e-Proceeding of Software Engineering Postgraduates Workshop (SEPoW) 2013

Theme: Innovative Software Engineering for Creative and Competitive Software Solution

19 November 2013
City Campus
Universiti Teknikal Malaysia Melaka

Organizer

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Software Engineering Department,
Faculty of Information and Communication Technology (FTMK)

Co-Organizer

MySEIG.org

Malaysian Software Engineering Interest Group

e-ISBN 978-967-0257-33-4

Sponsor

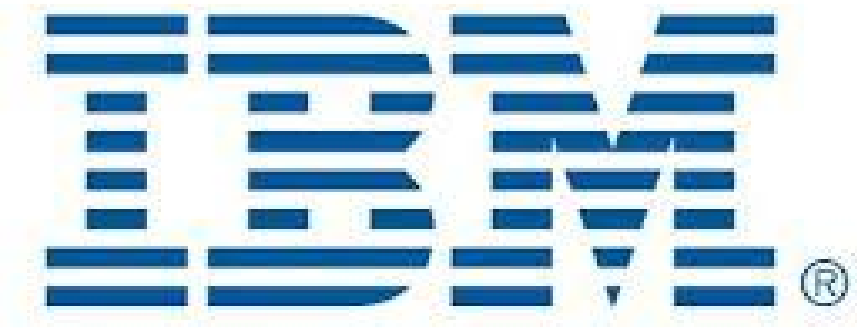

Table of Contents

Foreword from SEPoW 2013 Patron

It is my great pleasure to welcome you to the Fourth Software Engineering Postgraduates Workshop (SEPoW) 2013 - Innovative Software Engineering for Creative and Competitive Software Solution.

This year's workshop continues its tradition of being the preferred forum for presentation of postgraduate research activities relevant to software engineering. SEPoW gives researchers, experts and practitioners a unique opportunity, not only to share their research activities, but also to update and seek synergistic collaborations in the various facets of software engineering including software models, architectures, requirements, design and development, education and testing.

We also encourage attendees to attend the tutorial presentations. These valuable and insightful talks can and will guide us to a better understanding of software engineering research by Prof. Dr Ali Selamat (Universiti Teknologi Malaysia) and Dr Norsaremah Salleh (International Islamic University Malaysia).

Putting together SEPoW 2013 was a team effort. We first thank the speakers and authors for providing the content of the program. We are grateful to the organizing committee who worked very hard in reviewing papers and providing feedback for authors. Finally, I thank the co-organizer, the Malaysian Software Engineering Interest Group (MySEIG), our sponsors and corporate sponsors, IBM.

I hope that you will find this program interesting and thought-provoking and that the workshop will provide you with a valuable opportunity to share ideas with other researchers and practitioners from institutions around the world. I wish you all an exciting program and an unforgettable stay in Melaka.

PROF. DATUK DR. AHMAD YUSOFF HASSAN
Vice Chancellor
Universiti Teknikal Malaysia Melaka

Welcome from the SEPoW 2013 Chair

Dear SEPoW 2013 delegates,

On behalf of the committee of SEPoW 2013, I would like to extend a warm welcome to the participants of the fourth Software Engineering Postgraduate Workshop (SEPoW 2013) at UTeM, located at the historical city of Melaka. We hope that you will take this opportunity to gain valuable knowledge and experience from the extensive Technical Programs and Exhibits, while at the same time enjoying the charm and hospitality of this beautiful state.

This workshop is a co-located event with MySEC 2013, and it focuses in the area of Software Engineering (SE). In conjunction to our theme this year: *"Innovative Software Engineering for Creative and Competitive Software Solution",* this workshop is organised to enable us to exchange innovative and creative ideas in SE, both from the research/academic standpoint as well as the practical application perspective; thus, it helps us to develop creative and competitive research and products in various application of domains. We hope that you will take full advantage of the many benefits offered by this workshop.

The main objective of this workshop is to provide a platform to SE postgraduates students from all universities in Malaysia to present their research work to the SE academic community. An important audience in the parallel sessions will our guest panellists from the Malaysian Software Engineering Interest Group (MySEIG). They will give valuable insights and feedbacks in validating the novelty of your research and improving your research projects; thus, this will assist you to successfully complete your research project. Further, two presentations, entitled: "*How to Produce and Publish a Good SE Research?"* and *"Writing a Systematic Literature Review for SE research*" provide participants with the essential knowledge on how to conduct innovative SE research as needed by the industry.

Finally, we hope you will enjoy the content of the workshop and make full use of this avenue to build network with other researchers in the area of SE. Thanks to all authors and presenters for your fine work in this area, and most importantly, to the committee of SEPoW 2013 for your commitment in organising this Conference and putting together this fine program.

Dr. Massila Kamalrudin
Chair of SEPoW 2013
Universiti Teknikal Malaysia Melaka

Tutorial Talks

Prof Dr. Ali Selamat

Ali Selamat has received a B.Sc. (Hons.) in IT from Teesside University, U.K. and M.Sc. in Distributed Multimedia Interactive Systems from Lancaster University, U.K. in 1997 and 1998, respectively. He has received a Dr. Eng. degree from Osaka Prefecture University, Japan in 2003. Currently, he is the Dean of Research Alliance in Knowledge Economy (K-Economy RA) UTM. He is also a professor at the Software Engineering Department, Faculty of Computing UTM. Previously he was an IT Manager at School of Graduate Studies (SPS), UTM. He is also a head of Software Engineering Research Group (SERG), K-Economy Research Alliance, UTM. He is the editors of International Journal of Digital Content Technology and its Applications (JDCTA) and International Journal of Advancements in Computing Technology (IJACT). His research interests include software engineering, software agents, web engineering, information retrievals, pattern recognitions, genetic algorithms, neural networks and soft-computing.

He was a conference chair of Malaysian Conference in Software Engineering 2011 (MySEC2011) and Asian Conference in Intelligent, Information and Database Systems 2013 (ACIIDS2013). He is currently a treasurer of IEEE Malaysia Computer Society Chapter and also the auditor of IEEE Malaysia Section. He also involved as the organizing committee and advisor of Malaysia Conference in Software Engineering 2013 (MySEC2013) and 2nd International Conference on Interactive Digital Media (ICIDM 2013).

He was the invited speaker for Summer School in Intelligent Systems organized by Quang Binh University, Vietnam and IEEE SMC Chapter in November 2012.

More details about his research can be found in http://se.fsksm.utm.my/aselamat/

Tutorial Talks

Dr. Norsaremah Salleh

Norsaremah Salleh received the PhD degree in Computer Science from the University of Auckland, New Zealand in 2011. Previously she worked nearly five years as an analyst programmer in the manufacturing industry. She is currently an Assistant Professor of Computer Science Department and also Deputy Director of Information Technology Division at International Islamic University Malaysia. Her research interests include the areas of empirical software engineering (SE), evidence-based software engineering research, computer science and software engineering education, and social network sites research. Her work appears in IEEE Transactions on Software Engineering (TSE), Empirical Software Engineering Journal and premier SE conferences including ESEM, ICSE, and CSEET.

