set of activities that compares a system or system elements against the requirements, architectures [29] and design characteristics and other properties to be verified including deadlock freeness, safety, fairness and aliveness. System verification is used to establish that the design, product or system under consideration possesses the requisite properties.

The system verification methods can be broadly classified into three categories based on the system components or implementation, namely: (a) hardware verification [34] (b) software verification, and (c) System Integration verification. Quality expectations/predications are very high need in hardware systems due to higher fabrication cost and testing involved. Formal Verification uses mathematical reasoning to guarantee the absence of errors. It is an effective bug hunting technique [20]. Testing checks that system behaves correctly under a finite number of test cases, whereas formal verification is designed to be exhaustive.

Different types of Formal Verification (FV) Techniques

- Model Checker (e.g.: SPIN, UPPAAL, PAT, FDR, NuSMV)
- Theorem Proving (e.g.: PVS, HOL [35])

Theorem Proving is used for system with infinite number of states. Commercial use of automated theorem proving is in integrated circuit design and verification. Some of the companies like Intel and AMD use automated theorem proving to verify that division and other operations are correctly implemented in their processor design. But in this paper our aim is describe model checking techniques.

FV is used in various domains for distinct purposes as below.

- Development of Integrated Circuits
- For verifying Electronic Design Automation (EDA) tools
- Stringent regulations for certifications
- To check the correctness of device drivers, cryptographic and communication protocols
- Embedded Control systems [14].
    - Medical devices such as pacemakers and sensors
    - Communication networks
- Safety critical system
    - Avionic Industry, Nuclear energy, Process control, Robotics, Transport, Medical.

5. TYPES OF MODEL CHECKING

Model Checkers **(MC)** are classified as Modern Model Checkers (Software MC) and Traditional Model Checkers [1]. Another category is based on types of verification algorithms as explained below.

- Explicit Model Checker
- Symbolic Model Checker
- Bounded Model Checker
    - SAT (Propositional Satisfiability): In BMC, Model checking complexity is reduced to a propositional satisfiability problem that can be solved with SAT (Satisfiability) solvers. So the size of state space will decrease and increases the performance (speed).
    - SMT (Satisfiability Modulo Theories) [27]: It is an extension of propositional satisfiability (SAT) which is the most well know constraint-satisfaction problem. It generalizes Boolean satisfiability by adding equality reasoning, arithmetic, fixed-size bit vector, arrays, quantifiers and other useful first-order theories.

- On-line Model checkers [28]
- Parallel Model checkers
- Software Model checkers

Depending upon the choices of the system model [FSM, Process Algebra, GTS] and property specification methods [temporal Logic, automata], different Model checking approaches like Explicit MC, Symbolic MC, Bounded MC and On-line MC (OMC). Parallel model checkers are used in Semiconductor Industry for verifying multicore processor.These distinct model checker are typically classified by how states are stored and manipulated. Table 4 gives brief knowledge about the characteristics of the different model checkers and listed reduction methods used in different types of model checking tools.

Table 4. Model checkers types and their characteristics

Explicit Model Checker	**Symbolic Model Checker**	**Bounded Model Checker**	**On-line Model checkers**
States are indexed directly	Second generation MC. Here state of the system is represents by Boolean Functions.	Bounded because only states reachable within a bounded number of times. If length of the path can't be found at a given length K, then the search is continued for higher than K.	Here parameters of models are continually adjusted to remedy possible modelling faults [18] [17].
Graph algorithms are used to explore the state space starting from the initial state	To resolve the problem of state space explosion by enumerating states symbolically	It is a successor of propositional SAT solvers	It drops the need for models to be accurate far into the future.
It construct State Transition Graph by reclusively generating successor of the Initial state	Boolean formulas are represented using the data structure BDD or OBD to improve the efficiency	Used in semiconductor industry	OMC offers safety assurances for short time frames only and renews these assurances continually during operation
Graph is created using DFS (Depth First Search), BFS (Breadth First Search) or in Heuristic manner		Applicable to system level software	It is used in Medical domain because patient model is like to be inaccurate as the physiology of the human body is complex and differs between individuals.
Methods: Here Partial Order Reduction (POR) is used for state-space exploration problem		Best to find shallow bugs [programs without deep loops]	Example: Heart rate and oxygen in the blood depend on the patient's condition
It depends on extensive search through explicit representation of reachable system states [7].		It supports full counterexample trace. SAT based Bounded Model Checking is typically quicker in finding bugs compared to BDDs.	A generalized model will always slip individual characteristics

Table 5, list the tools developed under each model checking type and their limitation. Depending upon the system types and types of properties we have to verify, we have select more than one model checking tools in system development. DIVINE [30[31] is an example for Parallel Model checker. Limitations of parallel Model checkers are communication and load balancing.

Table 5. List of Model Checking Tools and their limitations

Explicit Model Checker	Symbolic Model Checker	Bounded Model Checker	On-line Model checkers
Tools: SPIN, ZING [7][23] PRISM [21][22] DiViNe	**Tools**: SAL, NuSMV BEBOP MOPED SMART MC CadenceSMV,	**Tools**: **SAT-based:**SAL (Symbolic Analysis Laboratory) NuSMV **SMT**-based [27] Z3, CVC4	**Tools**: Java PathExplorer (JPaX) [32]
Limitation: It suffers from the state-space explosion problem due to the exponential growing of the explicit state space	**Limitation**: Human verifier manually adjusts the order of the state variable. This can impact the size of the Binary Decision Diagram (BDD).So performance will reduce. Solution is BMC	**Limitations**: In SAT procedure, the variables must be Boolean type. Due to this it is inexpressive for industrial problems. Example: Computer Programs variables of type other than Boolean must be encoded into Boolean/bit variables which can result in a large formula. **Solution**: Alternative new technique is SMT (Satisfiability Modulo Theories).	
		Why BMC? SMT-based BMC are more expressive than SAT-based BMC.	**Why OMC:** It permits safety assignment at all times and provides means to react before safety violations occur.

Table 6, showcases few software model checkers and their purpose for different programming langauges as well as solutions to the state space explosion and their purpose. SLAM project is used for making reachability analysis for large sequential C programs mainly to Device Driver [1]. BLAST is to prevent all memory safety violations. JPF is used for verification and testing environment for Java programs.

Various abstraction techniques are used depending upon the usage. Some of the abstarction techniques are modular, lazy, Process counter abstraction, parallel modle checking and slicing methods. Few reduction methods are symmetric reduction, and partial order reduction. To achive good performance advanced optimization methods are now a days adapted in various model checking tools.

Table 6. Software model checkers

Year	Tools	Developed by	State-explosion solutions used
2000	(SLAM project) CHESS MODIST	Microsoft	Modular Abstraction BDD
2002	BLAST	University of California Berkeley	Lazy abstraction BDD
2002	JPF (Java PathFinder) [5]	NASA	• Compression techniques is used to handle big states • Partial Order Method • Symmetric Reduction method • Slicing Abstraction • Runtime analysis Techniques

Some of the demerits of Model checking are listed as below.

- Temporal logic specifications are complex.
- Writing specifications is difficult.
- State explosion is a major problem.

6. MODEL CHECKER TOOLS

Depending upon the characteristics of the system, application domain, here are the some of the examples of application of tools for a particular system as shown in the Table 7. The objectives for applying a formal method to a project must be clearly identified and documented. Though we have described in the previous sections about system specification and modelling methods, for detailed understanding of formal languages, we are describing model checking languages in the next section 7. Various model checking tools are used for distinct domain like communication, embedded system, Software Engineering, hardware and healthcare so on. Here we have listed few purposes like to model concurrent software, real-time systems, clock synchronized protocols, synchronous digital logic, asynchronous systems, consistency of software data structure. Table 7, list few model checking tool names and their abbreviation. For SPIN model checker, Promela is the specification language. It is based on the process algebra. In UPPAAL, for formal modelling timed automata is used for system modelling.

Table 7. MC tools and their abbreviation

Tool	Abbreviation	Developed By
SPIN	Simple Promela Interpreter	Bell Labs
UPPAAL	Uppsala and Aalborg	Uppsala and Aalborg University
NuSMV	New Symbolic Model Verifier	Carnegie Mellon University
FDR	Failures Divergences Refinement	University of Oxford
PRISM	Probabilistic Model Checker	University of Birmingham

Table 8. List of model checker tools and their applicability

MC	Purpose	Domain
SPIN	Used to model concurrent software or asynchronous processes.	Communication protocols
UPPAAL	Used to model real-time systems Formal model and analysis of clock-synchronised protocols in sensor networks based on timed automata [36].	Timed systems
NuSMV	Used to model synchronous digital logic.	Digital Circuits
FDR	Used to model asynchronous systems	
Alloy	Used to analyse consistency of software data structures	Requirement Analysis
Simulink Design Verifier	Used to verify models created in Simulink, a data-flow and state-machine simulation	Hardware Circuits
SAT Solvers[19]	Used to in Electronic Design Automation (EDA) community for checking correctness of Hardware designs mainly in synthesis and verification.	Hardware design
PRISM	Formal model of flooding and gossiping protocols for analysing their performance probabilistic properties [35],Cardiac Pacemakers [37] For automatically verify whether STRAC (Spatio-Temporal Access Control based on Reputation, one policy for the IoT) policies conform to security properties [38]	Healthcare

7. MODEL CHECKERS LANGUAGES

There are two different types of specification languages to construct a system model. One is state based and another is Event based languages [2]. Both the methods are specified in Figure 1.

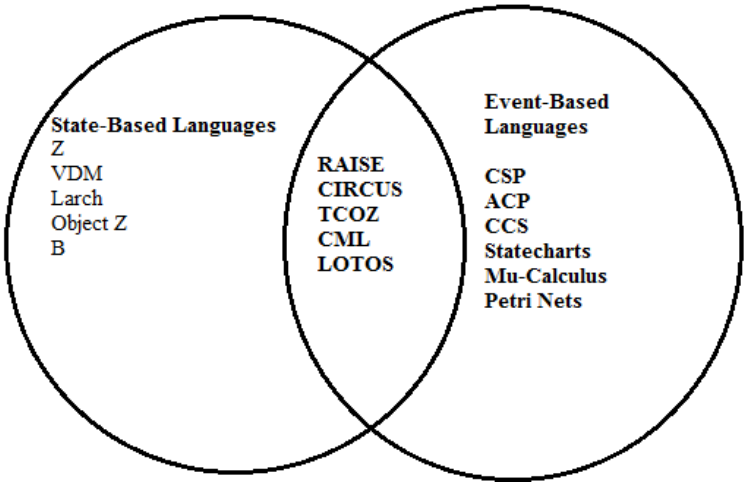


Figure 1. State-based and event-based Specification Languages

Figure 2, lists the important concepts used to develop formal specification languages. Sets, Relations, functions are used in state based and Trees. Graph and automata are used in event based formal specification languages.

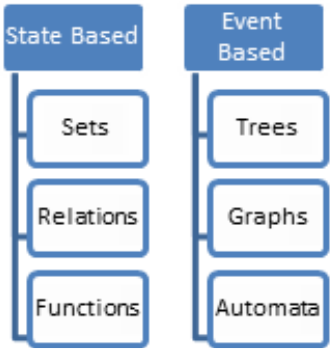


Figure 2. Concepts used in formal specification languages.

Current need is to integrate the distinct specification languages to handle different features of the systems as itemized in Table 8. RAISE is used for handing rich state space. CIRCUS is used for specification, programming, and verification by refinement. Its semantics is grounded on Hoare. It motivated by the need for a notation and techniques to reason about designs and implementations of state-rich reactive processes. Used for reasoning about Safety-critical Java programs (SCJ), avionics control systems, and Systems of Systems. TCOZ supports Object-Oriented principles. CML is a combination of Circus and VDM developed for the modelling of Systems of Systems (SoS).

LOTOS is used for handling concurrency complexities. LOTOS is used for protocol specification in ISO OSI standards LOTOS is an algebraic language that consists of two parts: a part for the description of data and operations, based on abstract data types, and a part for the description of concurrent processes, based on process calculus.

Table 9. Integration of specification languages

State-based Language	**Event-based Language**	**Extension**	**Integrated Specification Language**
VDM	CCS	VDM+CCS	RAISE
Z	CSP		CIRCUS
Object Z	Timed CSP	Object Z +Timed CSP	TCOZ
VDM	CIRCUS	VDM+CIRCUS	CML (COMPASS Modelling Language)
	CSP, CCS	CSP+CCS	LOTOS (Language Of Temporal Ordering Specification) [2]

8. CONCLUSION

As systems get complex, a seamless flow becomes imperative. A formal method helps to reduce errors, cost and ensure that the developed system will meet all expectations. Understanding the role of tools and languages irrespective of the application domain will invariably be a great asset to all concerned. The role of Model Checkers in design and verification of systems with relevant tools and languages is presented in detail. Suitability of Model Checkers tools and languages in various application domains is mainly based on few characteristics of the domain including sequential, parallel, timed, untimed, etc. Novices in the field will get a broad view of role of Model Checkers in system design and verification. Showcasing the usage of respective tool and

language in different domains would be appropriate but is outside the scope of this paper and can be considered for future work.

REFERENCES

[1] Strunk, E. A, Aiello, M. A & Knight, J. C. (2006) "A survey of tools for model checking and model-based development", *University of Virginia*.

[2] Clarke, E. M & Wing, J. M. (1996) "Formal methods: State of the art and future directions", *ACM Computing Surveys (CSUR)*, Vol. *28, No.*4, pp626-643.

[3] Kern, C & Greenstreet, M. R. (1999) "Formal verification in hardware design: a survey", *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, Vol. *4,* No.2, pp123-193.

[4] Liu, Y, Sun, J & Dong, J. S. (2011) "Pat 3: An extensible architecture for building multi-domain model checkers",*IEEE 22nd International Symposium on Software Reliability Engineering* pp190-199.

[5] Zhang, H, Aoki, T, & Chiba, Y. (2015) "Yes! You can use your model checker to verify OSEK/VDX applications", *IEEE 8th International Conference on Software Testing, Verification and Validation (ICST)* pp1-10.

[6] Lamsweerde, A. V. (2000) "Formal specification: a roadmap", *Proceedings of the Conference on the Future of Software Engineering* pp147-159. ACM.

[7] Chen, T, Diciolla, M, Kwiatkowska, M, & Mereacre, A. (2012) "Quantitative verification of implantable cardiac pacemakers", In *Real-Time Systems Symposium (RTSS), 2012 IEEE 33rd* pp263-272

[8] Qadir, J, & Hasan, O. (2015) "Applying Formal Methods to Networking: Theory, Techniques, and Applications", *IEEE Communications Surveys & Tutorials*, Vol. *17. No.*1, pp256-291.

[9] Gaudel, M. C. (1994). "Formal specification techniques", *Proceedings of the 16th international conference on Software engineering*. pp 223-227. IEEE Computer Society Press.

[10] Belinfante, A. (2010) "JTorX: A tool for on-line model-driven test derivation and execution", In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems* pp266-270. Springer Berlin Heidelberg.

[11] Medvidovic, N., & Taylor, R. N. (2000) "A classification and comparison framework for software architecture description languages", *IEEE Transactions on software engineering*, Vol *26*. No.1, pp70-93.

[12] Clements, P. C. (1996) "A survey of architecture description languages", *Proceedings of the 8th international workshop on software specification and design* pp16). IEEE Computer Society.

[13] Hermerschmidt, Lars, et al (2015) "Generating Domain-Specific Transformation Languages for Component & Connector Architecture Descriptions", *2nd International Workshop on Model-Driven Engineering for Component-Based Software Systems (ModComp)* Workshop Proceedings.

[14] Jiang, Y., Zhang, H., Li, Z., Deng, Y., Song, X., Gu, M., & Sun, J. (2015), "Design and optimization of multiclocked embedded systems using formal techniques", *IEEE Transactions on Industrial Electronics*, Vol.*62* No.2, pp1270-1278.

[15] Havelund, K., & Pressburger, T. (2000) "Model checking java programs using java pathfinder", *International Journal on Software Tools for Technology Transfer*, pp366-381.

[16] Corbett, J. C, Dwyer, M. B, Hatcliff, J, Laubach, S, Pasareanu, C. S., & Zheng, H. (2000), "Bandera: Extracting finite-state models from Java source code", *Software Engineering, 2000. Proceedings of the 2000 International Conference on* pp439-448. IEEE.

[17] Ammann, P. E, Black, P. E, & Majurski, W (1998), "Using model checking to generate tests from specifications", *Formal Engineering Methods, 1998. Proceedings. Second International Conference on* pp46-54. IEEE.

[18] Utting, M., & Legeard, B. (2010). "*Practical model-based testing: a tools approach*", Morgan Kaufmann.

[19] Baier, C., Katoen, J. P., & Larsen, K. G. (2008) *Principles of model checking*. MIT press.