More details about her research can be found in https://sites.google.com/site/norsaremahsalleh/

SEPoW 2013 Conference Committees

Patron
Prof. Datuk Dr. Ahmad Yusoff Hassan, Vice Chancellor, UTeM

Advisors
PM Dr. Mohd Khanapi Bin Abd Ghani, Dean of FTMK, UTeM
Dr.Nurul Akmar Emran, Head Department of Software Engineering

Organising committee
Dr. Massila Kamalrudin (Program Chair)
Prof. Dr. Mokhtar Mohd Yusof
PM Norhaziah Saleh
PM Dr. Burhanuddin Bin Mohd Aboobaider
Dr. Azah Kamilah Draman@Muda
Dr. Noraswaliza Abdullah
Dr. Norashikin Ahmad
Dr. Sabrina Ahmad
Abdul Razak Hussain
Amir Syarifuddin Kasim
Azlianor Abdul Aziz
Hidayah Rahmalan
Kasturi Kanchymalay
Khuzaifah Khamis
Mashanum Osman
Maslita Abd Aziz
Mohammad Uzaini Ab Rahim
Mohd Adili Norasikin
Mohd Fadzil Zulkifli
Mohd Nizam Said
Mohd Rif'an Abdul Rahman
Muhammad Suhaizan Sulong
Muhammad Suryanata
Noor Azilah Draman@Muda
Nor Hafizah Hamyadi
Nor Haslinda Ismail
Nor Mas Aina Md Bohari
Norasmara Mahmud
Norhafizan Md Yusof
Nur Ain Zakirah Bahari
Nuridawati Mustafa
Rosleen Abdul Samad
Rosmiza Wahida Abdullah
Safiza Suhana Kamal Baharin
Yahya Ibrahim
Yahaya Rahim
Zahriah Othman
Zarita Mohd Kosnin
Zubaidah Ab Hamid

Requirements Engineering

Aspect Refactoring Workflow to Improve Crosscutting Concern Extraction Effectiveness at Requirement Level **Hema Subramaniam, Hazura Zulzalil, Marzanah A. Zabar and Saadah Hassan**
Characterisation of Healthcare Domain Library to Capture Accurate Requirements of Healthcare Software **Nadiah Daud, Massila Kamalrudin, Safiah Sidek and Rosmiza Wahida Abdullah**
The Use of Essential Use Cases (EUCs) to enhance the Process of Capturing Security Requirements for Accurate Secure Software **Syazwani Yahya, Massila Kamalrudin and Safiah Sidek**
Metaheuristic Retaining New Case in ASEAN Tele-Diagnose Framework **Daniel Hartono Sutanto, Nanna Suryana Herman and Mohd. Khanapi Abd. Ghani**
A Framework for Accessing Patient Health Records Through Multi Channel of Devices **Noorayisahbe Mohd Yaacob, Mohd Khanapi Abd Ghani and Abd. Samad Hasan Basari**
A Framework for Gathering Electronic Medical Record through Telehomecare **Nor Afirdaus Bt. Zainal Abidin and Dr. Mohd Khanapi Abd Ghani**

Aspect Refactoring Workflow to Improve Crosscutting Concern Extraction Effectiveness at Requirement Level

Hema Subramaniam
Universiti Putra Malaysia (UPM),
Universiti Selangor (UNISEL)
Selangor, Malaysia
hema@unisel.edu.my

Hazura Zulzalil, Marzanah A. Zabar, Saadah Hassan
Department of Software Engineering & Information System
Universiti Putra Malaysia (UPM)
Selangor, Malaysia
{hazura;marzanah;saadah}@upm.edu.my

***Abstract* - The maximum benefits offered by modularization have increases the need for refactoring of an existing object oriented programs. However, refactoring needs lot of efforts since object oriented programs contain extensive amount of crosscutting concerns and it's hard to ensure the continuity of crosscutting concern flow throughout the software life cycle. Previous studies have shown that there is no workflow indicating the crosscutting concern flow throughout the software life cycle. Also, there is no standard classification and crosscutting concerns relationship observer that able to comprehensively extract crosscutting concern at requirement level. Lack of absolute crosscutting concern mining method makes refactoring too difficult. Indeed without solving the crosscutting concern conflict at earlier stage, developers will spend countless effort to reform the aspect from extracted crosscutting concern. Therefore, this research is to propose refactoring workflow that will facilitate the crosscutting concern transition continuity throughout software processes stages. It consists of a crosscutting concern mining method that emphasize on crosscutting concerns prioritization at requirement level. The research will begin with classifying the crosscutting concern at requirement level. Next, the information on relationship will be gathered to identify the negative relationship between the crosscutting concerns at requirement level. Once complete with the requirement stage, information on crosscutting concern transition elements into design and development will be facilitated. Finally, the refactoring workflow will be established whereby the element of workflow will be validated with the help of prototype. The main contribution of this study will be an aspect refactoring workflow which can be used by software practitioners in an effort to modularize the existing object oriented programs.**

***Index Terms*—crosscutting concern, requirement refactoring**

I. INTRODUCTION (RESEARCH BACKGROUND)

Modularization is one of the important software quality factors that would enable effective software maintenance [1]. However, majority of object oriented programs facing modularization problem since it contains extensive amount of crosscutting concerns that scattered and tangled in software programs [2]. Crosscutting concern is anything that exists in the program that apart from the business functionality which will perform certain functionality in the program [3]. Since the majority of the existing application is using object oriented approach, restructuring process known as refactoring become essential in increasing the program modularity. In that case, refactoring help in identifying the crosscutting concern that reside in programs and help in isolating it from the main program. Even though refactoring becomes the solution for this yet it does not seems to increase the modularity of a software program. This is due to lack of an aspect mining method which include the continuity of aspect mining from the requirement up to source code, extraction of crosscutting concern based on specific classification and crosscutting concern interdependency study.

Several attempts had been made to extract the crosscutting concern and resolve the negative interdependency between the extracted concerns throughout the software lifecycle especially at requirement stage. In particularly existing aspect mining techniques at requirement level had examined the requirement artifact such as use case model [4], problem domain [5], problem frame [6] and use case diagram [7] in an effort to extract the crosscutting concern. However, all this aspect mining techniques create a tendency to miss out the crosscutting concern which exist in the requirement document since there is no existence of classification for the crosscutting concerns. Moreover the situation become worst when there is a conflict in terms of negative dependency between the crosscutting concerns. Preliminary work on conflict handling which using the Hamilthonian paths [8], EA analyzer [9], syntactic concept and semantic approach[10] has been indicated that it will ensure the effectiveness of the refactoring and help in experiencing the whole benefit of aspect orientation. However these existing techniques only cater for coupling between the aspect and base concern rather than between the crosscutting concerns. Apart from that, the flow of the crosscutting concern which extracted from requirement

up to source code has not been explored as a whole [9], [11]. In particular, concern driven development [12] and UML collaboration [13] indicate the flow of crosscutting concern from design into source code only. Thus it resulted in incomplete crosscutting concern continuity from requirement stage up to source code level.

In that case a workflow which can show the continuous identification of crosscutting concern throughout the lifecycle should be in place. Furthermore, a technique to discover crosscutting concerns at use case scenario which based on classification strategy will help in creating the flow of crosscutting concern from requirement up to source code. Apart from that a relationship between the concerns will be observed in an effort to resolve the negative relationship between the crosscutting concerns which will help in aspect composition. A literature survey needs to be conducted in collecting the crosscutting concern that might exist at the design or source code level. Next, the relationship between the crosscutting concerns needs to be studied and compiled. A systematic literature review (SLR) will be used to systematically map the conceptual workflow elements with the literature data that been collected from the literature survey. This technique known as systematic mapping where each workflow element supported by the relevant literature. Prototype which executes the crosscutting concern extraction method and relationship study will be produced. Effectiveness of each proposed strategy evaluated empirically using experimentation.

As a result, a comprehensive conceptual workflow expected to be produced which may guide the software practitioner on the refactoring processes. Moreover the proposed conceptual workflow starts from requirement stage up to source code level which will help in remaining the continuity of extracted concern from requirement until source code.

II. Problem Statement

Software practitioner need to ensure the software program modularity at high level since it helps in improving the software quality. However, modularization realization becomes problematic with the existence of extensive amount of crosscutting concerns[14]. Crosscutting concern extraction become tedious when go through coding level since there is a tendency to miss out the crosscutting concern in the coding[15] [16]. Apart from that, relationship between the crosscutting concerns also play a main role while modularizing software programs. After crosscutting concerns being extracted from base code, depedency between the crosscutting concerns need to be observed for a succesful creation of aspect. Moreover, contra relationship between the crosscutting concern expected to give huge impact during this process[17][18]. As such, relationship between the crosscutting concern need to be study at the requirement level itself before it move to the next stage of software process. Furhtermore, crosscutting concern which has been extracted out from requirement document need to go through design and source code stage. In most cases crosscutting concern extraction initiated from the source code level, in which it need to deal with a large number of LOC (Line of Code). Untill now there is no compilation of crosscutting concern flow from requirement level up to design and development stage. In fact the crosscutting concern extraction should be initiated from requirement and need to be flow to the next stage[19].

III. Literature Review

A crosscutting concern in requirement document is a concern that is found in one place/multiple places in a requirement document but has broadly scoped influence, either explicit /implicit, on multiple requirement artifacts [20]. Moreover, crosscutting concern which exist in one place known as scattered type meanwhile crosscutting concern that exist on multiple places considered as tangled [21]. There are considerable amount of research has been conducted on the extraction of crosscutting concern at requirement document. First natural language processing (NLP) technique used as technique to identify crosscutting concern at requirement document. This technique utilize speech analysis, word frequency analysis and dominant verb analysis in finding the repeating crosscutting concern at the requirement document. However this technique face difficulties in identifying scattered type of crosscutting concern [4], [11]. Together with that, a technique known as EA miner and Theme/Doc also tend to explore the NLP approach in identifying crosscutting concern at the requirement level [22]. Apart from that, a technique known as use case driven formal concept analysis used in extracting crosscutting concern at use case model. Concept of lattices used in this approach by invoking computational units. Eventhough the technique covering scattered and tangled but it still did not specify the flow of extracted concern further into design and source code [23]. Use case also been explored by previous research in an effort to extract the crosscutting concern. This technique merge the formal concept with software stability model whereby only the main use case used and its non functional requirement being matched while executing the extraction process. In this case the concept of crosscutting concern which can exist anywhere in the system being unattended when the extraction only concentrate on main use cases rather than all use cases [24]. Similarly, crosscutting concern has been map with the core features again by another researcher whereby traceability table has been produced as outcome of the mapping. Although this technique cater for both non-functional and functional requirements yet the flow of mapping result into source code not been specified [6]. Likewise, template has been used by previous researcher in identifying functional crosscutting concern by associating the key association point in use case model. Because of template has been used in this use case driven approach it limit the identification of crosscutting concern which exist in use case scenario [7]. In general, the crosscuting concern identification at requirement level considered unusable if it does not support the continuity among the software life cycle stage [25]. In that case a

crosscutting concern identification technique at requirement level need to strive for continuity throughout the life cycle rather than stop at the requirement stage.

Other than that, there are number of literature by the previous researcher on resolving the contra relationship between the crosscutting concern been discussed. Interaction among the extracted crosscutting concerns analysed by evaluating the relationship between the apsects. Interaction among the aspects and between the base concerns help in revealing the conflict. In such cases the Hamiltonian paths used in order to resolve the conflict among the aspect and base concern. Conflict among the base concern and aspect resolved by specifying the proper joinpoint that can link the base code with the specific concerns [8] [6]. Apart from that, early aspect analyzer also has been used as the conflict resolver whereby each of the requirement has been clasified as 'conflict' or 'harmony'. In this case, bayesian method used in classifying the text since this is a well known technique to classify the text. Eventhough the early aspect analyzer had succesfully identify the conflict, the futher research still needed to specify the solution in order to overcome the identified conflicts [9]. Another technique that identify the crosscutting concerns relationship at the requirement level known as semantic approach. This approach use sytactic concept in an effort to identify the aspect and base concern conflict. Reasoning on the coupling between the base concern and the among the crosscutting concerns has been analyzed using formal method. Although the reasoning has been formulated, yet the relation between the composition towards aspect formulation less explored[10].

In addition to that, early aspect concept indicate the importance of identifying crosscutting concern as early as possible such as at requirement level itself[19]. Eventhough the concern identified earlier, but if the concern has not been channel properly to the next stage it will make the whole refactoring process fail. In that case, it is important to show the crosscutting concern flow throughout the software life cycle[26]. For this reason, a technique known as concern driven development has been proposed in order to show the concern extraction from requirement to design. This technique use the goal oriented concept whereby the goal of each requirement has been analyzed and the use case has been relate into the goal. Once the concern extracted from use case it has been map with the design phase. Though the flow of crosscutting concern been covered from requirement to design, concern driven development fail to relate it into source code[12]. Likewise, a method known as UML collaborations help in transmitting the recurring messages pattern as aspect into UML sequence diagram. Use case diagram used as input artefact at requirement level while sequence diagram used as artefact at design level. Once the crosscutting concern succesfully extracted from use case it has been transimitted into sequence diagram. Although it shows the flow from requirement into design, yet it does not include the flow to the source code level[13]. Thus a comprehensive refactroring workflow which can cover from requirement up to source code should be in place. This will ensure the effectiveness of the aspect refectoring process.

IV. Proposed Solution

Since the crosscutting concern is important and need to be extracted out at the requirement stage itself, crosscutting concern classification is essential. Moreover the provided classification need to be adhere with the source code level classification. This is to ensure the continuity among the crosscutting concern can be maintain throughout the software life cycle. For that reason, the classification considered into two level; based on the generic properties [3][27][28] and core functionality. Each of the generic properties consist of dedicated core functionality (Refer to the figure 1 for the proposed classification tree of crosscutting concern at the requirement document). With this classification methods, crosscutting concern will be comprehensively extracted out from the requirement document. Each of the requirement document classification has been mapped to the object oriented coding level idioms. Natural Language Processing (NLP) technique will be used to read the requirement document based on the classification list.