[20] Norman, G & David P (2014) "Quantitative Verification: Formal Guarantees for Timeliness, Reliability and Performance".

[21] Barnat, J, Brim, L, Ceska, M, & Rockai, P (2010),"Divine: Parallel distributed model checker", *Parallel and Distributed Methods in Verification, 2010 Ninth International Workshop on, and High Performance Computational Systems Biology, Second International Workshop on* pp4-7. IEEE.

[22] Kwiatkowska, M, Norman, G, & Parker, D. (2011), "PRISM 4.0: Verification of probabilistic real-time systems", *International Conference on Computer Aided Verification.* pp. 585-591. Springer Berlin Heidelberg.

[23] Andrews, T, Qadeer, S, Rajamani, S. K, Rehof, J, & Xie, Y. (2004), "Zing: Exploiting program structure for model checking concurrent software", *International Conference on Concurrency Theory.* pp. 1-15. Springer Berlin Heidelberg.

[24] Petri, C. A. (1986)"Concurrency theory", *Advanced Course on Petri Nets",* pp4-24, Springer Berlin Heidelberg.

[25] Markov Reward Model Checker. www.mrmc-tool.org, October 2016

[26] CADP Home page. www.inrialpes.fr/vasy/cadp/, October 2016.

[27] Phan, Q. S., & Malacaria, P. (2015) "All-Solution Satisfiability Modulo Theories: applications, algorithms and benchmarks' *Availability, Reliability and Security (ARES), 2015 10th International Conference on* pp100-109. IEEE.

[28] Jard, C., & Jeron, T. (1989). "On-line model-checking for finite linear temporal logic specifications", *International Conference on Computer Aided Verification,* pp189-196).Springer Berlin Heidelberg.

[29] Shaw, M, & Garlan, D. (1996), *Software architecture: perspectives on an emerging discipline,* Prentice Hall.

[30] Weiqiang, K. O. N. G., Shiraishi, T., Katahira, N., Watanabe, M., Katayama, T., & Fukuda, A. (2011). An smt-based approach to bounded model checking of designs in state transition matrix. *IEICE transactions on information and systems*, Vol.*94* No.5, pp946-957.

[31] Barnat, J., Brim, L., Havel, V., Havlíček, J., Kriho, J., Lenčo, M., ... & Weiser, J. (2013). "DiVinE 3.0–an explicit-state model checker for multithreaded C & C++ programs", *International Conference on Computer Aided Verification* pp. 863-868. Springer Berlin Heidelberg.

[32] The M¨obius Tool. www.mobius.illinois.edu, accessed on October 2016.

[33] PAT Web site http://pat.comp.nus.edu.sg/ accessed on October 2016

[34] Ardakani, H. H, Gharehbaghi, A. M, & Hessabi, S(2007), "A performance and functional assertion-based verification methodology at transaction-level", *2007 Internatonal Conference on Microelectronics* pp133-136. IEEE.

[35] Fehnker, A, & Gao, P(2006), "Formal verification and simulation for performance analysis for probabilistic broadcast protocols", *International Conference on Ad-Hoc Networks and Wireless* pp28-141. Springer Berlin Heidelberg.

[36] Heidarian, F, Schmaltz, J, & Vaandrager, F. (2012), "Analysis of a clock synchronization protocol for wireless sensor networks", *Theoretical Computer Science*, Vol. *413* No.1, pp87-105.

[37] Chen, T, Diciolla, M, Kwiatkowska, M, & Mereacre, A. (2012) "Quantitative verification of implantable cardiac pacemakers", *Real-Time Systems Symposium (RTSS), 2012 IEEE 33rd* pp 263-272. IEEE.

[38] Yunchuan, G, Lihua, Y, & Chao, L.(2014) "Automatically verifying STRAC policy", *Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on* pp141-142. IEEE.

AUTHORS

Shubha Raj K B She is an Assistant Professor at PESIT-BSC, Bengaluru, Karnataka, India in the Department of Computer Science and Engineering. Currently she is pursuing PhD at VTU Belagavi in Formal Methods. She obtained her MTech degree from RVCE, Bengaluru. Her research interests are in the field of Formal Verification, Model checking, Software Architecture and Architecture Description Languages.

Suryaprasad J He is Director/Principal at PESIT-BSC, Bengaluru, Karnataka, India. He obtained his PhD from Florida Atlantic University, Boca Raton. His research interests are in the field of System level design methodologies, Hardware Software Co-design Embedded System Design, Power Optimal Design, Verification Methodologies and Advanced Programming Metodologies.

INTENTIONAL BLANK

A SURVEY OF MARKOV CHAIN MODELS IN LINGUISTICS APPLICATIONS

Fawaz S. Al-Anziand Dia AbuZeina

Department of Computer Engineering, Kuwait University, Kuwait City, Kuwait
fawaz.alanzi@ku.edu.kw,abuzeina@ku.edu.kw

ABSTRACT

Markov chain theory isan important tool in applied probability that is quite useful in modeling real-world computing applications.For a long time, rresearchers have used Markov chains for data modeling in a wide range of applications that belong to different fields such as computational linguists, image processing, communications,bioinformatics, finance systems, etc. This paper explores the Markov chain theory and its extension hidden Markov models (HMM) in natural language processing (NLP) applications. This paper also presents some aspects related to Markov chains and HMM such as creating transition matrices, calculating data sequence probabilities, and extracting the hidden states.

KEYWORDS

Markov chains,Hidden Markov Models, computational linguistics, pattern recognition, statistical

1. INTRODUCTION

Markov chains theory is increasingly being adopted in real-world computing applications since it provides a convenient way for modeling temporal, time-series data. At each clock tick, the system moves into a new state that can be the same as the previous one. A Markov chain model is amathematical tool that capture the patterns dependencies in pattern recognition systems. For this reason, Markov chain theory is appropriate in natural langue processing (NLP) where it is naturally characterized by dependencies between patterns such as characters or words.

Markov chains are directed graphs (a graphical model) that are generally used with relatively long data sequences for data-mining tasks. Such tasks include prediction, classification, clustering, pattern discovery, software testing, multimedia analysis, networks, etc. Reference [1] indicated that there are two reasons of Markov chains popularity; very rich in mathematical structure and work well in practice for several important applications. Hidden Markov models (HMM) is an extension of Markov chains that used to find the hidden system's states based on the observations.

In order to facilitate the research in this direction, this paper provides a survey of this so popular data modeling technique. However, because of the wide range of the research domains that use this technique. We specifically focuson the linguistics related applications. Reference [2] list some domains that utilize Markov chains theory which include: physics, chemistry, testing, speech recognition, information sciences, queueing theory, internet applications, statistics, economics and finance, social sciences, mathematical biology, genetics, games, music, baseball,

David C. Wyld et al. (Eds) : ICAITA, CDKP, CMC, SOFT, SAI - 2016
pp. 53– 62, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.61305

Markov text generators, bioinformatics. Reference [3] lists the five greatest applications of Markov chains that include Scherr's application to computer performance evaluation, Brin and Page's application to PageRank and Web Search, Baum's application to HMM, Shannon's application to information theory, and Markov's application to Eugeny Onegin.

This paper is organized as follows. The next section presents a background of Markov chains theory. Section 3 highlights the main concepts of HMM followed by a literature review of Markov chains and HMM in section 4. Finally, we conclude in section 5.

2. MARKOV CHAINS

Markov chains are quite useful in modeling computational linguistics. A Markov chain is a memorylessstochastic model that describes the behaviour of an integer-valued random process. The behaviour is the simple form of dependency in which the next state (or event) depends only on the current state. According to [4], a random process is said to be Markov if the future of the process, given the present, is independent of the past. To describe the transitions between states, a transition diagram is used to describe the model and the probabilities of going from one state to another. For example, Figure 1 shows a Markov chain diagram with three states (Easy, Ok, and Hard) that belong to exam cases(i.e. states). In the figure, each arc represents the probability value for transition from one state to another.

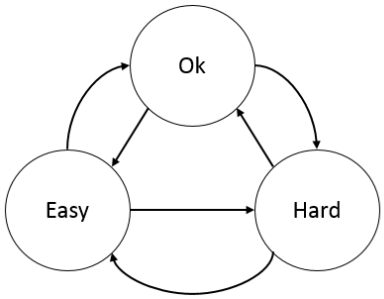


Figure 1. A Simple Markov chain with three states

The Markov chain diagrams are generally represented using state transition matricesthat denote the transition probabilities from one state to another. Hence, a state transition matrix is created using the entire states in the system. For example, if a particular textual application has a training datathat contains N states (e.g. the size of lexicon), then the state transition matrix is described by a matrix A= {aij} of size N*N. In matrix A, the element aij denote the transition probability from a state i to a state j. Table 1 shows how the state transition matrix used to characterize the Markov diagram shown in Figure 1. That is, the matrix carries the state transitions probabilities between the involved states(Easy, Ok, and Hard). For illustration, the P(E|H) denote to the probability of the next exam to be Easy given that the previous exam was Hard.

Table 1. A state transition matrix of three states

State		**Next Exam**		
		Easy (E)	**Ok (O)**	**Hard (H)**
Previous Exam	**Easy (E)**	P(E|E)	P(O|E)	P(H|E)
	Ok (O)	P(E|O)	P(O|O)	P(H|O)
	Hard (H)	P(E|H)	P(O|H)	P(H|H)

In Table 1, the sum of the probability values at each row is 1 as the the sum of the probabilities coming out of each node should be 1. Hence,P(E|E)+P(O|E)+P(P(H|E) equal 1. Markov chain is a worthy topic that has many details. For examples, it contains discrete-time, continuous-time,

time-reversed, reversible, and irreducible Markov chains. The case shown in Figure 1 is irreducible case, also called ergodic, where it is possible to go from every state to every state.

To illustrate a simple Markov chain data model, a small data set contains two English sentences used to create a transition matrix based on the neighbouringcharacters sequences. The sentences are inspirational English quotes picked from [5]:

(1) Power perceived is power achieved. (2) If you come to a fork in the road, take it.

Figure 2 shows the transition matrix of these quotes by counting the total number of occurrences of the adjacent two character sequences. It is a 19 × 19 matrix where the value 19 is the total number of unique characters appeared in thesentences (i.e the two quotes). In this example, creating transition matrix is case insensitive where D is same as d, as an example. In addition, a space between two words discarded and not considered in the transition matrix. Figure 2 also shows that the maximum number in the matrix's entries is 3 (a highlighted underlined value) which means that moving from character e to r (e➔r) is the most frequently sequence appeared in this small corpus. The words that contains this sequence are :{ Power (two times) andperceived}.

	a	c	d	e	f	h	i	k	m	n	o	p	r	s	t	u	v	w	y
a	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
c	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
e	0	0	2	0	0	0	1	0	0	0	0	0	**3**	0	0	0	1	0	0
f	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
h	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
i	0	0	0	1	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0
k	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
m	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
o	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	2	0
p	0	0	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
r	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
t	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
u	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
v	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
w	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0

Figure 2. A transition matrix of two characters sequences

Based on the information provided in the transition matrix shown in Figure 2. It is possible to answer some questions related to the given data collection. Among inquires, what is the total number of the two characterssequences appeared in the given data set?What are the two characters sequences that did not appear in the data collection?What is the least frequently two characters sequences in the data set? Accordingly, Markov chains are used as prediction systems such as weather forecasting. Therefore, it is possible to predict the tomorrow's weather according to the today's weather. For example, if we have two states (Sunny, Rainy), and the requirement is to find the probability P(Sunny|Rainy), Markov chains make it possible based on the information provided in the probability transition matrix. Another example of the using Markov chains is banking industry. A big portfolio of banks is based on loans. Therefore, Markov chainsare used to classify loans to different states such as Good, Risky, and Bad loans.

For simplicity, the information presented in Figure 2 shows the transition matrixbased on total number of occurrences. Figure 3 shows the same information but using probabilities instead of

the number of occurrences. That is, it contains the probability of moving from one character to another. As previously indicated, the sum of entries at each row is equal 1. In Figure 3, any matrix entry that has 0 means that there is no transition at that case. Similarly, if the matrix entry is 1, it means that there is only one possible output of that state. For example, the character “o” comes after “y”, and this is the only possible arc of the state “y”.

	a	c	d	e	f	h	i	k	m	n	o	p	r	s	t	u	v	w	y
a	0	0.33	0.33	0	0	0	0	0.33	0	0	0	0	0	0	0	0	0	0	0
c	0	0	0	0.33	0	0.33	0	0	0	0	0.33	0	0	0	0	0	0	0	0
d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
e	0	0	0.29	0	0	0	0.14	0	0	0	0	0	0.43	0	0	0	0.14	0	0
f	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
h	0	0	0	0.5	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0
i	0	0	0	0.17	0.17	0	0	0	0	0.17	0	0	0	0.17	0.17	0	0.17	0	0
k	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
m	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
o	0.17	0	0	0	0	0	0	0	0.17	0	0	0	0.17	0	0	0.17	0	0.33	0
p	0	0	0	0.33	0	0	0	0	0	0	0.67	0	0	0	0	0	0	0	0
r	0	0.33	0	0	0	0	0	0.33	0	0	0.33	0	0	0	0	0	0	0	0
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
t	0.33	0	0	0	0	0.33	0	0	0	0	0.33	0	0	0	0	0	0	0	0
u	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
v	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
w	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0

Figure 3. A probability transition matrix of two characters sequences

3. HIDDEN MARKOV MODELS

Hidden Markov models (HMM) is an extension to Markov chains models as both used fortemporal data modeling. However, the difference is that the states in Markov chain models are directly observed while they are hidden in the case of HMM.We explain the concept of HMM based on Figure 1 that shows athreeexam’s states Markov diagram. As a very simple example, supposed that a student’s parents want to know the levels (i.e the difficulty) of theirson’s exams, naturally, it is possible to recognize the exam as Easy or Ok if the son feels Fine. Similarly, it is possible to recognize the exam as Hard if the son looks Scared. From the parents’ point of view, the required states (i.e. Easy, Ok, or Hard) are hidden. However, they directly observe the student’s reaction or feeling. Hence, the parents might use the observed reactionas an indication to know the hidden states. HMM is described using three matrices: the initial probability matrix, the observation probability matrix, and the state transition matrix. Figure 4 shows a HMM diagram that shows the states and the observations. In the figure, each arc represents the probability between the states and between the states and the observations.

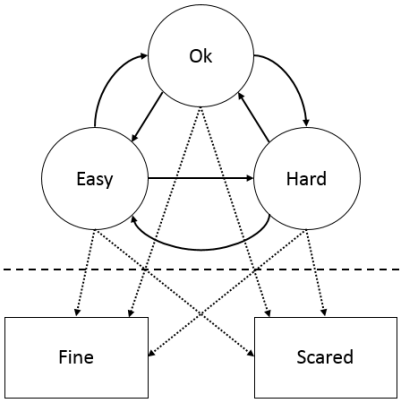


Figure 4. A HMM diagram with the transition and the observation arcs

Based on the information provided in the matrices, either Baum-Welch (also called any path) or Viterbi (also called best path) algorithms used to find the probability scores during recognition phase. Figure 5 shows the trellis diagram forexam states HMM. While Baum-Welch algorithm is used to compute the recognitin probability of a sequence, Viterbi is used to find the best-state sequence associated with the given observtatin, this procoss is also known as back-trakcing. Hence, after computing the observations sequence probability and finding the maximumprobability (supposed the star in Figure 5), the Viterbi algorithm leads the process back to identify the states (sources) from which the observations sequence have been emitted. In Figure 5, the maximum probalities supposed to be achieved at the states shown using the dotted lines: Ok, Easy, Hard, respectively.