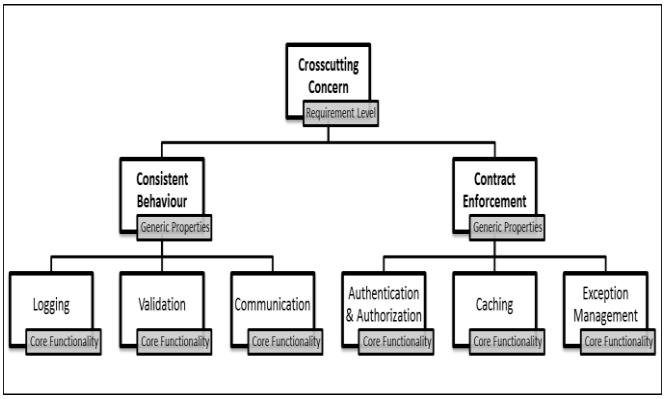

Fig 1. Proposed crosscutting concern classification at requirement level

Crosscutting concern which has been extracted out from the base code need to be analyzed in order to anticipate the relationship between them. Relationship between the crosscutting concern studied based on the interdependecy of one crosscutting concern with another. For instance, security and time responsiveness has the negative relationship among them whereby in order to acheive high security, response time recorded to be slower. This type of negative relationship classify as the 'conflict' in crosscutting concern [8]. For this purpose each crosscutting concern will be assigned with the weightage. Based on the weightage given, the strength level will be calculated and the prioritization will be produced. Based on the prioritization, user able to make decision which crosscutting concern will be emphasize more.

Once the extraction and crosscutting concern complete, the aspect creation process will be countinued by directing the

crosscutting concerns to the next level of the software life cycle. Therefore, a workflow to show the flow of crosscutting concern from one level to another level is conceptualized. Hence, this workflow need to indicate the flow of crosscutting concern from requirement to design and to source code level. Furthermore, it will include the previously proposed classification strategy and interdependency study method at requirement level (refer appendix c). Then, design level expected to have crosscutting concern mapping with the sequence diagram which will help in extracting function which contain the crosscutting concerns. In fact, the method/function that will be extracted out expected to help in isolating concern from source code level. Once the isolation process is complete, the aspect will be successfully created by specifying the pointcut and advice.

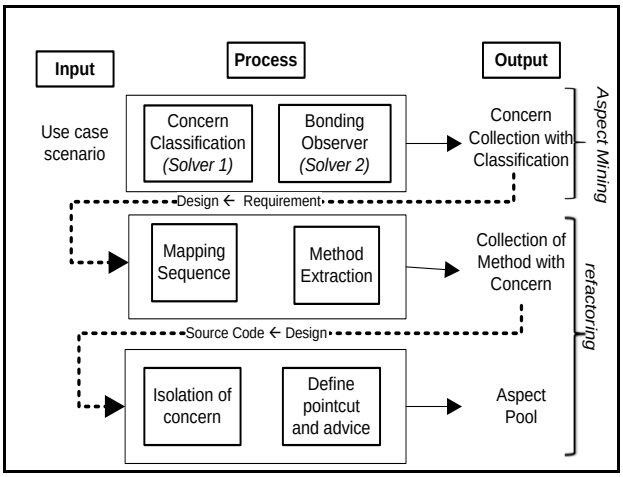

Fig 2. Conceptual Aspect Refactoring Workflow

V. Research Activities

This research divided into five stages (refer fig 3). As a start, literature survey will be conducted by identifying and classifying crosscutting concerns at requirement level. Generally, literature survey is a method to collect data by systematically analyzed the literature. In this case, most relevant paper will be choosen and reading process will be conducted in order to extract the different type of crosscutting concerns that likely to be exist at the requirement documents [29]. For now, the list of crosscutting concerns which listed in terms of generic properties and core functionality (refer fig 1 and Table 1). Healthwatcher requirement document used as tools to extract the crosscutting concerns.

Once the literature survey is complete it will resulted in list of crosscutting concern based on classificationn consideration. Then, the relationship study among the crosscutting concern will be conducted in order to know the interdependency strength level among the crosscutting concerns at the requirement level. The strength level expected to be in the positive relationship or negative relationship. Based on that relationship study, the prioritization among the crosscutting concern will be identified.

Next, to construct the aspect refactoring workflow at the requirement level systematic mapping strategy will be used. Systematic mapping is the method whereby outcome of systematic literature review will be used as qualitative data that needs coding (the codes map to concepts). For instance each element/component in the workflow supported by the literature data and relation between workflow element mapped with literature as well [30].

Following into that, a prototype which consist of crosscutting concern extraction method, interdependency study and aspectization based on the prioritization will be created. Indeed the prototype will be used for the purpose of emperical study in evaluating the effectiveness of the proposed solution. Meanwhile, evaluation of the proposed solution divided into three main categories. The first test is on the crosscutting concern extraction technique effectiveness whereby requirements from requirement document will be entered into the prototype which implement the classification technique. Number of crosscutting concerns extracted out need to compare with other techniques. Then, crosscutting concerns relationship identification technique effectiveness test will be conducted in identifying the depedency between the crosscutting concerns. Finally, crosscutting concerns extraction workflow effectivenss experiment will anticipate the effectiveness of the conceptual workflow. Experiment will be evaluating the number of crosscutting concerns that being extracted using the steps specified in the workflow is it more than the coding level extraction.

VI. Expected Result and Contribution

As a contribution of this research, crosscutting concern mining technique at requirement level based on classification strategy whereby classification consider in terms of generic properties and core functionality will be constructed. Apart from that, crosscutting concern dependency between them and suggestion the appropariate prioritization at the requirement level also expected to produced at the end of the research. Finally, Requirement Level Aspect Refactoring (ReqAR) Conceptual Workflow expected to depict the overall flow of crosscutting concern from requirement level up to coding level.

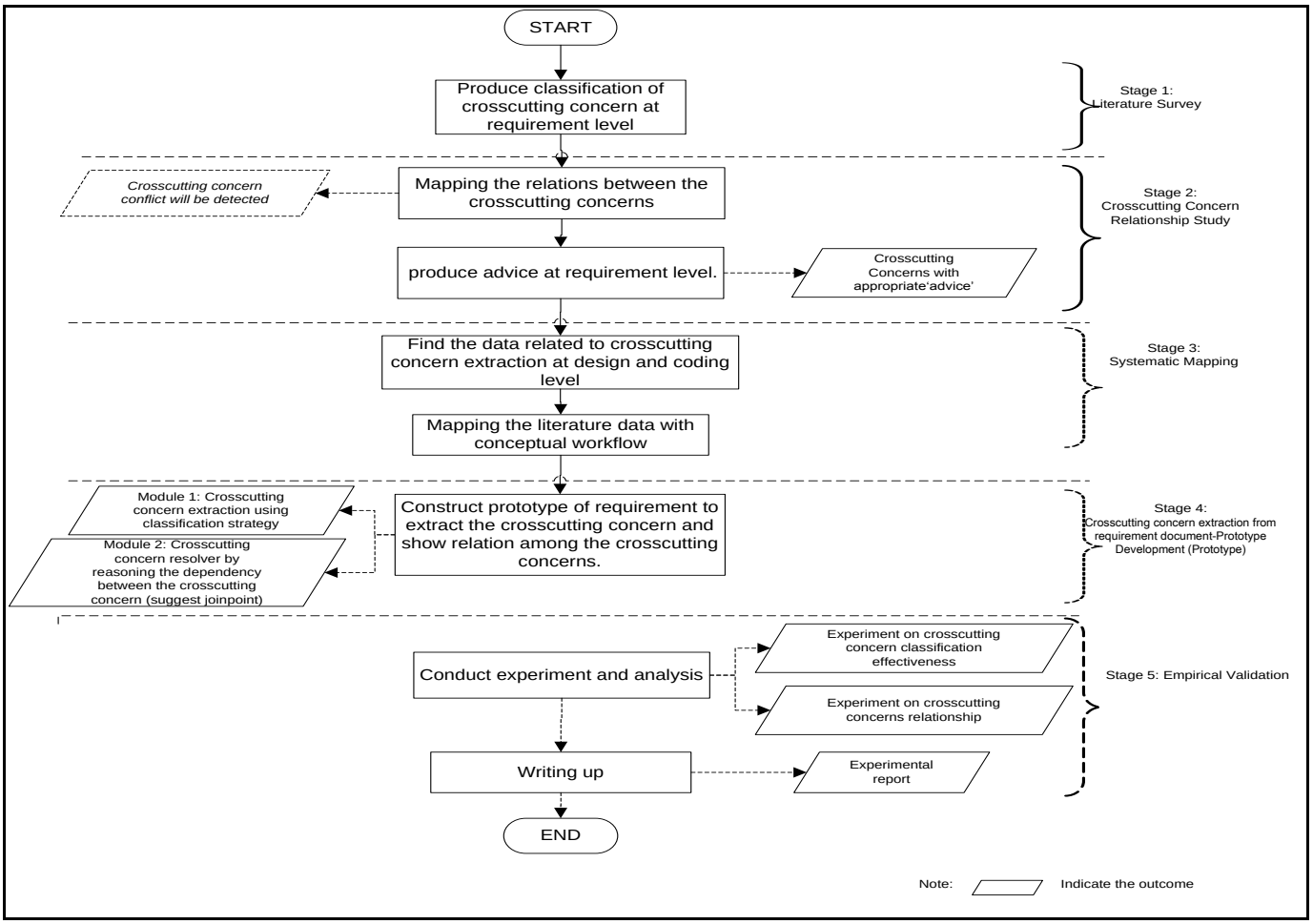

Fig 3. Research activities

TABLE 1. CROSSCUTTING CONCERN CLASSIFICATION (GENERIC PROPERTIES)

Concern Classification consideration: (Generic Properties)	**Object Oriented Coding Level Idiom [1], [2], [3]**	**Aspect Idioms**	**Requirement (use case scenario Idioms) – Contribution of this research**	**Implicit/Explicit Scattered / Tangled**	**Sample requirements (*relate with HEALTH WATCHER)**
Consistent Behaviour	Method calls to the desired functionality	Pointcut and Advice	Main flow event that been repeated for the same use case: (According to the given diagram we know that call a method can be based on the use case scenario that help to call and display the output the desired functionality). Eg: Display	Scattered because dedicated to one use case. And this type of concern mix with the core functionality of the system (sample refer to Page 6/19 whereby 1.1.1 (core functionality) go along with 1.1.2 (concern)	Requirement document: Health Watcher (Page 6/19) Req number: 1.1.2, 1.2.2, 1.3.2
Contract Enforcement	Method calls to method implementing the condition checking	Pointcut and Advice	Alternative flow for one use case similar to another use case. Eg: authentication, authorization and etc.	Tangled because the same kind of checking occurs in most of the use cases. So this indicates the code would be repeated for a multiple methods.	Requirement document: Health Watcher (Page 7/19 & 10/19) Req number: 1.x, 1.x.1 and etc.

REFERENCES

[1] Y. Singh and B. Goel, "Volume 32 Number 4 A Step Towards Software Preventive Maintenance ACM SIGSOFT Software Engineering Notes," vol. 32, no. 4, pp. 1–5, 2007.

[2] Y. Zhang, J. Zhang, Y. Chen, and Q. Wang, "Resolving Synchronization and Analyzing Based on Aspect-Oriented Programming," *2008 International Symposium on Computer Science and Computational Technology*, pp. 34–37, 2008.

[3] M. Marin, L. Moonen, and A. Van Deursen, "Documenting Typical Crosscutting Concerns," *14th Working Conference on Reverse Engineering (WCRE 2007)*, pp. 31–40, Oct. 2007.

[4] B. S. Ali and Z. M. Kasirun, "3CI: A Tool for Crosscutting Concern Identification," *2008 International Conference on Computational Intelligence for Modelling Control & Automation*, pp. 351–355, 2008.

[5] B. S. Ali, "Developing Tool for Crosscutting Concern Identification using NLP 2 . What are Crosscutting Concern and," 2008.

[6] Y. Chernak, "Requirements Composition Table explained," in *20th IEEE International Requirements Engineering Conference (RE)*, 2012, pp. 273–278.

[7] X. Liu, S. Liu, and X. Zheng, "Adapting the NFR Framework to Aspectual Use-Case Driven Approach," in *7th ACIS International Conference on Software Engineering Research, Management and Applications, 2009. SERA '09.*, 2009, pp. 209–214.

[8] A. Boubendir and A. Chaoui, "Towards a Generic Technique for Analysing Interactions Between Aspects at Requirement Phase," pp. 507–512, 2010.

[9] A. Sardinha, R. Chitchyan, N. Weston, P. Greenwood, and A. Rashid, "EA-Analyzer : Automating Conflict Detection in Aspect-Oriented Requirements," no. i, 2009.

[10] N. Weston, R. Chitchyan, and A. Rashid, "A Formal Approach to Semantic Composition of Aspect-Oriented Requirements," *2008 16th IEEE International Requirements Engineering Conference*, pp. 173–182, Sep. 2008.

[11] B. S. Ali and Z. M. Kasirun, "A Review on Approaches for Identifying Crosscutting Concerns," *2008 International Conference on Advanced Computer Theory and Engineering*, pp. 855–859, Dec. 2008.

[12] D. Amyot, S. Leblanc, J. Kealey, and J. Kienzle, "Concern-driven development with jUCMNav," *2012 20th IEEE International Requirements Engineering Conference (RE)*, pp. 319–320, Sep. 2012.

[13] K. Concerns and A. Mining, "Understanding Crosscutting Concerns From Various Perspectives in Software Reverse Engineering," pp. 145–150, 2010.

[14] M. Marin, L. Moonen, and A. Van Deursen, "An Approach to Aspect Refactoring Based on Crosscutting Concern Types," 2005.

[15] B. Mosser, G. Mussbacher, M. Blay-Fornarino, and D. Amyot, "From aspect-oriented requirements models to aspect-oriented business process design models," in *In Proceedings of the tenth international conference on Aspect-oriented software development (AOSD 11)*, 2011, pp. 31–42.

[16] A. Marot and R. Wuyts, "Composability of aspects," in *In Proceedings of the 2008 AOSD workshop on Software engineering properties of languages and aspect technologies(SPLAT '08)*, 2008.

[17] S. Yang, "Extracting traversing features at requirements level from software legacy system," in *Sixth International Conference on Networked Computing and Advanced Information Management (NCM)*, 2010, pp. 139–144.

[18] S.-H. Lee;, B.-H. Cho, and Y.-J. Song, "A Study on Crosscutting Refactoring Using Progam Dependency Relation," in *9th International Conference on Computer and Information Science (ICIS)*, 2010, pp. 684–689.

[19] A. Rashid, P. Sawyer, A. Moreira, and J. Araújo, "Early Aspects : a Model for Aspect-Oriented Requirements Engineering," pp. 1–4, 2002.

[20] J. Lee, K.-H. Hsu, S.-J. Lee, and W.-T. Lee, "Discovering Early Aspects through Goals Interactions," *2012 19th Asia-Pacific Software Engineering Conference*, pp. 97–106, Dec. 2012.

[21] N. Niu, S. Easterbrook, and Y. Yu, "A Taxonomy of Asymmetric Requirements Aspects," pp. 1–18, 2007.