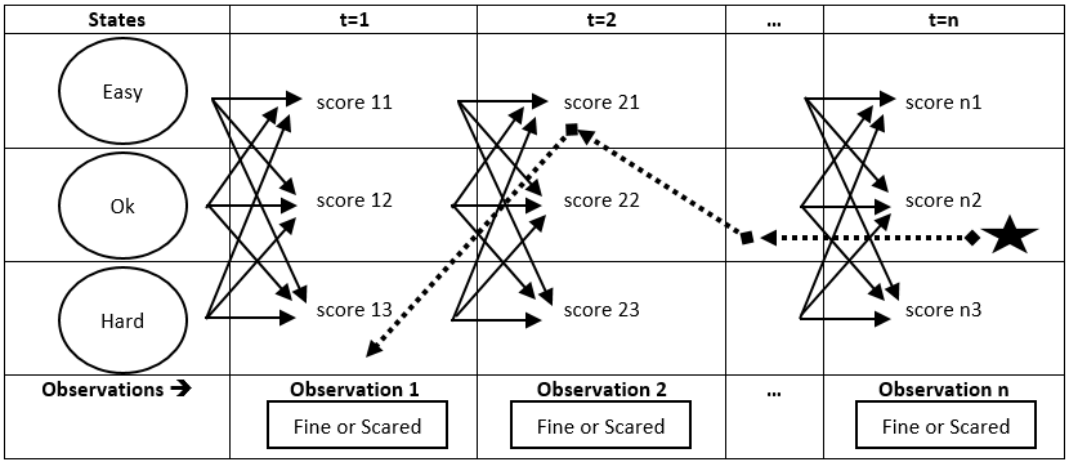


Figure 5.Trellis diagram of three states HMM

4. LINGUISTIC APPLICATIONS

In the literature, there are quite many works on modelingcontent dependencies for linguistics applications. Markov chain models and HMMs are of great interest to linguistic scholar who primarily work on data sequences. Even though this study focuses on linguistic applications, however, Markov chains used to model a variety of phenomena in different fields. The following are some of studies employed Markov chains. We intentionally ignored the references as the literature hastoo many studies employed Markov chains:

orecasting,
iformatics,
er quality,
nagement,
, business

The following two subsections include some of the linguistic studies that utilized Markov chain theory. Linguistic applications topics mainly include (but not limited) speech recognition,speech emotion recognition,part-of-speech tagging, machine translation, text classification, text summarization, optical character recognition (OCR), named entity recognition, question answering,authorship attribution, etc.For the reader who interested in NLP, Reference [6] is a good reference as it demonstrates a thorough study of NLP (Almost) from Scratch.

4.1. Markov chains based research

The literature has a large number of studies that employ Markov chains for NLP applications. The following are some linguistic related applications. Reference [7] proposed a word-dividing algorithm based on statistical language models and Markov chain theory for Chinese speech processing. Reference [8] presented a semantic indexing Markov chains algorithm that uses both audio and visual information for event detection in soccer programs. Reference [9] investigated the use of Markov Chains and sequence kernels for the task of authorship attribution. Reference [10] implemented a probabilistic framework for support vector machine (SVM) that allows for automatic tuning of the penalty coefficient parameters and the kernel parameters via Markov chain for web searching via text categorization. Reference [11] demonstrated an automatic video annotation using multimodal Dirichlet process mixture model by collecting samples from the corresponding Markov chain. Reference [12] used a linguistic steganography detection method based on Markov chain models. Reference [13] showed how probabilistic Markov chain models can be used to detect topical structure in large text corpora.

Reference [14] proposed a method of recognizing location names from Chinese texts based on Max-Margin Markov Network. Reference [15] utilized Markov chain and statistical language models in a linguistic steganography detection algorithm. Reference [16] proposed a Markov chain based algorithm for Chinese word segmentation. Reference [17] presented two new textual feature selection methods based on Markov chains rank aggregation techniques. Reference [18] proposed a Markov chain model for radical descriptors in Arabic Text Mining. Reference [19] presented statistical Markov chain models for the distributions of words in text lines. Reference [20] proposed a method for handwritten Chinese/Japanese text (character string) recognition based on semi-Markov conditional random fields (semi-CRFs). Reference [21] presented a Markov chain method to find authorship attribution on relational data between function words. Reference [22] utilized a probabilistic Markov chain model to infer the location of Twitter users. Reference [23] proposed a Markov chain based technique to determine the number of clusters of a corpus of short-text documents.Reference [24] proposed a Markov chain based method for digital document authentication. Reference [25] used Markov chain for authorship attribution in Arabic poetry.

4.2. Hidden Markov modelsbased research

Linguistic HMM based research has been for long an active research area due to the rapid development in NLP applications. The literature has many studies as follows. Reference [26] proposed to extract acronyms and their meaning from unstructured text as a stochastic process using HMM. Reference [27] proposed a morphological segmentation method with HMM method for Mongolian.Reference [28] employed HMM for Arabic handwritten word recognition based on HMM. Reference [29] presented a scheme for off-line recognition of large-set handwritten characters in the framework of the first-order HMMs. Reference [30] proposed the use of hybrid HMM/Artificial Neural Network (ANN) models for recognizing unconstrained offline handwritten texts. Reference [31] used HMMs for recognizing Farsi handwritten words.

Reference [32] describes recent advances in HMM based OCR for machine-printed Arabic documents.Reference [33] proposed a HMMbased method fornamed entity recognition. Reference [34] combined text classification and HMM techniques for structuring randomized clinical trial abstracts. Reference [35] employed HMM for medical text classification. Reference [36] propose text (sequences of pages) categorization architecture based on HMM.Reference [37] described a model for machine translation based on first-order HMM.Reference [38] introduced speech emotion recognition by use of HMM.Reference [39] presented a HMMbased method for speech emotion recognition. Reference [40] discussed the role of HMM in speech recognition. Reference [41] indicated that almost all present day large vocabulary continuous speech

recognition (LVCSR) systems based on HMMs.Reference [42] presented a text summarization method based on HMM. Reference [43] presented a method for summarizing speech documents using HMM. Reference [44] used HMM for part-of-speech tagging task. Reference [45] presented a second-order approximation of HMM for part-of-speech tagging task.

5. CONCLUSIONS

This work demonstrates the potential and the size of Markov chains research. The study reveals that the Markov chain and HMM is of high important for linguistic applications. Similarly, Markov chains are also widely used in many other applications. For future work, it worthy to explore the power of Markov chain in new linguistic and scientific directions with more details.

ACKNOWLEDGEMENTS

This work is supported by Kuwait Foundation of Advancement of Science (KFAS), Research Grant Number P11418EO01 and Kuwait University Research Administration Research Project Number EO06/12.

REFERENCES

[1] Rabiner, Lawrence R. "A tutorial on hidden Markov models and selected applications in speech recognition." Proceedings of the IEEE 77.2 (1989): 257-286.

[2] Markov_chain. (2016, August). Retrieved from https://en.wikipedia.org/wiki/Markov_chain

[3] Von Hilgers, Philipp, and Amy N. Langville. "The five greatest applications of Markov Chains." Proceedings of the Markov Anniversary Meeting, Boston Press, Boston, MA. 2006.

[4] Leon-Garcia, Alberto, and Alberto. Leon-Garcia. Probability, statistics, and random processes for electrical engineering. Upper Saddle River, NJ: Pearson/Prentice Hall, 2008.

[5] California Indian Education. (2016, August). Retrieved from http://www.californiaindianeducation.org/inspire/world/

[6] Collobert, Ronan, et al. "Natural language processing (almost) from scratch."Journal of Machine Learning Research 12.Aug (2011): 2493-2537.

[7] Bin, Tian, et al. "A Chinese word dividing algorithm based on statistical language models." Signal Processing, 1996., 3rd International Conference on. Vol. 1. IEEE, 1996.

[8] Leonardi, Riccardo, PierangeloMigliorati, and Maria Prandini. "Semantic indexing of soccer audio-visual sequences: a multimodal approach based on controlled Markov chains." IEEE Transactions on Circuits and Systems for Video Technology 14.5 (2004): 634-643.

[9] Sanderson, Conrad, and Simon Guenter. "On authorship attribution via Markov chains and sequence kernels." 18th International Conference on Pattern Recognition (ICPR'06). Vol. 3. IEEE, 2006.

[10] Lim, Bresley Pin Cheong, et al. "Web search with text categorization using probabilistic framework of SVM." 2006 IEEE International Conference on Systems, Man and Cybernetics. Vol. 4. IEEE, 2006.

[11] Velivelli, Atulya, and Thomas S. Huang. "Automatic video annotation using multimodal Dirichlet process mixture model." Networking, Sensing and Control, 2008. ICNSC 2008. IEEE International Conference on. IEEE, 2008.

[12] Chen, Zhi-li, et al. "Effective linguistic steganography detection." Computer and Information Technology Workshops, 2008. CIT Workshops 2008. IEEE 8th International Conference on. IEEE, 2008.

[13] Dowman, Mike, et al. "A probabilistic model of meetings that combines words and discourse features." IEEE Transactions on Audio, Speech, and Language Processing 16.7 (2008): 1238-1248.

[14] Li, Lishuang, Zhuoye Ding, and Degen Huang. "Recognizing location names from Chinese texts based on max-margin markov network." Natural Language Processing and Knowledge Engineering, 2008. NLP-KE'08. International Conference on. IEEE, 2008.

[15] Meng, Peng, et al. "Linguistic steganography detection algorithm using statistical language model." Information Technology and Computer Science, 2009. ITCS 2009. International Conference on. Vol. 2. IEEE, 2009.

[16] Baomao, Pang, and Shi Haoshan. "Research on improved algorithm for Chinese word segmentation based on Markov chain." Information Assurance and Security, 2009. IAS'09. Fifth International Conference on. Vol. 1. IEEE, 2009.

[17] Wu, Ou, et al. "Rank aggregation based text feature selection." Web Intelligence and Intelligent Agent Technologies, 2009. WI-IAT'09. IEEE/WIC/ACM International Joint Conferences on. Vol. 1. IET, 2009.

[18] El Hassani, Ibtissam, AbdelazizKriouile, and Youssef BenGhabrit. "Measure of fuzzy presence of descriptors on Arabic Text Mining." 2012 Colloquium in Information Science and Technology. IEEE, 2012.

[19] Haji, Mehdi, et al. "Statistical Hypothesis Testing for Handwritten Word Segmentation Algorithms." Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on. IEEE, 2012.

[20] Zhou, Xiang-Dong, et al. "Handwritten Chinese/Japanese text recognition using semi-Markov conditional random fields." IEEE transactions on pattern analysis and machine intelligence 35.10 (2013): 2413-2426.

[21] Segarra, Santiago, Mark Eisen, and Alejandro Ribeiro. "Authorship attribution using function words adjacency networks." 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013.

[22] Rodrigues, Erica, et al. "Uncovering the location of Twitter users." Intelligent Systems (BRACIS), 2013 Brazilian Conference on. IEEE, 2013.

[23] Goyal, Anil, Mukesh K. Jadon, and Arun K. Pujari. "Spectral approach to find number of clusters of short-text documents." Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2013 Fourth National Conference on. IEEE, 2013.

[24] Shen, Jau Ji, and Ken Tzu Liu. "A Novel Approach by Applying Image Authentication Technique on a Digital Document." Computer, Consumer and Control (IS3C), 2014 International Symposium on. IEEE, 2014.

[25] Ahmed, Al-Falahi, et al. "Authorship attribution in Arabic poetry." 2015 10th International Conference on Intelligent Systems: Theories and Applications (SITA). IEEE, 2015.

[26] Osiek, Bruno Adam, Geraldo Xexéo, and Luis Alfredo Vidal de Carvalho. "A language-independent acronym extraction from biomedical texts with hidden Markov models." IEEE Transactions on Biomedical Engineering 57.11 (2010): 2677-2688.

[27] He, Miantao, Miao Li, and Lei Chen. "Mongolian Morphological Segmentation with Hidden Markov Model." Asian Language Processing (IALP), 2012 International Conference on. IEEE, 2012.

[28] Alma'adeed, Somaya, Colin Higgens, and Dave Elliman. "Recognition of off-line handwritten Arabic words using hidden Markov model approach." Pattern Recognition, 2002. Proceedings. 16th International Conference on. Vol. 3. IEEE, 2002.

[29] Park, Hee-Seon, and Seong-Whan Lee. "Off-line recognition of large-set handwritten characters with multiple hidden Markov models." Pattern Recognition 29.2 (1996): 231-244.

[30] Espana-Boquera, Salvador, et al. "Improving offline handwritten text recognition with hybrid HMM/ANN models." IEEE transactions on pattern analysis and machine intelligence 33.4 (2011): 767-779.

[31] Imani, Zahra, et al. "offline Handwritten Farsi cursive text recognition using Hidden Markov Models." Machine Vision and Image Processing (MVIP), 2013 8th Iranian Conference on. IEEE, 2013.

[32] Prasad, Rohit, et al. "Improvements in hidden Markov model based Arabic OCR." Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. IEEE, 2008.

[33] Zhou, GuoDong, and Jian Su. "Named entity recognition using an HMM-based chunk tagger." proceedings of the 40th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2002.

[34] Xu, Rong, et al. "Combining Text Classification and Hidden Markov Modeling Techniques for Structuring Randomized Clinical Trial Abstracts." AMIA. 2006.

[35] Yi, Kwan, and JamshidBeheshti. "A hidden Markov model-based text classification of medical documents." Journal of Information Science (2008).

[36] Frasconi, Paolo, Giovanni Soda, and Alessandro Vullo. "Hidden markov models for text categorization in multi-page documents." Journal of Intelligent Information Systems 18.2-3 (2002): 195-217.

[37] Vogel, Stephan, Hermann Ney, and Christoph Tillmann. "HMM-based word alignment in statistical translation." Proceedings of the 16th conference on Computational linguistics-Volume 2. Association for Computational Linguistics, 1996.

[38] Schuller, Björn, Gerhard Rigoll, and Manfred Lang. "Hidden Markov model-based speech emotion recognition." Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on. Vol. 2. IEEE, 2003.

[39] Nwe, Tin Lay, Say Wei Foo, and Liyanage C. De Silva. "Speech emotion recognition using hidden Markov models." Speech communication 41.4 (2003): 603-623.

[40] Juang, Biing Hwang, and Laurence R. Rabiner. "Hidden Markov models for speech recognition." Technometrics 33.3 (1991): 251-272.

[41] Gales, Mark, and Steve Young. "The application of hidden Markov models in speech recognition." Foundations and trends in signal processing 1.3 (2008): 195-304.

[42] Conroy, John M., and Dianne P. O'leary. "Text summarization via hidden markov models." Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2001.

[43] Maskey, Sameer, and Julia Hirschberg. "Summarizing speech without text using hidden markov models." Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers. Association for Computational Linguistics, 2006.

[44] Kupiec, Julian. "Robust part-of-speech tagging using a hidden Markov model." Computer Speech & Language 6.3 (1992): 225-242.

[45] Thede, Scott M., and Mary P. Harper. "A second-order hidden Markov model for part-of-speech tagging." Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics. Association for Computational Linguistics, 1999.

CONTROL OF AN INDUCTION MOTOR WITH DOUBLE ANN MODEL BASED DTC

Fatih Korkmaz

Department of Electric-Electronic Engineering, Çankırı Karatekin University,
Uluyazı Kampüsü, Çankırı, Turkey
fkorkmaz@karatekin.edu.tr

ABSTRACT

Direct torque control (DTC) is preferably control method on high performance control of induction motors due to its advantages such as fast dynamic response, simple and robust control structure. However, high torque and current ripples are mostly faced problems in this control method. This paper presents artificial neural network (ANN) based approach to the DTC method to overcome mentioned problems. In the study, by taking a different perspective to ANN and DTC integration, two different ANN models have been designed, trained and implemented. The first ANN model has been used for switch selecting process and the second one has been used for sector determine process. Matlab/Simulink model of the proposed ANN based DTC method has created in order to compare with the conventional DTC and the proposed DTC methods. The simulation studies have proved that the induction motor torque and current ripples have been reduced remarkably with the proposed method and this approach can be a good alternative to the conventional DTC method for induction motor control.

KEYWORDS

Direct torque control, Induction motor control, Artificial neural networks, Vector control

1. INTRODUCTION

In the early 1970s, vector control method was firstly proposed by Blaschke and it was named as field oriented control (FOC). About ten years later, another vector based control method was presented by Takahashi and it was called as direct torque control (DTC). Control of motors with parameters which converted to veal identities, can be acceptable prominent features of the both control methods. [1-2].

Due to rapid development in the field of power switches and microcontrollers technologies , the both methods have shown great development, since they were first presented. When compared with the FOC, the DTC method has some structural advantages such as simple control algorithm, robust controller character due to independency of motor parameters, needs only stator phase resistance, and fast dynamic response. However, the DTC method has some handicaps that can be listed as, high torque and current ripples, variable switching frequency behavior, and implementation limitations owing to necessity of low sampling time [3].

David C. Wyld et al. (Eds) : ICAITA, CDKP, CMC, SOFT, SAI - 2016
pp. 63– 71, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.61306

Over the last decades, different types of the DTC algorithms have been proposed by researchers and academics to overcome mentioned handicaps and most of them have focused on torque ripples side in their studies. In [4], researchers proposed matrix converter based model predictive control approach and they proved that it can be effective in minimizing torque ripples. In [5], authors present a modified DTC algorithm for permanent magnet synchronous motor drives with fast torque dynamics and constant switching frequency. The authors presents the DTC method using fuzzy controller to minimize torque ripples for BLDC in [6] and for induction motors in [7]. Artificial neural network model based DTC systems have also been investigated to reduce torque ripples in [8].

This paper presents a different approach to ANN based DTC model to reduce torque ripples on DTC controlled induction motor drives. Two different ANN models have been designed, trained and performed in the same model. The first ANN model has been used for switch selecting process and the second one has been used for sector determine process. Matlab/Simulink model of the proposed ANN based DTC method was created in order to compare with the conventional DTC and the proposed DTC methods. Basics of the DTC method was explained in Section 2 and proposed ANN based DTC method was documented in detail in Section 3. Simulation results were presented and commented in Section 4.