[22] J. Herrera, I. Macia, P. Salas, R. Pinho, R. Vargas, A. Garcia, J. Araujo, and K. Breitman, "Revealing Crosscutting Concerns in Textual Requirements Documents: An Exploratory Study with Industry Systems," *2012 26th Brazilian Symposium on Software Engineering*, no. ii, pp. 111–120, Sep. 2012.

[23] Y. Su and X.-Z. Pan, "Recovering Crosscutting Concern from Legacy Software Based on Use-Cases Driven Formal Concept Analysis," *2010 International Conference on Computational Intelligence and Software Engineering*, pp. 1–4, Dec. 2010.

[24] H. S. Hamza and D. Darwish, "On the Discovery of Candidate Aspects in Software Requirements," in *Sixth International Conference on Information Technology: New Generations, 2009. ITNG '09*, 2009, pp. 819–824.

[25] P. Zhang and Y. Su, "Understanding the aspects from various perspectives in aspects-oriented software reverse engineering," *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, no. Iccasm, pp. V11–311–V11–314, Oct. 2010.

[26] I. Conference, C. Application, S. U. Yang, E. Technology, A. Police, and F. Xi, "Research on Recovering Early Aspects in Aspect-Oriented Software Reserve Engineering," no. Iccasm, pp. 202–206, 2010.

[27] M. Marin, L. Moonen, and A. Van Deursen, "An Integrated Crosscutting Concern Migration Strategy and its Application to JHOTDRAW," *Seventh IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM 2007)*, pp. 101–110, Sep. 2007.

[28] Y. Su, X.-W. Zhou, and M.-Q. Zhang, “Approach on Aspect-Oriented Software Reverse Engineering at Requirements Level,” *2008 International Conference on Computer Science and Software Engineering*, pp. 321–324, 2008.

[29] E. Fielt, “A SYSTEMATIC , TOOL-SUPPORTED METHOD FOR CONDUCTING LITERATURE REVIEWS IN INFORMATION,” 2011.

[30] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, “Systematic literature reviews in software engineering – A systematic literature review,” *Information and Software Technology*, vol. 51, no. 1, pp. 7–15, 2009.

Characterisation of Healthcare Domain Library to Capture Accurate Requirements of Healthcare Software

Nadiah Daud
Faculty of Information & Communication Technology
Universiti Teknikal Malaysia Melaka,
Melaka, Malaysia
nadiahdaud@yahoo.com

Massila Kamalrudin, Safiah Sidek, Rosmiza Wahida Abdullah
Innovating Software System and Service Group
Universiti Teknikal Malaysia Melaka,
Melaka, Malaysia
{massila, safiahsidek, rosmiza} @utem.edu.my

***Abstract*— Capturing accurate requirements is crucial in healthcare domain as it is considered as the most sensitive domain due to its direct impacts on people's well-being and lives. The common practice for capturing requirements is found to be subjected to errors, inconsistencies, redundancy, incompleteness, and omissions. Although frequent face-to-face communication between the requirements engineer and various stakeholders can minimize these problems, it is found to consume too much time. Hence, this research suggests that a healthcare domain library comprising compilation of structural expressions and glossary of functional requirements derived from various stakeholders can assist requirements engineer to capture accurate requirements as well as reduce the time taken. This research involves three stages: analysis, design development as well as testing and evaluation. This paper reports the progress of this research approach which is at its analysis stage.**

***Index Terms*—healthcare software, domain library, accurate requirements**

I. INTRODUCTION

In response to the need to deliver efficient and effective healthcare services together with the advancement in Information technology, there is a growing demand to integrate software information technology in healthcare services. However, the utilization of this software has been very restrictive among the users due to the complexities of usage, lack of integration and lack of cost effectiveness of the system. Further, developing healthcare software is found to be problematic as the healthcare sector is one of the complex domains with many subtleties.

When developing software for healthcare, it is crucial for software developers to capture accurate requirements. This is due to the fact that the healthcare is a sensitive domain since all activities in this domain have direct impacts on people's well-being and lives. Specifically, when developing software, any misinterpretation or incorrect elicitation of requirements may jeopardize one's life. In this regards, software developers (requirements engineers) who are unfamiliar with the myriad kinds of requirements face difficulties to capture accurate requirements as they have to very careful to avoid misinterpretations that may lead to mishaps.

Requirements are a crucial specification before any system/software development starts. At present, in the healthcare business, requirements engineers often use some forms of natural language, written either by the clients or themselves to capture healthcare requirements from the healthcare workers and healthcare vendors. This forms a human centric representation of the requirements accessible to both the engineer and client.

However, due to both the ambiguities and complexities of natural language and the capturing process, these requirements often have inconsistencies, redundancy, incompleteness and omissions. Further, most of the developers in the healthcare industry are facing problems to understand the information provided by the stakeholders [6]. Other common issues are the complexity of healthcare requirements [1] and the conflicting of verbal statements among the stakeholders [6]. Resulting from these problems, developers must have frequent face-to-face communication with different stakeholders to capture accurate requirements. The frequent communication between two parties, however, consumes too much time in confirming the healthcare requirements.

This research suggests that a compilation of functional requirements derived from the various stakeholders can assist requirements engineer to capture accurate requirements. As such, this research aims to examine the following research questions:

i. *What are the characteristics of the functional requirements in the healthcare domain?*
ii. *How accurate is the healthcare domain library in eliciting healthcare requirements?*

This paper describes a research to develop a healthcare domain library that can be used as reference by the healthcare developers and vendors in the elicitation process of requirements. Focusing on the functional requirements, this domain library comprises of categorizations and compilations of the requirements relevant to healthcare domain. We expect that this library will be useful to requirements engineers

because it can help them to elicit accurate requirements by focusing on the relevant information from the proposed library domain. To achieve these aims, the requirements from the healthcare industries and healthcare vendors will be collected. These requirements will be analysed to identify the different categories of structures/patterns and compile a list of common terms in healthcare domain. Then, the domain library will be tested by the experts to determine its accuracy. Finally, this research can provide a positive impact towards the development of effective healthcare software, especially for the requirement engineers to elicit accurate healthcare requirements.

II. RELATED WORKS

There are works done in enhancing the elicitation process of requirements in healthcare. For example, Martin et al. [2] attempted to improve the elicitation process in healthcare requirements. Their work focused on eliciting the requirements for developing a medical device. They used fully interview technique to elicit the requirements. However, the main objective of the elicitation was found to be unsuccessful as they did not effectively establish users' preferences for the design of the device [2].

Another related work conducted by Kaiya et al. [4] proposed a method for enhancing requirements elicitation using domain ontology. They added some values to the ontology value and used web mining as well as lightweight natural language processing techniques. A supporting tool called OREW was also developed to perform this method. Nevertheless, this work did not focus on the requirements in healthcare domain. It also did not develop any domain library to help the elicitation process [4].

In a similar vein, Rumayan at el. [8] developed instruments to elicit all personal values and their relationships to software requirements by interviewing nurses and physicians. Their results were used to develop two small test cases. Yet, their results showed that the instruments are impossible to elicit value-related information and correlate it with the software features.

There have been several other works done to better capture the requirements. For example, Kamalrudin et al. developed a lightweight approach to capture consistent requirements from the client-stakeholders. Here, a library pattern that supports various application domains was developed to store the essential requirements following the EUC methodology. However, the library pattern did not focus on the usage of the healthcare requirements [5],[7].