2. DIRECT TORQUE CONTROL

Mathematical model of the DTC method contains simpler and fewer equations when compared the FOC. Unlike the FOC, the DTC model needs only one motor parameter in the modelling of the control system. Thus, the DTC is considered to be more simple and reliable [9].

The both control techniques methodology base on mathematical transformation that transforms three phase motor parameters to two phase components. The mathematic model of the DTC system bases on Clarke transformation. Thus, some measured three phase parameters of the motor are converted to two phase components, such as voltages and currents. Instant values of the stator flux and produced motor torque can be calculated with these components as defined with following equations. The Clarke transformation matrix is given as:

$$\begin{bmatrix} f_\alpha \\ f_\beta \end{bmatrix} = \frac{2}{3}\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}\begin{bmatrix} f_a \\ f_b \\ f_c \end{bmatrix} \quad (1)$$

Where, f_α, f_β are $\alpha - \beta$ components of motor parameters, and f_a, f_b, f_c are the abc frame components[10-11].

$\alpha - \beta$ components of the stator voltages and flux are expressed in following equations.

$$v_{s\alpha} = R_s i_{s\alpha} + L_s \frac{di_{s\alpha}}{dt} \quad (2)$$

$$v_{s\beta} = R_s i_{s\beta} + L_s \frac{di_{s\beta}}{dt} \quad (3)$$

$$\lambda_{s\alpha} = \int (v_{s\alpha} - R_s i_{s\alpha}) dt \quad (4)$$

$$\lambda_{s\beta} = \int \left(v_{s\beta} - R_s i_{s\beta}\right) dt \tag{5}$$

The magnitude of the flux can be calculated with:

$$\lambda = \sqrt{\lambda_{s\alpha}^2 + \lambda_{s\beta}^2} \tag{6}$$

and position of the stator flux vector is calculated with:

$$\theta = arctan \frac{\lambda_{s\beta}}{\lambda_{s\alpha}} \tag{7}$$

And finally produced motor torque equation can be written as:

$$T_e = \frac{3}{2} p\left(\lambda_\alpha i_\beta - \lambda_\beta i_\alpha\right) \tag{8}$$

In the DTC, motor torque and flux can be controlled directly with control of stator flux vector. In this control, rotation speed of the flux vector controls produced torque while the length controls value. The DTC control algorithm is bases on selecting the appropriate inverter switching state to directly control stator flux vector's speed and length that explain why the method is named as DTC. In order to keep the torque and flux errors within the predefined limits, hysteresis comparators are used [12].

In this control process, six active voltage vectors are used to keep stator flux vector in predefined hysteresis band limit. The flux vector rotation is divided to six different sectors. Figure 1 represents six sector and active voltage vectors.

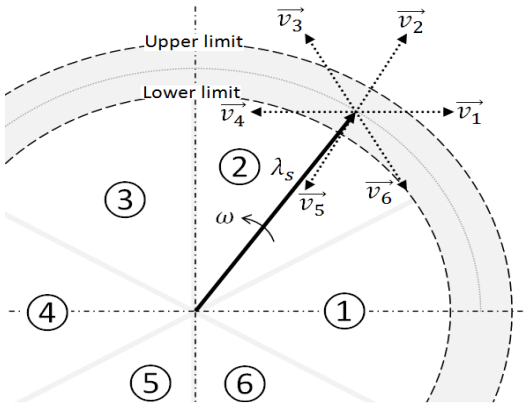


Figure 1. Six sector and active voltage vectors

3. ANN MODEL BASED DTC SYSTEM

ANN models of the control systems are inspired by biological nervous systems and interconnected points, that called neurons, are designed to solve complex problems. In substance, this structure bases on actually working principles of human brain to solve problems.

In recent decades, ANN models have rising popularity in different kind of control systems due to their learning abilities, robust structures, beside their modeling success on non-linear and complex mathematical models.

In the proposed ANN based DTC model, two different ANN models have designed to solve complex processes in conventional DTC approach; selecting appropriate inverter switching states and determining stator flux vector region.

$\alpha - \beta$ components of stator flux were used in the ANN model to determining stator flux vector region process. The ANN model structure is shown in Figure 2

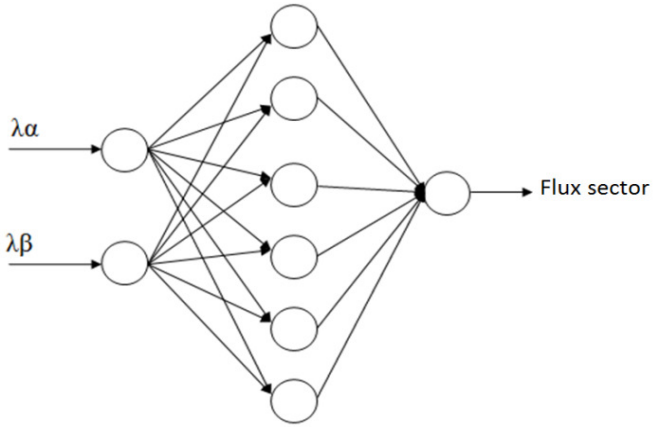


Figure 2. Structure of ANN model for determining stator flux sector

As shown in Figure 2, flux sector ANN model has two neurons in input layer, six neurons in hidden layer and one neuron in output layer. The flux sector ANN was trained about 100000 input data which were obtained with conventional DTC model. 70000 input data were used to train the network while 15000 input data were used for validation and 15000 input data were used for test processes after training. Simulink model of the flux sector ANN given in Figure 3.

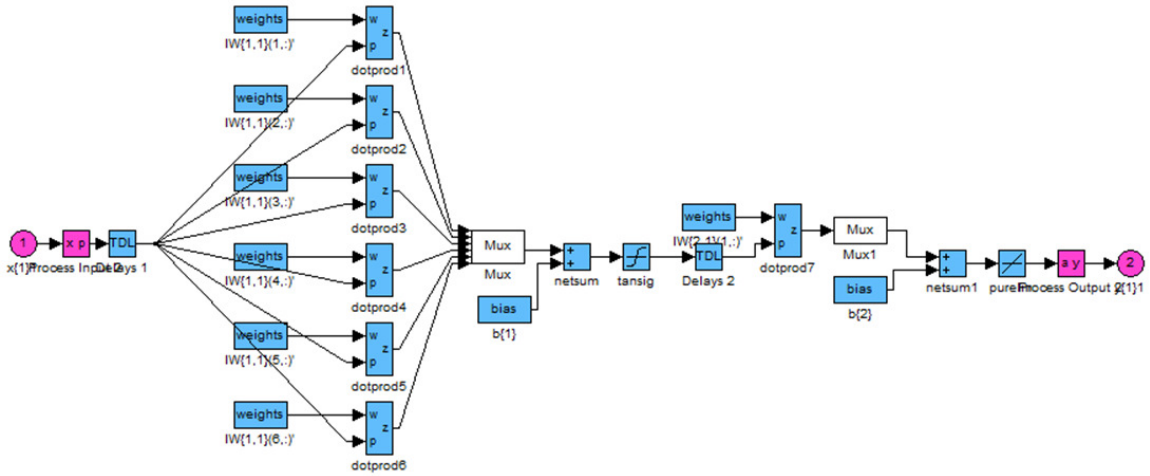


Figure 3. Simulink model of the flux sector ANN unit

Outputs of the flux hysteresis, torque hysteresis and flux region data were used as inputs of the second ANN model which was created to determine optimum inverter switching states.

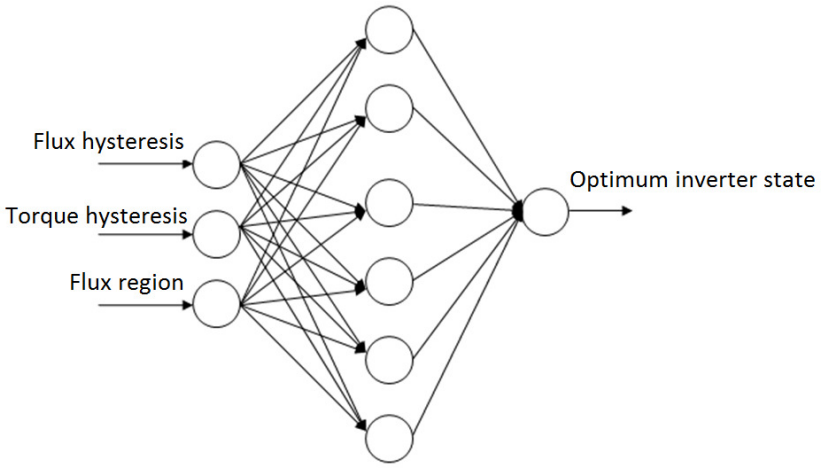


Figure 4. Structure of ANN model for determining optimum inverter switching states

As shown in Figure 4, inverter state ANN model has three neurons in input layer, six neurons in hidden layer and one neuron in output layer. The flux sector ANN was trained about 50000 input data which were obtained with conventional DTC model. 35000 input data were used to train the network while 7500 input data were used for validation and 7500 input data were used for test processes after training. Overall Simulink block diagram of the proposed ANN based DTC system is shown in Figure 5.

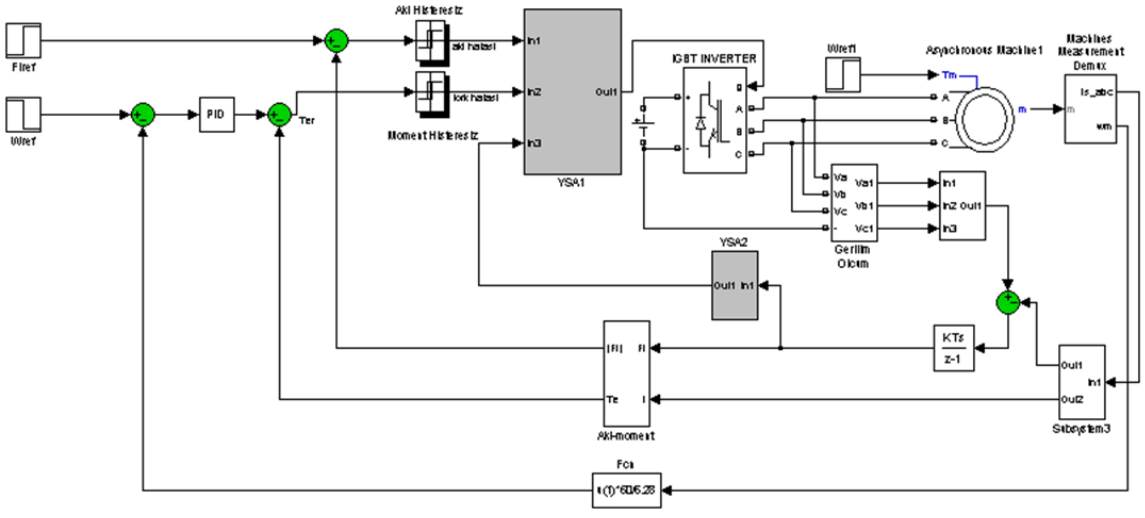


Figure 5. Overall Simulink block diagram of the proposed ANN based DTC system

4. SIMULATIONS

This section presents comparatively simulation results between conventional DTC (C-DTC) and proposed ANN based DTC (ANN-DTC) models. Two different scenarios were applied and investigated to obtain fare comparison between both methods. Induction motor and simulation parameters were presented in Table 1. Total simulation time was 5 sec. and motor load was 0 Nm at the first 3 sec. then switched to 3 Nm reference torque after 3. sec

Table 1. The induction motor and simulation model parameters

Motor Power (kW)	4
Bus Voltage (V)	300
Phase resistance(ohm)	1,54
Flux Reference (Wb)	0,5
Flux Hysteresis Band	±0,05
Torque Hysteresis Band	±0,5
Sample time (µs)	10

In the first scenario, motor was tested at low speed condition with no-load and loaded conditions. The motor speed and torque responses are given in Figure 6 and Figure 7, respectively.

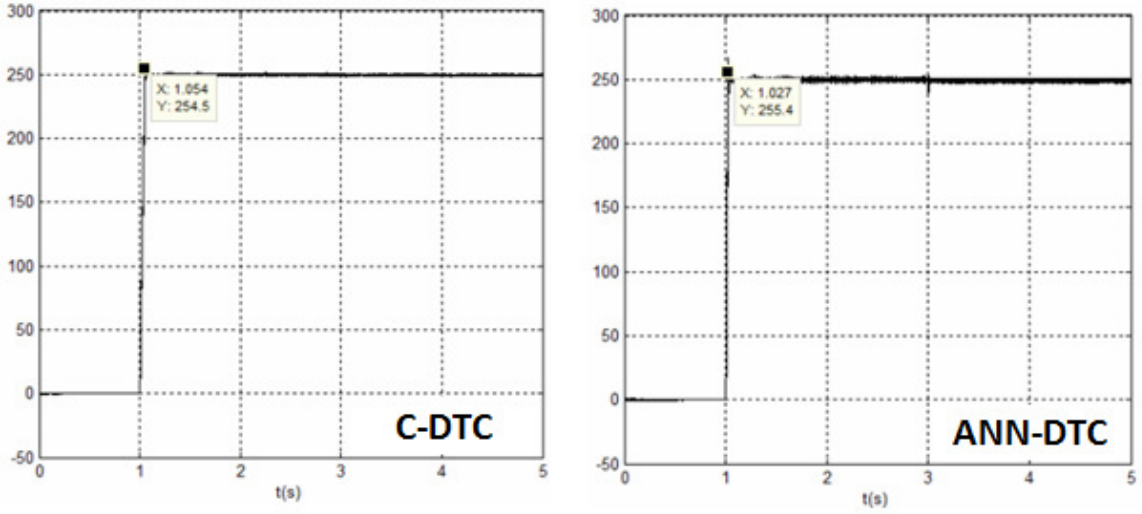


Figure 6. Motor speed curves at 250 rpm reference

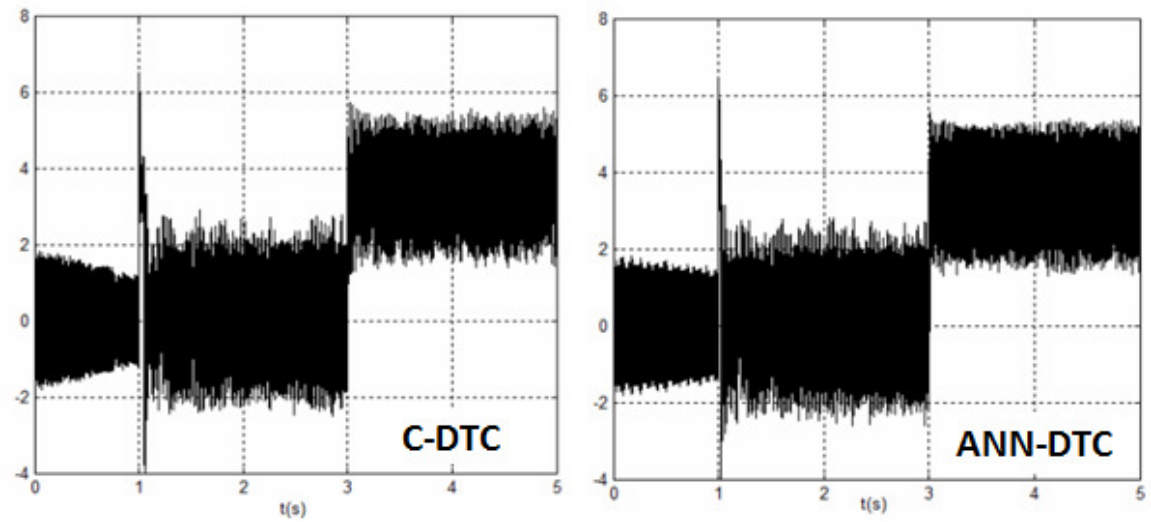


Figure 7. Motor torque curves at 250 rpm reference

As shown in Figure 6 and Figure 7, motor speed and torque performances are almost same in low speed working conditions. In both control approach, torque ripples are about ±2 Nm for unloaded period while torque ripples are about +5/1,7 Nm for loaded conditions. Briefly, it can be said that there are no meaning difference between the both methods for low speed conditions.

In the second scenario, motor was tested at rated speed (2800rpm) with no-load and loaded (3Nm) conditions. The motor speed and torque responses are given in Figure 8 and Figure 9, respectively.

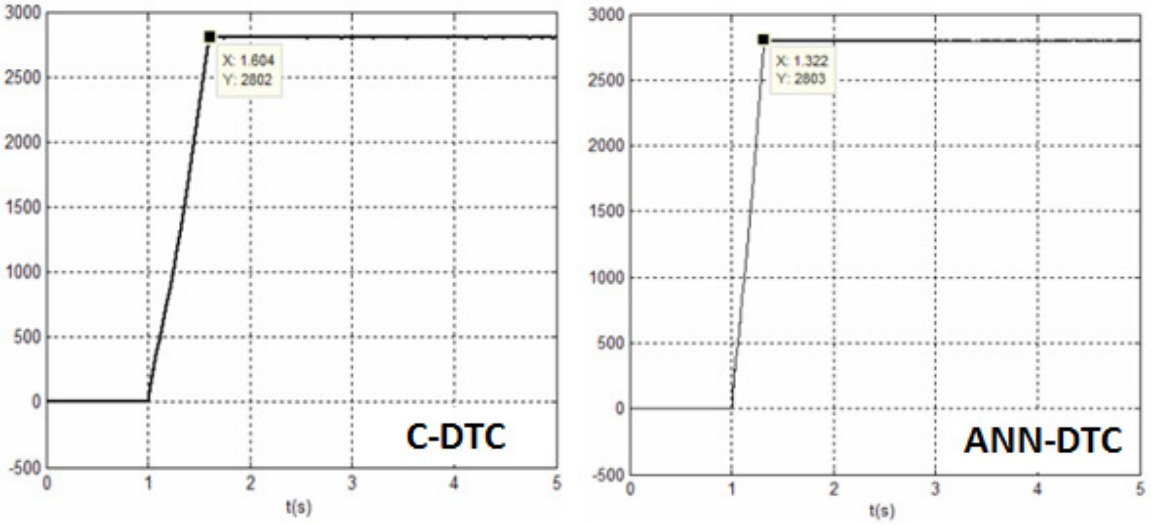


Figure 8. Motor speed curves at 2800 rpm reference

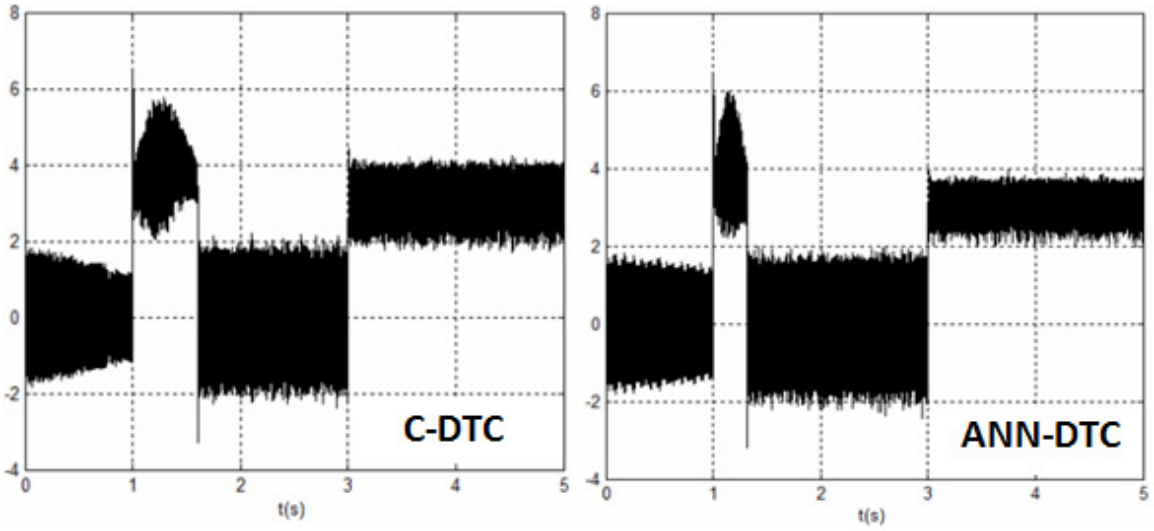


Figure 9. Motor torque curves at 2800 rpm reference

The main difference for the proposed ANN-DTC method emerged in the secondary group of tests as can be seen clearly in Fig 8 and Fig 9. The speed response of the motor proved that ANN-DTC method had faster dynamic responses when compared C-DTC method. Time to reach the reference speed was about %50 shorter in the proposed method. On the other hand, when the torque curves examined closely, it can be seen that torque ripples of the motor were reduced remarkably with the ANN-DTC method.

5. CONCLUSIONS

This paper presents a new artificial neural networks based approach to conventional direct torque control method for high performance control of induction motors. The aim of the paper is simplifying the complex mathematical process in conventional direct torque control method by usage of neural network models. For this aim, two independent neural models have been designed, trained and tested. The first neural model was designed for sectors determine process and the second one were designed for optimum switching select unit. Numerical simulations have been performed to investigate the effects of the proposed method on the overall system performances. Low speed and rated speed performances of the motor were simulated at zero loads and rated loads working conditions. Simulation results have been comparatively presented and they have been showed that proposed neural network based DTC method has better dynamic performance especially closer to rated speed values. The proposed method has faster acceleration time in transient state region and torque ripples were reduced with the proposed method in steady-state region. The practical works on the proposed method as a real time application are still going on and experimental results will be presented by researchers in the future works.

REFERENCES

[1] Blaschke, F.,(1972) “The Principle of Field Orientation Applied to The New Transvector Closed-Loop Control System for Rotating Field Machines”, Siemens-Rev., Vol. 39, 217–220.

[2] Takahashi Takahashi, I. & Noguchi. T. (1986) “A new quick-response and high efficiency control strategy of an induction motor,” IEEE Transactions on Industrial Applications, vol.I A-22 ,No.5. , pp. 820–827.

[3] Fatih Korkmaz. & M. Faruk Çakır. & Yılmaz Korkmaz. & Ismail Topaloglu, (2012) “Fuzzy Based Stator Flux Optimizer Design For Direct Torque Control” International Journal of Instrumentation and Control Systems (IJICS) Vol.2, No.4, pp 41-49.

[4] S.Kannan , S.Chinnaiya, S.U.Prabha Torque Ripple Minimization of Matrix Converter-Fed PMSM Drives Using Advanced Direct Torque Control International Journal of Soft Computing, Mathematics and Control (IJSCMC), Vol. 3, No. 4, November 2014pp. 51-68

[5] G. H. B. Foo and X. Zhang, "Constant Switching Frequency Based Direct Torque Control of Interior Permanent Magnet Synchronous Motors With Reduced Ripples and Fast Torque Dynamics," in IEEE Transactions on Power Electronics, vol. 31, no. 9, pp. 6485-6493, Sept. 2016.

[6] D.Premalatha, A.Sakthi Rubini “Direct Torque Control of Brushless DC Motor Using PI and Fuzzy Controller” International Journal of Science, Engineering and Technology Research (IJSETR) Volume 4, Issue 4, April 2015 pp.922-926

[7] Korkmaz, Fatih; Topaloğlu, İsmail; Mamur, Hayati; ,Fuzzy Logic Based Direct Torque Control of Induction Motor with Space Vector Modulation,"International Journal on Soft Computing, Artificial Intelligence and Applications",2,5/6,31-40,2015

[8] Bhim Singh, Pradeep Jain, A.P.Mittal, J.R.P.Gupta “Torque Ripples Minimization of DTC IPMSM Drive for the EV Propulsion System using a Neural Network” Journal of Power Electronics, Vol. 8, No. 1, January 2008 pp. 23-34

[9] F. Korkmaz “A New Approach to DTC Method For BLDC Motor Adjustable Speed Drives” The Fourth International Conference on Instrumentation and Control Systems(CICS-2016), vol. 6, no 5, pp-37–44, April 2016

[10] F. Korkmaz, I. Topaloglu, M.F. Cakir, R. Gurbuz, "Comparative performance evaluation of FOC and DTC controlled PMSM drives", Power Engineering, Energy and Electrical Drives (POWERENG), 2013 Fourth International Conference on , vol., no., pp.705–708, 13-17 May 2013.

[11] M. Masmoudi, B. E. Badsi and A. Masmoudi, "DTC of B4-Inverter-Fed BLDC Motor Drives With Reduced Torque Ripple During Sector-to-Sector Commutations," in IEEE Transactions on Power Electronics, vol. 29, no. 9, pp. 4855-4865, Sept. 2014.

[12] Vas, P., Sensorless Vector and Direct Torque Control, Cambridge University Press, 1998.

INTENTIONAL BLANK

THE IMPACT OF LEGO MINDSTORMS NXT ROBOT IN SCIENCE ON SIXTH GRADE STUDENTS IN BAHRAIN

Dr. Minimol Anil Job[1], Hasna'a Jaffer Mohammed Al Saeed[2]

[1]Assiatnt Professor, ITC Department, Arab Open University, Kingdom of Bahrain
[2]ITC Teacher, Al-Andalus Primary Girls School,
Ministry of Education, Kingdom of Bahrain

ABSTRACT

This research was conducted to gather information about the impact of usingLEGO MINDSTORMS NXT ROBOT in science on sixth grade students in primary schools, the need ofit in education, its advantages, disadvantages, hardware and software used to develop it and whether it fits with the physiological and cognitive developments characteristics of the students. The participants attended a LEGO robotic training programin three science experiments to study the motion, distance and speed principles. The researchers applied a science test, a student's questionnaire, a teacher's interview and a teacher's interview scale to gather the needed information. The results show that the LEGO robot is an efficient technology in teaching science, increases the students' academic achievement as the science test results show and the findings emphasize that this robot is a valuable technology in education and in primary schools as it develops the 21st century skills of the students.

KEYWORDS

Robot, Lego Mindstorms NXT Robot,Critical Thinking, Mindtool

1. INTRODUCTION

Educational field is one of the primary fields that benefits from new technologies. As we see in many countries education becomes rich of effective technologies. For example, in Bahrain, the big project, King Hamad Project for Future Schools, which was launched by Ministry of Education in 2004, opened the door for education in Bahrain to use new technologies (Smart board, Attractive board, Data Show, Camera….etc.) in order to use ICT in the learning process[1]. The learning process changes day by day, new strategies and new words come up such as collaborative learning, problem-solving strategy, students learning style…etc. Each of these strategies needs to use technological tools. So, as a new technology in education within these new strategies, it is the chance to enter the world of robotics as a new door for an attractive ICT tool in education in Bahrain. The researchers searched about using robotics in education and

David C. Wyld et al. (Eds) : ICAITA, CDKP, CMC, SOFT, SAI - 2016
pp. 73– 82, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.61307

found many exciting and successful experiments and studies about using the educational LEGO MINDSTORMS NXT ROBOT for teaching students' mathematics and science principles and how it is effective in problem-solving, creative thinking and critical thinking strategies [2] [3].LEGO Mindstorms is a line of programmable robotics/construction kits, manufactured by the Lego Group (LEGO Education, 2011). Lego Mindstorms is primarily used in secondary education but most universities also use the Lego Mindstorms NXT for their introductory courses in mobile robotics or for projects. There is now a new Lego Mindstorms Robot, called the EV3.In this research, the researchersused the Lego Mindstorms NXT 2.0 (8547), the second set from LEGO's Lego Mindstorms. The set contains 619 pieces, including a new sensor that can detect colors [4]. It provides students to control the behavior of a tangible model by means of a virtual environment and conduct science experiments, in which young students investigate a socio scientific issue using their scientific process skills both in and out of the classroom[5]. The aim of using this set is to study its impact in science on sixth grade students in primary schools in Bahrain.The aim of this research is to know more about robotics and the educational Lego Mindstorms NXT robot .What is the impact of using it in teaching science at sixth grade students? What are the advantages and disadvantages of using robotics in education? Does robotics fit with the physiologies and cognitive developments characteristics of sixth grade students? What are the software and hardware needed to apply this technology? Can Bahrain schools apply it or not?

2. OBJECTIVES OF THE STUDY

The objectives of this researchare to know the different types of robots, to identify the impact of educational Lego Mindstorms NXT robot in science classes of sixth grade students, to emphasize the advantages and disadvantages of robotics in schools, to know how robotics fit with the physiological and cognitive developments characteristics of sixth grade students and to know the software and hardware needed to use educational robotics in schools.

3. SIGNIFICANCE OF THE STUDY

The word "robot" expresses the relationship between the child and the games, especially[6][7]. It is seen as a new educational object that causes advanced teaching and the growth of kids. The robot role is seen as an educational tool that encourages students at all levels (Primary, Intermediate and Secondary) to study science. It is also noticed that the students are guided to study the principles of programming and engineering. Most of the students try to collect robots and program them to do a specific task. So, it opened up doors to the students to create, and apply scientific principles[8].Consequently, the importance of directing students to use the robot in the learning process can be summarized in the following points:

• Student participation in the educational process (Student-Centered learning process).

• Teaching the principles of science and mathematics through experimentation.

• Learn programming through programming the robot in order to perform a specific task.

• Develop problem-solving skills by building and programming the robot.

• Spark student's interest in Science and Technology.

4. REQUIREMENTS SPECIFICATIONS

The robot that used in this research is built using the LEGO Mindstorms NXT 2.0 (8547) kit which contains an intelligent NXT Lego brick with 32-bit microprocessor and a large matrix display, three interactive servo motors; four sensors (Ultrasonic Sensor, 2 Touch Sensors and the all-new Color Sensor).The color sensor has triple functionality: distinguishes color and light settings, and function as lamp [9][10].The software that will be used in this research is the MINDSTORM Software which lets the user to program the robot. The software is used to create a program, and then it is downloaded to the NXT by using the USB cable or the wireless Bluetooth connection. Mac- and PC-compatible drag- and- drop software gives the user 16 building and programming challenges, from beginner to expert [11][12].The researchers used as hardware the laptops with Microsoft Windows 7 to program the robot as well as the data show and the Smart board to display the work of the students.

5. RESEARCH METHODOLOGY

The researchers found that the MINDSTORM NXT robot is suitable for primary schools students as they can build it easily using the Lego parts to design it according to the function they want it to perform. Besides, this type of robot comes as a kit with a software CD that has the LEGO NXT-G programming language, a graphical language understandable by the small students[13] [14].First, we built the robot by hand programmed it using the NXT-G programming language software and tested it. Then it is tried it in the school with four sixth grade students.After receiving approval from the higher authority to apply this tools and the application on Bahrain government's schools, the researchers prepared the research tools which are questionnaire for students, interview questions for teachers, interview scale for the teacher interview and a science test for the students in two languages in Arabic and English. At first the researchers introduced the concept and objectives of the research to the science teacher then applied the LEGO training program and science experiments in motion, distance and speed on a group of six sixth grade students which is the experimental group and the other 6 students. The control groups have not attended the training program and they studied motion, distance and speed by the regular methods. The researchersused two laptops; one of them connected to the data show and used most of the time by the researchers. The other laptop used by the students for programming and designing instructions. The researchers asked some teachers to attend the training program and the science lessons to see the application of the LEGO Mindstorms NXT robot in science [15][16]. The researchers then applied a summative evaluation test on groups, the experimental and the control groups. The questionnaires were distributed to the experimental group students and interviewed the teachers who intended the training program and the science lessons. Also, the researchers will ask the teachers to fill the interview scale. The information that collected from student's questionnaire, teacher's interview, teacher's interview scale is used for the data analysis both quantitative descriptive. The analysis has done using both Microsoft Excel and SPSS.

6. STUDY POPULATION AND SAMPLING PROCEDURE

The total sample is 12 students in one primary school from sixth grade class divided into two equal groups, each group has 6 students. One group is an experimental group and the other is a control group. The sample is non-homogeneous since the students are differentiating in learning style and academic achievements. The sample of this research is a probability sample.

a. LEGO Training Program and Science Experiments

Since the students of the experimental group don't know about this robot before, the researchers will conduct a training program on LEGO Mindstorms NXT robot to make the students familiar with this robot and to develop their skills in designing and programming it to achieve a specific task before starting using it in the science lessons and experiments. The training program was a robotic curriculum prepared by the milset ASIA (the International Movement for Leisure Activities in Science and Technology) which was established on July 1987 in Kuwait ("Milset ASIA," 2014). This training program was applied before on Kuwait intermediate and primary schools as an activity program during the students' free times. The researchers scheduled one week training program divided into two classes' everyday as described in the table below.The students have taken works as one team with the defined rules; Engineer (builder) - 2 students, Software Specialist (programmer)-2 students, Information Specialist (Gets the necessary information for the team to move forward) - 1 student and Project Manager (Whip-cracker) - 1 student.

6.2 Science Experiments

After the application of the LEGO robot training program, the students designed the robot as a simple vehicle according to the design guideline embedded within the LEGO Mindstorms NXT2.0 software and it comes as a manual with the LEGO kit and the robot used for applying the science experiments to learn the scientific principles in motion, distance and speed. In each experiment, the students programmed the robot using the NXT-G programming language according to the rules given and observed the results and record the data. Finally, they deduced the science principle learned from the experiment.The researchers applied three experiments. The first experiment is about the linear motion in which the students will use the robot to know the definition of a motion, the meaning of the start and end points and to understand what a linear motion is by programming the robot to move forward in a straight line at a given time. That is to integrate technology with science. The experiments can be summarized by follows:

Experiment#1: Linear Motion

The goals of this experiment are to know the definition of the motion, to know what a start point and an end point are and to understand the linear motion. Also, its goal is to improve the robot designing skills of the students. The team will design a robot in the shape of a vehicle and they program it to move forward at a given time. Then, they will define the start point and the end point of the robot motion and deduce what type of motion the robot moves with.