Works to develop a library or corpus for requirements have also been conducted. Tuzovsky at el. [3] developed a thesaurus for a vocabulary control which demonstrates the relationship between two or more languages. This thesaurus corresponds to common standards and satisfies the requirements of domains knowledge. Yet, this thesaurus is not suitable for the healthcare domain although it is useful for multiple domain knowledge.

Overall, there have been a number of works done to elicit requirements in healthcare domain. However, these works have limitations and difficulties. There has been a study to develop a library or corpus but it does not directly related to the healthcare domain. It is also found that works to develop a domain library specifically for healthcare requirements to enhance the process of eliciting the healthcare requirements are non-existence. Hence, this is the gap that this research aims to address

III. OUR RESEARCH

A. *Our Approach*

Motivated by the related works, we propose a healthcare domain library that can ease the requirements engineers to identify the functional requirements in the system. We propose a domain library tool and a glossary in healthcare software to define any synonym words in the healthcare domain. Figure 1 below shows the outline of the framework of our proposed domain library approach:

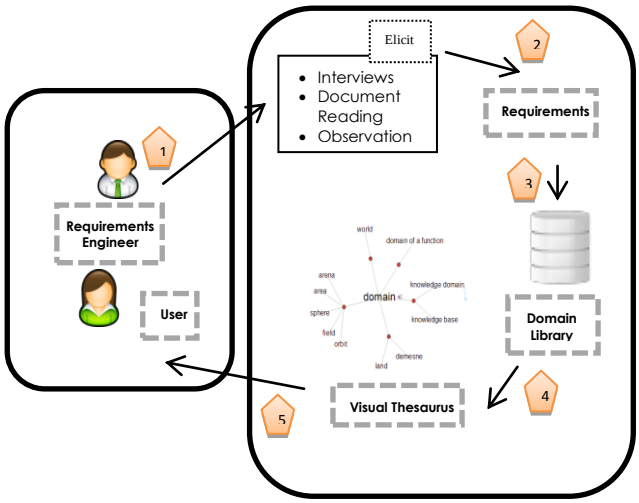

Fig. 1. The Domain Library Approach

The followings are the steps in our approach:

1. Requirements engineers and users collect and elicit the requirements by using the conventional methods.
2. After using the methods for elicitation, they get the healthcare requirements.
3. They access the domain library to identify whether the requirements are functional or non-functional requirements.
4. The domain library tool will identify the correct domain and the types of healthcare requirements and subsequently show the requirements characterisation.
5. Visual thesaurus will show all the synonyms of the words that the requirements engineers and users want to search. When they click to a specific word, the visual thesaurus will expand the spider web to show the synonyms of the word requested.
6. The illustrated synonyms of the requirements from the glossary provide an understanding of the uncommon terms of healthcare elements.

B. Research Methodology

This section outlines the activities conducted in this research. The aim of this research is to develop a new domain library and subsequently, investigate its accuracy in eliciting the healthcare requirements. Figure 2 shows the three main stages of our research methodology in addressing the research questions.

Fig. 2. Research Approach

The method employed in researching and developing the healthcare domain library consists of three main phases: Analysis, Design and Development as well as Testing and Evaluation. As a software engineering research, the method should not be seen as step-by-step but more of an iterative process, where some tasks can be intermeshed with one another depending on their needs. The phases of the research are shown in Figure 2.

1) Analysis

The analysis stage involves two main activities which are the literature review and the collection of healthcare requirements. The purpose of the literature review is to find a research gap in the domain of healthcare. It involves reading existing literature and gathering relevant information of the healthcare requirements elicitation process as well as the background and gaps of current approaches. The second activities involve collecting samples of requirements from selected healthcare industries and vendors to study the types of requirements that are currently used by the users. Based on the collections of requirements, the characteristics of the patterns and lexical items of healthcare requirements will be identified.

2) Design and Development

The second stage involves the design and development of the domain library of healthcare requirements which will be drawn from the analysis of the literature review and collection of healthcare requirements conducted in the earlier stage. At this stage, the functional requirements key textual structure together with a compilation of glossary of terms used in healthcare will be developed. Here, the researchers will design a checklist to evaluate the accuracy of the domain library in eliciting the requirements.

3) Testing and Evaluation

The final stage is the testing and evaluation of the developed healthcare domain library. A qualitative study will be conducted to evaluate the accuracy of the domain library using the designed checklist. Here, expert opinions will be referred to evaluate the accuracy of the requirements in the domain library.

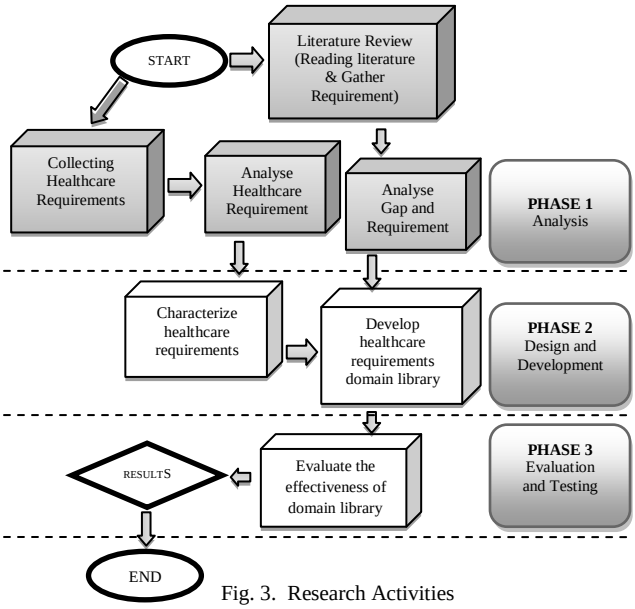

Fig. 3. Research Activities

IV. RESEARCH RESULTS AND PROGRESS

The research is at its first stage of the research process, namely the analysis. As shown in Figure 3, the completed research activities up to now, are colored in grey. Based on the literature review, we found several related works which have been summarized in section II of this paper. We are still conducting the literature review in order to find more related works.

For the collection of healthcare requirements, we have collected 16 requirements from the Internet and software house. Focusing on the functional requirements, we conducted a preliminary analysis using the expert opinions to validate the healthcare requirements. Upon validation from the experts, they were then inserted into Speech Tagging to identify the structural pattern of the healthcare requirements. Based on this analysis, we found interesting common structures of the healthcare requirements.

Figure 3 shows the whole activities of this research and the next stage of this research will concentrate on the Design and Development phase. At this stage, the characterization of the healthcare requirements and the development of a healthcare requirements domain library will be carried out. The highlighted flow of the research indicates the completion of this research.

V. RESEARCH IMPLICATION

This research is believed could provide a domain library that can be used as reference in the elicitation of healthcare. It helps to ensure the correctness of the healthcare software being developed and reduce the time spent to communicate with stakeholders.

VI. Acknowledgement

This research is funded by Universiti Teknikal Malaysia Melaka: Grant PJP: PJP/2013/FTMK(12D)/S01176.

References

[1] R. Razali, P. Najafi, and S.H Mirisaee. "Combining Use Case Diagram and Integrated Definition's IDEFO— A preliminary study." 2nd International Conference on, pp. 231-236. IEEE, 2010.