Experiment#2: Distance Measurement

The goals of this experiment are to measure the distance that a robot passes, to improve the student skills in designing and programming the robot to pass a certain distance in addition to improve the robot designing skills of the students. The team will specify a start point for the robot movement and the end point where the robot stops and they will measure the length of the track between these two points. They should conclude that the distance is the total length of a track between two points and it is measured in meters.

Experiment#3: Speed

This experiment consists of two parts. The first part goals are to calculate the speed of the robot by measuring the distance it passes within a given time and to explain the relationship between the power, the distance and the speed by programming the robot to move forward with a predefined power of the motors where the students change the power three times and measure the distance and then calculate the speed of the robot. The second part of this experiment is similar to the first part. However, it aims to know the time needed for the robot to pass a known distance, to calculate the speed and then explain the relationship between the time and the speed. In this experiment, the students should know the unit used to measure the time.

7. DATA COLLECTION PROCEDURES

The researchers collected quantitative information from the test and the students questionnaire which applied on the students, and the teachers interview scale which answered by the interviewed teachers and the researchers collected the qualitative information from the teacher's interviews answers, teacher's interview scale, and the test results. In addition, the researchers collected qualitative information from the observations in control and experimental groups about student's reaction in the learning process in the lessons that the robot used for teaching.

8. DATA ANALYSIS AND RESULTS

a. Analysis of Personal Demography of the Student Test

This part contains the analysis of the student's test sample demography, which are age, and academic achievements of students in marks. The test was done by the experimental and control groups. Each group had 6 students. The total number of the test sample was 12 students, 67% of the students were 11 years old and 33% of them are 12 years old. This means that all the students in the experimental and control group are within the age of sixth grade students according to the Ministry of Education.According to the final average achievements of the science test for the first and second terms of the year 2013/2014, 50% of the students got marks between 100-90 which is A grade, 17% of the students got marks between 89-80 which is B grade, 8% of the students got marks between 79-70 which is C grade, 8% of the students got marks between 69-60 which is D grade and 17% of them got marks less than 60 which is F grade.

b. Analysis of Student's Questionnaire

The questionnaire of the students consists of three major sections, which were: learning objectives, learning process and LEGO-robot efficiency and it had two open questions. The questionnaire was answered by the experimental group students. The total number of the students was6 and all the questionnaires are accepted.

c. Analysis of the Learning Objectives

The mean values of the responses of the questions were between 4.67 and 5.00, the highest mean was 5.00 which states that the LEGO-robot helped me to know the relationship between the distance,the time and the speed and the participants agreed that LEGO-robot helped them to explain the relationships between the power and the speed, to know the definition of a motion, to

know the definition of a distance, to know what a start point is and also to know what an end point is and to calculate the speed of an object.

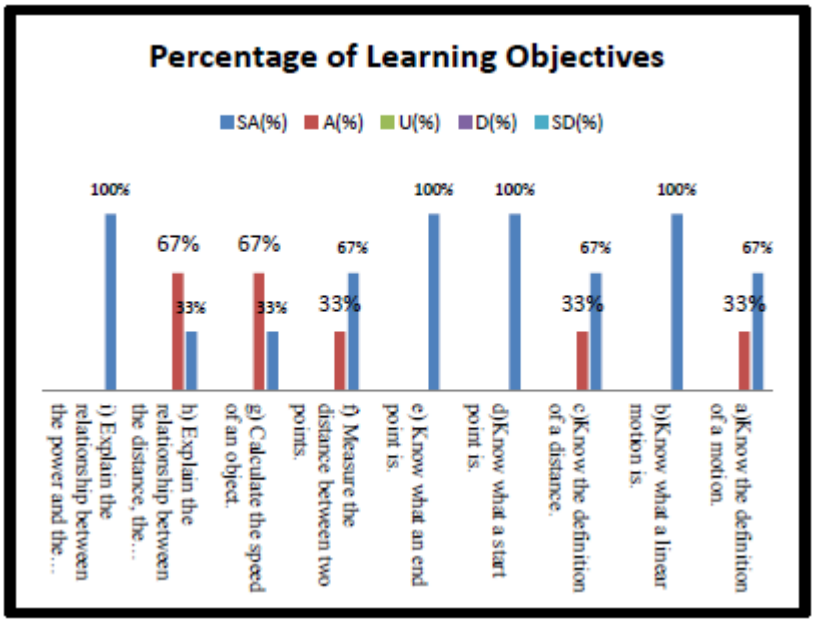


Figure1: Percentage of the LEGO-robot Efficiency Section in the Teacher's Interview Scale

d. Analysis of the LEGO-robot Efficiency

The mean values of the responses of the questions were between 1.17 and 5.00, the highest mean is 5.00 which states that the programming of the LEGO-robot was easy, the technology used (hardware and software) to build the LEGO-robot was adequate and the use of LEGO-robot need to know some computer basics, they we need to use LEGO-robot in future to learn other science basics and the programming language used to program the LEGO-robot was understandable. All the participants agreed that the design of the LEGO-robot was difficult and it is not just a toy.

9. PROGRAMMING OF THE SCIENCE EXPERIMENTS

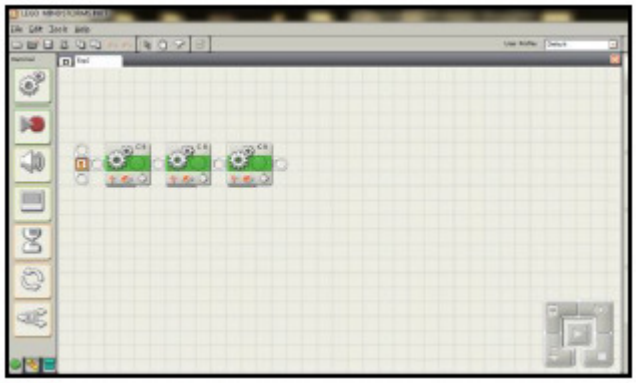


Figure 3: (Exp-1) Programming the LEGO robot to move forward for 3 rotations

Figure 3 shows the programming interface of the first experiment, where the student should programs the robot to move forward (as clear in the move block) for a given time, and in this experiment, the time is given as three rotations which is clear in the (Duration) part.

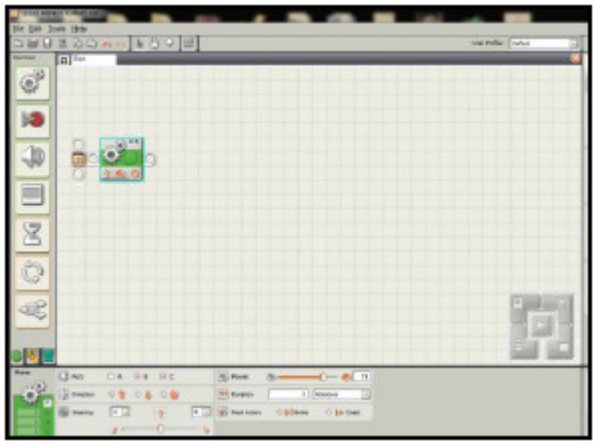


Figure 4: (Exp-2) Programming the LEGO robot to move forward three times with different durations

Figure 4 shows the programming interface of the second experiment, where the student should programs the robot to move forward from point A to point B for 10 seconds, then move it forward again from point B to C for 7 seconds , and finally move it forward from point C to D for 10 seconds.

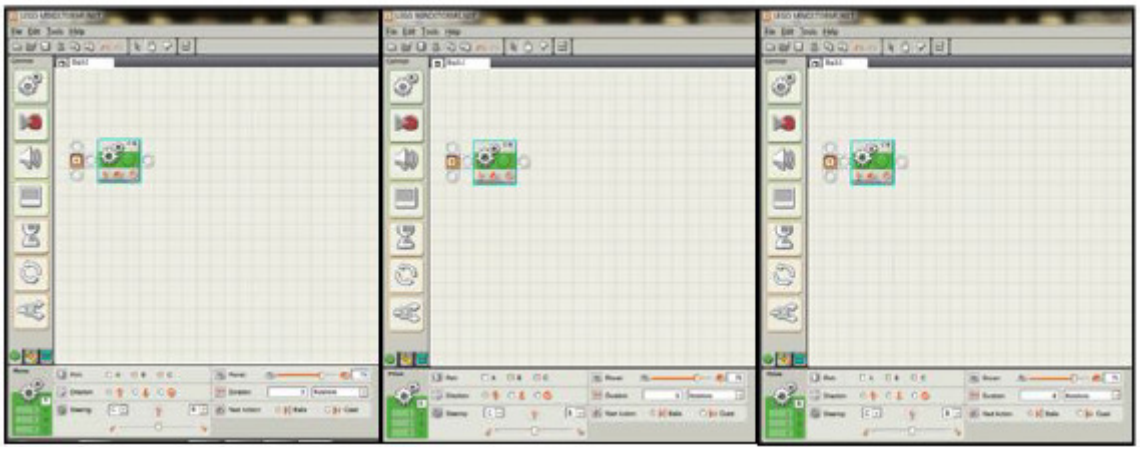


Figure 5: (Exp-3-1) Programming the LEGO robot to move forward three times with different powers for 10 minutes each time

Figure 5 shows the programming interface of the first part in the third experiment, where the student should programs the robot to move forward using different powers for three times, and she should change the power in each time. The power of the robot should be 50 at the beginning, then it should be changed to 75 and finally it should be changed to 100.

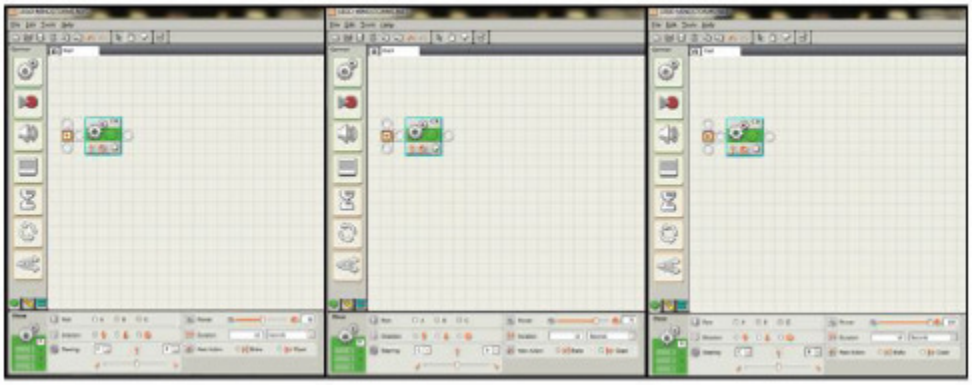


Figure 6 :(Exp-3-2) Programming the LEGO robot to move forward three times with different durations using different numbers of rotations

Figure 6 shows the programming interface of the second part in the third experiment, where the student should programs the robot to move forward using different numbers of rotations three times, and she should change the number of rotations in each time. The duration of the robot should be 3 rotations at the beginning, then it should be changed to 5 and finally it should be changed to 8.

10. DISCUSSION

The researchers conducts this research to determine the needs for the LEGO robots in primary schools, determine the advantages and disadvantages of using it in schools, how it fit with the physiological and cognitive development characteristics of sixth grade students, what are the softwares and the hardwares needed to build and program these robots and to determine whether Bahrain schools can apply it or not.

a. Research Findings and Conclusion

The key findings that the researchers collects from applying the LEGO MINDSTORMS NXT ROBOT, to collect information about the impact of using it in science on sixth grade students are:

- A robot is a machine used to perform jobs automatically, which is controlled by a computer, the basic parts of a robot are: sensors, effectors, and control system.

- The robots canbe divided into different types according to their applicationsas the following : industrial, household, medical, service, military, entertainment, hobby and copetition, and educational robots.

- LEGO MINDSTORMS NXT ROBOT is an educational robot that comes as a kit consists of different parts. The hardware of this robot consists of NXT brick, 619 lego parts, three interactive servo motors, and four sensors (2 touch sensors, 1 color sensor, 1 ultrasonic sensor) and the software is the LEGO MINDSTORMS NXT software that uses the NXT-G programming language.

- The LEGO robot is needed to use in education and in primary schools since, it fit with the new generation's requirements and fit with the 21th century learning skills (learning, literacy and life skills) as it improves the technology literacy and makes the students technology friendly, it develops the learning skills by improving the critical thinking, creative thinking , collaboration and communication of the students and it also improve the students' life skills in flexibility and leadership as they can manage the work of the team and learn independently.

- The advantages of the LEGO robot are: the flexibility in designing it in different shapes, easy to build and program, the cost of the kit as it can be bought by the students to continue practicing at home, its robust material and projects , the graphical interface of NXT-G programming language.

- The disadvantages of the LEGO robot are logistical , they are: cost of providing one robot for each student is expensive, the sensors problems, lost or damage of the lego parts and the problems with the power of batteries.

- The pedagogical theory that the LEGO robot fit with is: constructionism.

- The psychological theory that the LEGO robot fit with is: cognitive theory where the LEGO robot is suitable to be a mindtool.

11. RECOMMENDATIONS

The researchers recommends the following ideas for future works:

- To organize a robotic camp to study the impact of using the LEGO robot in STEM education by the participation of different STEM (science, technology, engineering and mathematics) teachers.

- To conduct a future research to evalute the effects of robotic activities on children's social skills.

- To use the LEGO robot to teach other subjects such as Computer , Maths, Physics, Art and English.

- To use the LEGO robot for different education stages, primary, intermediate and the secondary stages.

- To conduct a research to study the impact of using the LEGO robot with other smart technologies such as smartphones and IPads.

- To use the LEGO robot for further advanced research with smartphones to scan qr codes

REFERENCES

[1] Alimisis, D., Moro, M., Alergui, J.,Pina, A., Frangou, S.&Papanikolaou, K. (2007). Robotics and Constructivism in Education: the TERECoP project. euroLogo 2007 Bratisiava

[2] American Heritage Dictionary. (1995). http://americanheritage.yourdictionary.com/

[3] Barker, B. S., & Ansorge, J. (2007). Robotics as means to increase achievement scores in an informal learningenvironment. Journal of Research on Technology in Education, 39(3), 229-243.

[4] Benitti, F. B. V. (2012). Exploring the educational potential of robotics in schools: A systematic review. Computers & Education, 58(3), 978-988.

[5] Chambers, J. M., & Carbonaro, M. (2003). Designing, developing, and implementing a course on LEGO robotics for technology teacher education. Journal of Technology and Teacher Education, 11(2), 209-241.

[6] Christmas, M., Hanukkah, H., Kwanzaa, H. (2014). Educational Robots, Retrieved from: www.robots-and-androids.com/educational-robots.html

[7] Cocek, C., C. (2008). Lego Mindstorms and Critical Thinking Skills in the Elementary Classroom. Mount Saint Vincent University, p.12

[8] Framework for 21st Century Learning. (2011). Retrieved from: www.p21.org/our-work/p21-framework

[9] Jonassen, D. H. (2000). Computers as mindtools for schools. NJ: Prentice Hall.

[10] Kelly, J. (2010). LEGO MINDSTORMS NXT-G Programming Guide (Second Edition). USA, P1

[11] Korkmaz, O., Altun, H., Usta, E., &Ozkaya, A. (2014). THE EFFECT OF ACTIVITIES IN ROBOTIC APPLICATIONS ON STUDENTS' PERCEPTION ON THE NATURE OF SCIENCE AND STUDENTS' METAPHORS RELATED TO THE CONCEPT OF ROBOT. International Journal on New Trends in Education and Their Implications. 5 (2). Article: 05 . ISSN: 1309-6249

[12] Lego Mindstorms NXT Hardware Developer Kit. (n.d.). Retrieved May 30, 2014, from http://mindstorms.lego.com/en-us/support/files/default.aspx

[13] Milset ASIA. (2014). Retrieved from: www.milset.org/en/regions/asia.html

[14] Ruiz-del-Solar, J., & Aviles, R. (2004). Robotics Courses for Children as a Motivation Tool: The Chilean Experience. IEEE Transactions on Education, 47(4), 474-480. doi:10.1109/TE.2004.825063

[15] UCGUL, M. (2013). Hstory of Educational Potential of LEGO Mindstorms NXT, Mersin University Journal of the Faculty of Education. 9(2) pp.127-137

[16] Williams, D., Ma, Y., Prejean, L., & Ford, M. (2008). Acquisition of Physics Content Knowledge and Scientific Inquiry Skills in a Robotics Summer Camp, Journal of Research on Technology in Education, 40(2), 201-216

APPROACH MULTI-AGENTS EMBEDDED ALARM IN POTROOMS

Luis Carlos Costa Fonseca[1], Alan Robert da Silva Ribeiro[2] and Jonh Selmo de Souza do Nascimento[3]

[1,2,3]Department of Computer Engineering and System,
State University of Maranhão, São Luis, Maranhão, Brazil
[1]luiscarlos.uema@gmail.com
[2]alan.robert2617@gmail.com
[3]jonhselmo@pecs.uema.br

ABSTRACT

Industrial Shop Floor environments require fast intervention of the controller's computers and operators to ensure high industrial production efficiency. This work focuses the electrolytic potrooms process control efficiency. The main goal of this work is to design an embedded solution for detection and alarm, using multi-agents system technologies so that controllers can alert plant operators about the problems in the electrolytic pot still malfunctioning after controller intervention. If the controller action was unsuccessful due to a feeder and pot bus problem, an audio alarm is immediately issued to the potrooms so that the operator can be notified about the specific problem, independently of the potrooms location.

KEYWORDS

Alarm, Potrooms, Agents, Embedded Systems, Pot

1. INTRODUCTION

In the aluminum production process, the cost and the quality of the product is strongly affected by the amount of alumina added to the pot production. The resistance and variations of first and second orders are used to assess the percentage of the alumina concentration in the bath [1]. The resistance variation is the primary variable used as an evaluation metric of vessel behavior. In practical terms, the monitoring and control of alumina in the bath changes are implemented in accordance with the assessment of the resistive variation. There are situations, however, where the electrolytic process control computers, even sending signals to the alumina feeding actuators and movement of buses that regulate the resistance of the electrolytic cells allow better control of the amount of alumina in the pot; this guarantee control may be interrupted by any mechanical failures in actuators and busbars.

Therefore, rapid actuation of the vessel operator to reestablishment of normal operation of the equipment and continuity of the aluminum production process with higher production gain becomes a daily preoccupation of management, to the extent that an aluminum production room comprises hundreds of devices distributed over large production halls. Figure 1 shows the Alumar

David C. Wyld et al. (Eds) : ICAITA, CDKP, CMC, SOFT, SAI - 2016
pp. 83– 95, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.61308

potroom which has three production lines with 204, 250 and 256 electrolytic cells, each, for a total of 710 electrolytic cells.

Figure 1. Reduction Alumar

Note that potrooms are arranged in large halls in which the electrolytic cells are disposed in series, side by side. These pots are controlled by the process control computers, which send signals to drive the feeders allowing it the addition of alumina in the pot. In addition, the control computer is responsible for sending signals to drive the mechanical drive bus, responsible for the regulation of the electrolytic tank resistance variation.

When, perhaps a mechanical failure occurs in one of these devices, we have as a consequence a delay in the correction and re-establishment of the operation of the electrolytic pot, causing the excess in the bus drive and alumina supply in the pot, due to unauthorized drives by computer control generating the accumulation of alumina or the lack of it, causing increased resistance to passage of electric current and negatively affecting the efficiency in aluminum production.
Thus, it becomes necessary to adopt a technique that best fits the application of electrolytic process control computer, in order to provide the necessary information and alerts and can thus overcome and minimize problems that may arise, for example, anode effect, unscheduled shutdown pot, bored pot, or quality time requirements and life for the potrooms.

Thus, an embedded solution with communications technology based on multi-agent systems is built based on stimulus-response model, in order to map and detect all abnormalities prevent the correct operation of the electrolytic pot, even with sending signals performance controllers. If signs of activity are not sufficient to establish the acceptable level of resistance change, and consequently, of alumina in the electrolytic pot, then audio alarms are generated immediately for operators and technicians of the potrooms (factory floor) can understand the situation and carry out the necessary intervention to solve the physical problem of equipment. The system is accessible and can be easily implemented in the automation industries that have controllers such as PLC (programming logic controller) and process control computer.

The technologies investigated in this research are the Process Control System, Multi-Agent Systems and Embedded Software, all used in the construction of a monitoring and real-time alarm solution. To perform this analysis, it is used as a case study alarm management system for aluminum production halls, rooms of tanks, which will be built using these technologies.

Is approached, therefore, the need for detection and alarm automation environments and control in which controllers can not function correctly and precisely, due to mechanical problems in the equipment that receive the command signals for actuation, as the emitted pulses the controllers. In the case of large factory floor environments where each electrolytic tank requires their respective actuators work in real time, the extent to which the actuators are triggered, the delay in identifying problems of this nature can lead to loss of productivity and own electrolytic pot of aluminum production, generating as a result of damage to the production process.

The contribution of this article is to develop an embedded management solution and alarm based on the standard observe and react, using multi-agent systems for monitoring and conversion of messages generated by the controllers, allowing the broadcast in real time, state and environmental behavior, allowing adoption of preventive actions and operational in the shortest time interval, by the vats operators working there. It consists in building a software solution embedded in a raspberry PI plate for automation and control environment, based on the response stimulus model, able to communicate with controllers electrolytic cells through the use of multi-agent systems, allowing communication between centralizing agents and controllers for transmitting / receiving text messages sent, generated by computer control, warning of each of the potential problems identified in several electrolytic cells. These text messages are received by the centralizing agent that converts audio messages transmitted to speakers of channels distributed over the potrooms, allowing the technician and in charge of the process in it can identify a pot with problem despite being in a distant location of the problem.

2. CONTROL SYSTEM POTROOM

In general, the process of aluminum production is based on the amount of alumina (Al2O3), the raw material added to the electrolytic bath. In a bowl, electricity circulates between a positive (cathode) and a negative pole (anode). That of the alumina mixing process, existing electrolytic bath in the pot, and passing electric current anode cathode, it is apparent that liquid aluminum is removed from the vessel for the production of ingots. Figure 2 below shows the generic functional architecture of the aluminum production process, from the perspective of control systems. This architecture is applied to the modeling of a system of pots, specifically for monitoring purposes of their states and implementation of its controls.

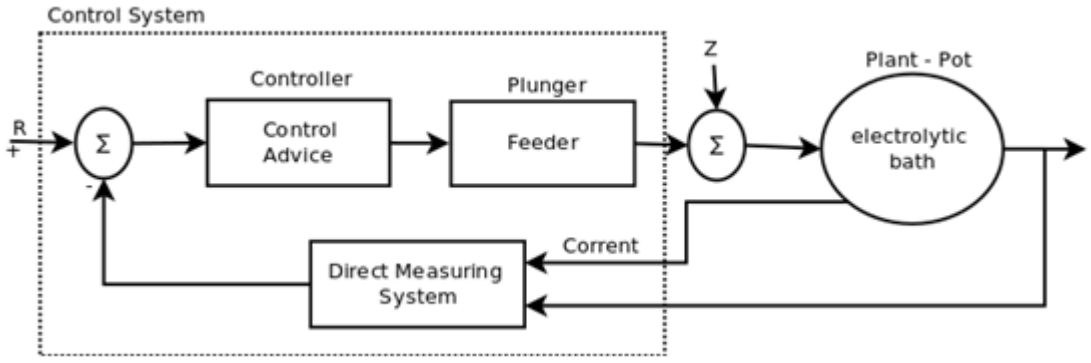


Figure 2. Block diagram for distribution of alumina in vessel

The control system comprises sensor, controller and actuator. First, it is an indirect measurement system, because from voltage and current signals and other information such as operation status pot, to estimate quantities or rates those are used by the control device. The control system processes the signals and these are used to making decisions that interfere with the production process behavior. For example, control can promote actions to modify the frequency of operation of Al2O3 feeder.

This control system is run on computer called process control computer, and has as main objective to control the supply of alumina to the electrolytic cells, whether pre-baked type or Soderberg point-feeder [2]. The intention is to keep a quantity of alumina dissolved in the bath to ensure maximum production of aluminum. In pots of pre-baked type, the optimum concentration of alumina in the bath is around 3% in the electrolytic bath. However, there are large variations in this concentration, depending on the operations carried out in vats and variations of other control parameters [3]. In practice, there is no method to directly measure this quantity or alumina concentration in the electrolytic bath in the process of speed control requires [4]. Therefore, the process computer controls the amount of alumina should meet the demand of alumina is added to the vessel within the rated power range. The feed interval set in seconds, is the time window in which the computer will send a pulse to drive the alumina feeders, wherein said range is from pot to pot. It means that each production cell (electrolytic tank) has its power gap target.

When the control computer sends a signal to drive the feeders, and for some reason, such feeders are not triggered, it means that some physical problem may be occurring. In this case, a physical intervention such equipment is required. For this task, they are designated process technicians who perform the repair and the reestablishment of the operation of the feeders and voltage busbars. The problem becomes complex when technicians deal in a potroom environment with large amounts of electrolytic cells arranged in large rooms and corridors in an aluminum production plant.

3. MULTI-AGENT SYSTEMS

The definition of an agent can be expressed in many different ways. [5] defines an agent as a computational entity software, located in a given environment, which has the perception that environment via sensors has reasoning ability and acts autonomously in that environment through actuators so to perform a specific function for which it was designed. [6], defined as something that can be seen as having the perception of an environment through sensors and acting in this environment through actuators. For [7] would be a piece of hardware or a computer system based on software that enjoys some properties, such as learning autonomy, responsiveness, pro-activity and social skills.

Figure 3 shows the interaction between two or more agents in a computer system and configures a Multi-Agent System where these work together to perform certain tasks or perform a set of goals. Scientific research and the practical implementation of Multi-Agent Systems is focused on building standards, principles and models that allow the creation of small and large companies of semi-autonomous agents, able to conveniently interact in order to achieve their objectives [8].

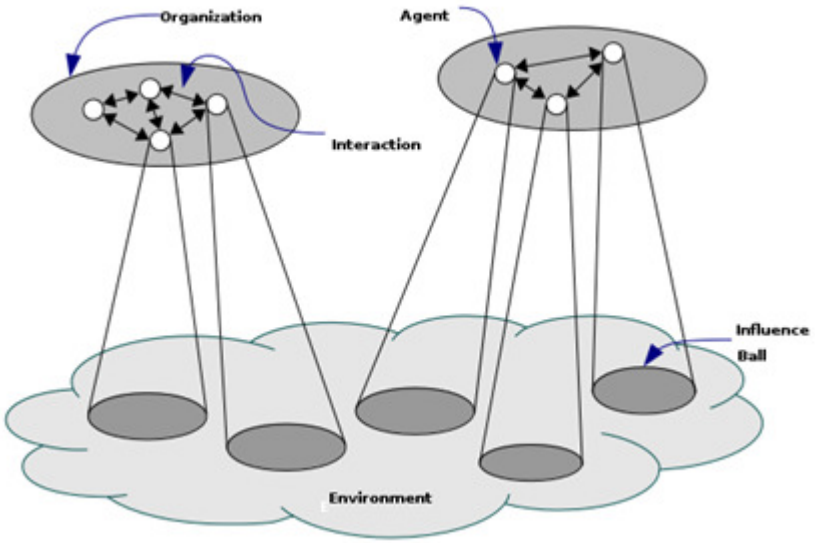


Figure 3. Structure of a Multi-Agent System

3.1. Jade

JADE tool (Java Agent Development Framework) began in 1998 through a CSELT (Telecom Italia), whose development was motivated by the need for validation of specifications of FIPA (Foundation for Intelligent, Physical Agents). AvJADE in 2000, became an open source platform, being distributed by Telecom Italia on the LGPL (Library Gnu Public License) [9].

JADE provides certain features, such as distributed applications that exploit the use of software agents. Being built on the Java platform, enables deployed agents using the JADE framework, which can run on any operating system (OS), which makes it ideal for environments where you can not get information about the OS used [10].

Other JADE's features are the development of agents and control of them via a graphical interface [9].

3.2. Auml

Unified Modeling Language Agent (Auml) is a standardized graphical modeling language by FIPA (Foundation Intelligent Physical Agents). Auml was proposed as an extension of UML (Unified Modeling Language), which uses decomposition, abstraction and organization to reduce the complexity of software development, decomposing a system in small parts of objects, models, use case or class, several operational actions. In relation to abstraction, it provides a specialized abstract view of modeling (class, use case diagram of interface etc.) and is used to create a set of semantics and operating service conditions and infrastructure [11].

The Auml offers structures as a class diagram and interface to show how agents can be erected in an agent system. The model focuses on a point at a time, increasing the ability to understand the issues to complex problems during the lifetime of the system design.

The main parts of Auml are the mechanisms for modeling multi-agent interaction protocols. This is accomplished by introducing a new class diagrams UML diagrams protocol. These extend diagrams and sequence diagrams of UML state, as shown in Figures 4 and 5.

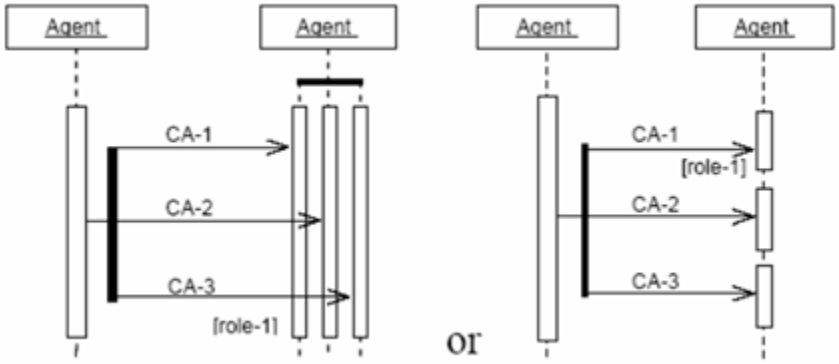


Figure 4. AUML – Diagram Sequence

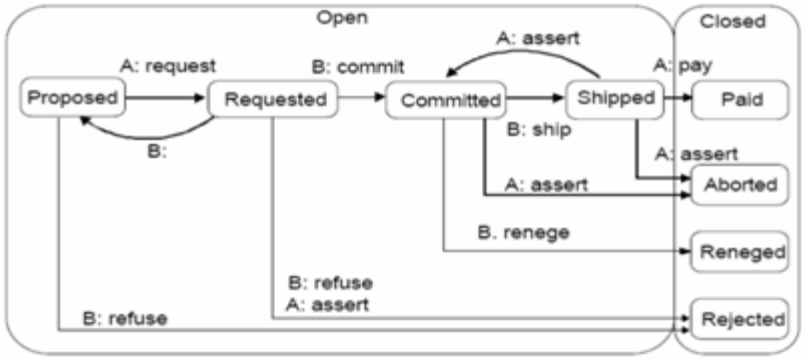


Figure 5. AUML – State Diagram

4. EMBEDDED SYSTEMS

The embedded systems are a combination of computer hardware and software, sometimes having mechanical parts designed to perform a dedicated function, wherein such systems are often only a small part of a larger system [12].

Typically, they are implemented from different technologies such as microprocessors, microcontrollers, DSP, reconfigurable circuits, analog circuits and microwave and even microelectromechanical systems (MEMS - Micro Electro Mechanical Systems) [13].

4.1. Raspberry PI

The Raspberry Pi model B is a small computer that uses the Broadcom BCM2835 multimedia processor, the SoC type (system-on-chip), with 700 MHz 32-bit, built on the ARM11 architecture [10]. The board is fed by a micro-USB port, with 5V voltage and requiring at least 700mA, with waste of energy varying in 3.5 W [14].

The Raspberry Pi, as shown inFigure 6, has no hard disk on your hardware composition, using only one memory card. For operating system installation it is necessary that the card has at least 2 GB capacity for all the necessary files. The operating system commonly used is the Raspibian, a distribution recommended by Raspberry Foundation, based on Debian [15].

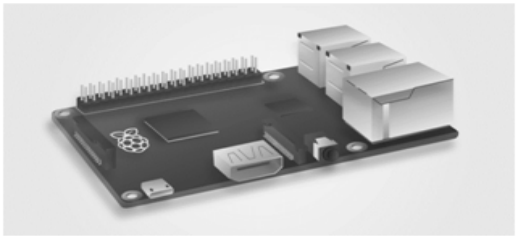


Figure 6. Raspberry B+ model

5. RELATED WORK

Binnicker [16] in 2015 presents a tool to detect and alert with temperature sensors for vehicles occupied by passengers. The system alerts drivers if a child is left unattended in the vehicle in order to prevent the child from being exposed to high or low temperatures inside the vehicle. To this end, sensors are used to capture data via occupancy sensors, temperature and distance. The data captured by the sensors determine the need to alert sent to the vehicle owner (driver) if an occupant has been left inside the vehicle, still controlling some functions of the vehicle, allowing foreign individuals to rescue him, with the unlocking of doors, windows opening, with possible activation temperature of the automobile system.

M. Kumar et al. in [17] designed and developed an intelligent detection and intrusion alert system designed to increase security, with real positive identification probability of attackers and intruders, compared to other electronic security systems already used. Through multiple sensors, the system evaluates the extent of danger exhibited by a person or animal within the confines of the residence, conveying the owner of the residence, the various information about critical events generated.

Ramya [18] presents a construction work involving containing a microcontroller system for detecting and warning of toxic gases, particularly propane and LGP. Specific sensors for identification of these two noxious gases are added to the microcontroller, which through the use of analog / digital converter capture and analyze the information, which are also shown in the display. If the gases exceed acceptable limits, the system immediately generates an alarm and also send SMS alert message to authorized persons using GSM technology.

Gaspar et. al. [19] addresses the adoption of an alarm management philosophy for incident management that avoids the generation of a huge amount of alarms in order to make life easier for technicians who must manage in real time complex and dynamic environments and good management alarms can provide crucial information to identify the cause of the fault and replace the plant in normal operating state.

Morales [20] created and patented a monitoring and alarm system in a network audio link that detects sound alarms triggered and inform the responsible authorities. The system and method

created based on the audio produced from alarms which are triggered. Implemented for PC type computers or other premises of processors, different sound coming from different alarms, classifies the sound as a specific alarm by FFT processing and distributes the alarm over the Internet, public telephone switches, or other communication links to central station, which then distributes the alarm to the authorities. In another embodiment, the alarm reports are generated directly from the PC to the residence or specific company where authority is located.

6. PRELIMINARY RESULTS

As preliminary results, diagrams were built in Agent UML (Auml) for multi-agent embedded proposed solution, which includes the deployment diagram, use cases, activities, block, and an embedded prototype raspberry pi with jade platform running on environment simulation, where the centralizing agent is embedded in the board and the controlling agents are available in environments that simulate controllers. All diagrams were constructed using the free tool-day standard 0.97.2 (http://live.gnome.org/Dia). The source code of the solution multi-agent was developed in Java with Netbeans IDE version 8.1, where the environment is multi-agent on the implementation platform JADE, in version 4.4.0.

6.1. Diagram Use Cases

Figure 7 shows the actors are placed in the context of the proposed approach and its use cases, showing the set of paths that agents can run through the system, each with a discrete goal. The actors representing the controlling system of tanks in practice can be represented by any controller technology, this PLC (programmable logic controllers) to real-time vats control applications that interact with Analog Stick / digital converters, analog and digital cards Mustangs.

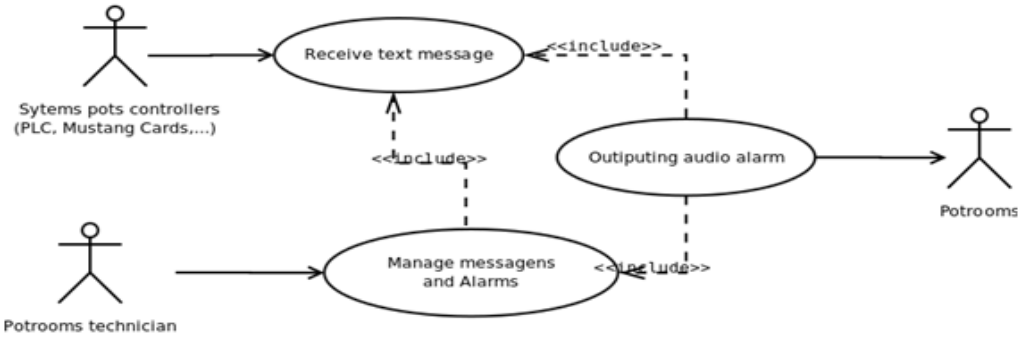


Figure 7. Diagram Use Cases of the proposed system

The use case Send audio alarm, converts the textual words received by dividing each of them into audio files (eg. .wav) that are sent to their rooms vats according to the parameter that identifies the number of pot with problem.

6.2. Transition State Diagram

The Transition State Diagram shown in Figure 8 represents the state and associated transitions to the communication protocol between the developed agents and between the controller and agents,

which at the end of processing, transfer to the vat room audible alarm identifying a potential problem in an electrolytic tank. Is represented in this case, the level of interaction between players by exchanging messages and activation through behaviors.

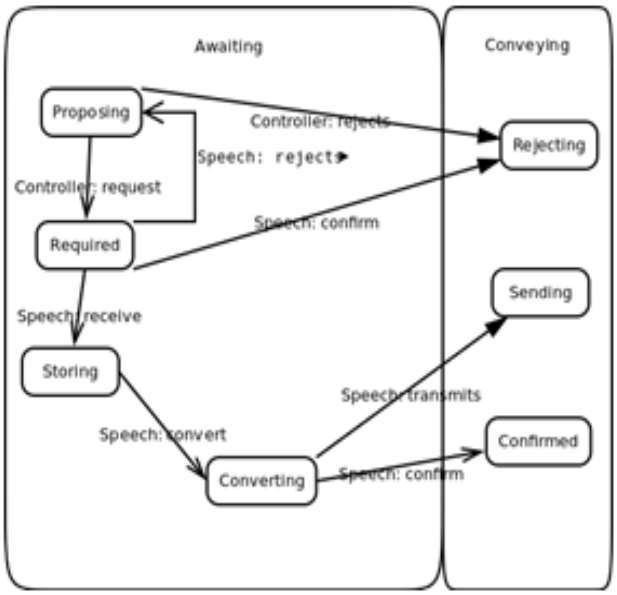


Figure 8. State Transition Diagram of the Proposed System

Through this diagram, we see the main actors interacting in the sending and receiving of text messages parameterized (Controller and Speech) that is validated, converted to audio and broadcast**.**

6.3.Diagram Deployment

The deployment diagram shown in Figure 9 detailing the distribution of the proposed solution in a generic environment potroom controlled by the controlling computer. It is observed that the bundled solution raspberry pi plate is arranged in the process control network, in the same environment controllers. The audio channel raspberry card is connected to speakers arranged in the potroom. Therefore, the power pulses, ascent and descent bus, when they jammed by mechanical failures in the vats can be signaled in the form of beep, in which the solution proposal report to the operator exactly what existing physical problem.

6.4. Block Diagram of the Proposed Solution

Each controller agent is responsible for receiving the parameterized messages generated by the control computer in which the agent is inserted. From the generation and reception of such text messages, with the use of multi-agent platform JADE, the controlling agents send text messages with the problem identified in a given tank to the centralizing agent, which queues the message and is in charge of is converted it audio, transferring them to the speakers of the respective room vats, where the pot with problem lies. The solution proposed is the block diagram shown in Figure 10.

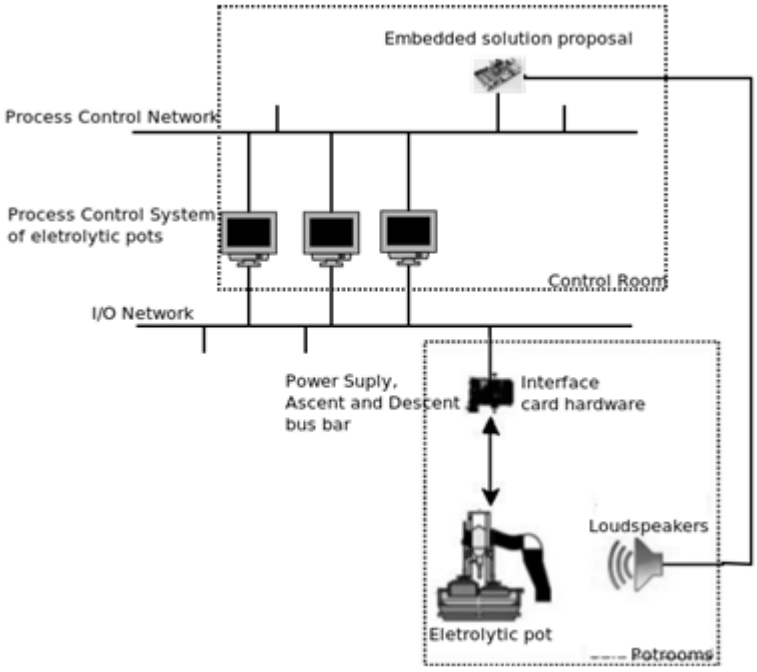


Figure 9. Deployment Diagram of the Proposed System

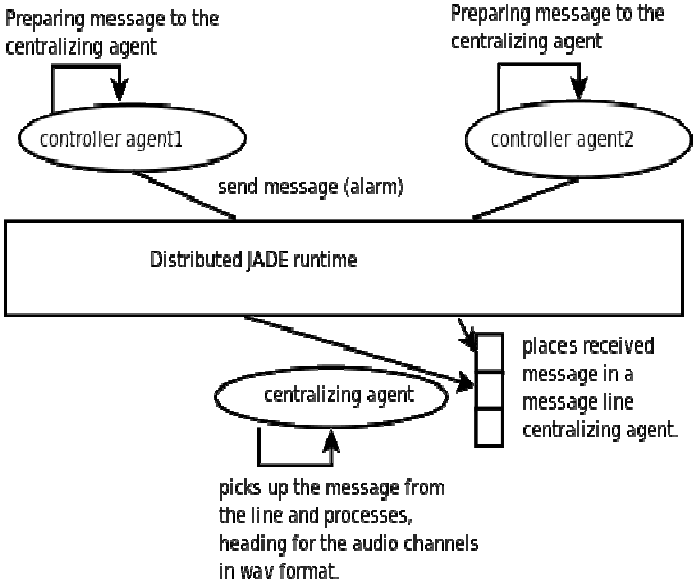


Figure 10. Solution Block Diagram proposal

7. IMPLEMENTATION

Alarms tanks with high voltage, much increase and / or bus down, swings, and its operating status were detected by the controlling agents and sent to the centralizing agent of the respective potroom, residing in raspberry pi. The centralizing agent identifies the queue textual alarm messages that are sent in FIFO order (first-in, first-out) to the potroom in an audible format, as a warning system, which helps in faster diffusion of the critical situation found. The prototype of the proposed system is shown in Figure 11.

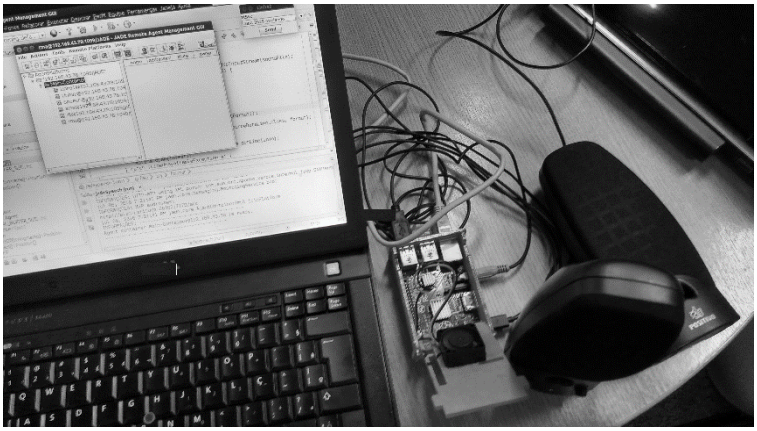


Figure 11. Prototype of the Proposed System

8. CONCLUSIONS

An embedded system using multi-agent platform for detection and alarm electrolytic cells in potrooms was implemented. For demonstration purposes they were made simulation controllers pot and controlling and centralizing agents responsible the generation and transmission of alarms in audio format for the tanks room, respectively. The extent to which the controllers identify some kind of problem in a particular pot, textual message of the problem in question, parameterized for the pot number identifier is generated and transmitted to the controller agent forwards to the centralizing agent. This solution may be suitable for any shop floor environment where there is a process automation for large equipment and controllable mechanisms that need to maximize their production with minimal stop as a function of real problems identified the equipment, enabling operators and technicians act to minimize downtime.

REFERENCES

[1] Fiona, Stevens McFaden & Bearne, Geoffrey (2001) Apliccation of advanced process control to aluminium reduction cell - a review, TMS- The Mineral, Metal & Materials Society.

[2] Braga, Carlos A. P., Martin, S., Nagem N. F, Ritter, C., Silva, Ari, Verlihay, Mark (2007) Faster Anode Effect Kill. TMS Letters, v. 50, p. 417-422.

[3] Silva, Ari (1995) Alumina feed control system of line 1 and line 2 potrooms at ALUMAR. ABAL – Associação Brasileira de Alumínio.

[4] Silva, Ari, Carlos Braga, Eliezer Batista (2002) Smelter personnel technical training on aluminum production. ALUMAR - Consórcio de Alumínio do Maranhão.

[5] Reis, Luis Paulo (2003) Coordination in Multi-Agent Systems: Applications in University Management and Robotic Soccer. PhD, Electrical and Computers Engineering, FEUP, Porto.

[6] Russel, Stuart. "NORVIG, Peter." (2013) Inteligência Artificial. Editora Campus. Tradução da 3a edição.

[7] Wooldridge, Michael, and Nicholas R. Jennings (1995) Intelligent agents: Theory and practice. The knowledge engineering review 10.02: 115-152.

[8] Lesser, Victor (1999) Cooperative Multi-Agent Systems: A Personal View of the State of the Art, IEEE Transactions on Knowledge and Data Engineering, Vol. 11, Nº 1.

[9] Telecom Italia - JADE - http://jade.tilab.com/

[10] Bellifemine, F., G. Caire, D. Greenwood, e I. NetLibrary (2007) Developing Multi-agent Systems with JADE. Springer

[11] Gatti, M., Von Staa, A., Lucena, C. (2007) AUML-BP: A Basic Agent Oriented Software Development Process Model Using AUML., PUC-Rio, n. 21/07, p. 25.

[12] William, Stallings (2010) Arquitetura e Organização de Computadores.

[13] Dias, Ailton Fernando (2001) "Concepção Conjunta Hardware/Software de Sistemas Embarcados de Processamento de Imagens." CDTN/CNEN. Belo Horizonte, Brasil.

[14] Richardson, Matt, and Shawn Wallace (2013) "Primeiros passos com o Raspberry Pi." Primeira Edição. Novatec Editora Ltda.

[15] Raspberry PI Foundation: http://www.raspberrypi.org/downloads/

[16] Binnicker, G. (2015) Temperature-Sensitive Vehicle Occupancy Detection and Alert System. https://www.google.com/patents/US9428109

[17] S. Kumar, et al. (2015) "A Pattern Matching Model For Misuse Intrusion Detection,", pp. 1-11, (found at http://www.cs.purdue.edu/coast/archive/data/categ24.html).

[18] Ramya V., Palaniappan B. (2012) Embedded system for Hazardous Gas detection and Alerting. International Journal of Distributed and Parallel Systems (IJDPS) Vol.3, No.3.

[19] Gaspar, D., Ferreira, D., Coutinho, J., Correia, J., Albuquerque, P. (2010). A Política de Alarmes na Gestão da Manutenção de uma Indústria Farmacêutica. Millenium, 39: 135-151.

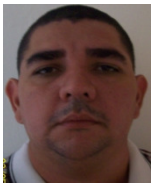
[20] Morales, F. (2001) “Network audio-link fire alarm monitoring system and method” https://www.google.com/patents/US6215404

[21] Braga, Carlos A., Nagem,N.F, Fonseca, J. V. (2007) “Bandwidth adjustment of digital kalman filter for indirect measurement of electrolytic bath state variables”. 2nd International Conference on Sensing Technology.

AUTHORS

Luis Carlos Fonseca graduated in Technology in the Data Processing of Maranhão University (2000), Master in Engineering Electricity from Federal University of Maranhão (2003) and PhD in Computer Science in Education from the Federal University of Rio Grande do Sul (2009). He is currently assistant professor at the State University of Maranhão. He has experience in computer science with an emphasis in Information Systems, mainly in the following areas: Artificial Intelligence, Information Technology in Education, Software Engineering, Information Retrieval, Recommender Systems.

Alan Robert da Silva Ribeiro graduated in Bachelor of Computer Science from Federal University of Maranhão (1999), specialization in Systems Analysis and Design from the Federal University of Maranhão (2006), specialized in Advanced Systems for Internet and Intranet at University Center Maranhão (2007) and is Masters Student of Computer Engineering and Systems at the State University of Maranhão. It is currently a Ministerial Analyst and Coordinator of the Public Ministry of the Information Technology area of the state of Maranhão, having worked as a Senior Process Control Systems Analyst at multinational ALCOA, which was part of SDG global group (Service Delivery Group), responsible for development and installation of process control systems in the company's factories in the world. He has experience in the area of Computer Science, Aluminum Process Control, Information Technology Management, acting on the following topics: Electrolytic Control Computer, Artificial Intelligence, Information Technology Governance, Database, Ubiquitous Computing and Driven Engineering by models.



John Selmo de Souza do Nascimento holds a degree in Computer Engineering with emphasis on Automation and Process Control at the State University of Maranhão (2015). Currently is Masters Student of Computer and Systems Engineering from State University of Maranhão. He has experience in Geosciences with an emphasis on Georeferenced systems, robotics, mechatronics and automation.

AUTHOR INDEX