[2] J.L. Martin, J.C. Daniel, S.P.Morgan, J.A. Crowe, and E. Murphy. "A user-centred approach to requirements elicitation in medical device development: A case study from an industry perspective." Applied ergonomics 43, pp. 184-190, 2012.

[3] A. Tuzovskyy, D. Bubnov, and S. Kozlov. "Multilingual thesauri development and application." In Science and Technology, 2003. 7th Korea-Russia International Symposium, pp. 405-409, 2003.

[4] H. Kaiya, Y. Shimizu, H. Yasui, K. Kaijiri, and M. Saeki. "Enhancing domain knowledge for requirements elicitation with web mining." In Software Engineering Conference (APSEC), 17th Asia Pacific, pp. 3-12. IEEE, 2010.

[5] M. Kamalrudin, J. Hosking, and J. Grundy. "Improving requirements quality using essential use case interaction patterns." In Software Engineering (ICSE), 2011 33rd International Conference on, pp. 531-540. IEEE, 2011.

[6] L.M Marcio. "Requirements engineering in the health care domain." IEEE Joint International Conference, pp. 350-356, 2002.

[7] M. Kamalrudin, J. Grundy, and J. Hosking. "Tool support for essential use cases to better capture software requirements." In Proceedings of the IEEE/ACM international conference on Automated software engineering, pp. 255-264. ACM, 2010.

[8] P. Rumyana, P. Barbara, H.K. Sven, W. Andreas, and W .Thomas. "Investigating the influence of personal values on requirements for health care information systems." Workshop Software Engineering in Health Care, pp. 48-55, 2011.

The Use of Essential Use Cases (EUCs) to enhance the Process of Capturing Security Requirements for Accurate Secure Software

Syazwani Yahya
Faculty of Information & Communication Technology
UniversitiTeknikal Malaysia Melaka,
Melaka, Malaysia,
yiend11@gmail.com

Massila Kamalrudin, Safiah Sidek
Innovative Software System and Services Group,
UniversitiTeknikal Malaysia Melaka,
Melaka, Malaysia,
{massila,safiahsidek}@utem.edu.my

***Abstract*—Complexity in security requirements requires requirements engineers to have a security experience and knowledge in the process of eliciting and analyzing requirements. Further, capturing correct security requirements at the early phase can contribute to the development of secure software. However, requirement engineers lacking sufficient experience in security are facing the risk of over-looking security requirement which might lead to inaccurate, inconsistent and incomplete requirements when capturing the security requirements. This situation may influence the quality of security requirements. In this paper, we describe the overview of our research plan towards developing our new approach that uses essential use cases (EUCs) model to enhance to process of capturing security requirements for accurate secure software.**

***IndexTerms*—Software Engineering, Capturing, Eliciting, Security, Requirements, Secure Software Development, essential use case (EUC)**

I. INTRODUCTION

Software application that contains sensitive data such as financial information, medical records as well as personnel, military and classified information are exposed to daily risks from hackers, viruses, and spywares and even from deliberate theft by disgruntled employees. In this respect, security control of software application is a serious business as it is associated with major cost and serious threats. Further, poor security control can lead to serious damages against the software and the corporate executives who do not ensure high security level. For modern security control of critical software applications, it requires a combination of specialized skills, sophisticated software tools, proper architecture and design, and coding practices and approaches as well as constant vigilance.

The most crucial part in software development is its first phase, which is the requirements elicitation. However, requirement engineers without sufficient experience in security, face the risk of over-looking the security requirement. This might lead to security vulnerabilities that can be exploited in practice. The captured security requirements may also encounter the possibility of being inaccurate, inconsistent and incomplete, resulting in insecure software systems. To overcome these problems, formal or semi-formal models are currently used in the process of capturing the security requirements as they allow for better checking, analysis and structured representations.

This research investigates the usage and effectiveness of a semi-formalised model called Essential Use Cases (EUCs) in capturing security requirements. EUCs are previously identified in several studies to be fruitful in the process of capturing and validating business requirements. It is anticipated that this study will develop a new effective approach to support requirements engineers in capturing security requirements for accurate secure software. To achieve its objectives, this study addresses the following key research questions.

a) Do EUCs help to improve the process of capturing security requirements from the business requirements?

b) What are the appropriate essential requirements patterns that can be used for security requirements?

c) Are the EUCs captured in security requirements different from the EUCs in business requirements?

This paper reports the progress of our research to develop new security library patterns that can be used to capture security requirements from the normal/original business requirements. Specifically, it is organized in six sections. After the introduction section, the second section presents the motivations of this study. This is followed by the third section that summarizes the related works of this research. The fourth section describes our proposed research approach and this is followed by the fifth section that presents our research progress and results. This paper ends with a section on the implication of the research.

II. RESEARCH MOTIVATIONS

Requirements are crucial specification at the early development of any system or software. Thus, knowledge and reasoning of requirements engineering at the early phases are necessary [3], particularly in the process of capturing security requirements. Security requirements can be defined as a system specification of its required security, such as specifications towards the types and levels of protection that are necessary for the data, information, and application of the systems. Examples

of security requirements are the authentication requirements, authorization requirements, intrusion detection requirements, and many others [1]. Capturing security requirements is crucial for the security of a software development. It needs to be accurately defined because poor elicited security requirements can cause a failure to the software development, leading to a costly undertaking [23]. However, the common requirements capturing techniques, such as scenarios and interviews may not contribute to the identification of security requirements as the systematic way may lead to less security exposures [23].

At present, when capturing security requirements from clients, requirements engineers often use some forms of natural language, written either by the clients or themselves. This practice forms a human-centric representation of the requirement accessible to both the engineers and clients. However, due to both the ambiguities and complexities of the natural language and the process of capturing, these requirements often have inconsistencies, redundancy, incompleteness and omissions which can lead to inaccurate security software.

Requirements modeling such as goal-oriented, aspect-driven and system requirements modeling have also been used to capture business requirements [14] and security requirements. UML models are a common way of capturing software requirements including security requirements [13]. Specifically, the use case diagrams have been widely used by developers and requirements engineers to elicit and capture requirements. For UML models, use cases are utilized to capture functional requirements, and as applied in software engineering, it deals with actor/system interaction [14]. However, use cases are found to provide limited analysis to the requirements. To overcome the limitations of use cases, Constantine and Lockwood [14] have developed the Essential Use Cases (EUCs) models. In comparison to a conventional UML use case, an equivalent EUC description is generally shorter and simpler since it only comprises of the essential steps (core requirements) of intrinsic user interest. Kamalrudin et al. [14, 5] have also proven that EUCs is useful to capture and validate the quality of requirements. The use of EUCs in the development process is also advantageous as they fit a ñproblem-oriented rather than ñsolutionï orientedò approach, thus they allow the designers and implementers of the user interface to explore more possibilities. The use of EUCs eventually allows more rapid development since it is no longer necessary to design an actual user interface when using EUCs. In comparison to the conventional UML use cases, the EUCs have the advantages to simplify captured requirements as well as integrate the requirements and design [15]. However, they have not been fully explored in the process of capturing security requirements.

III. Related Works

The process of capturing, elaborating, analysing and documenting security requirements is still not robustly explored [7]. Security requirement is often neglected in business process models as the requirements engineers are not the expert in the security [1][11]. Therefore, previous studies have identified the needs to develop new techniques, methods, tools and notations for practitioners in capturing the security requirements.

There are many frameworks that have been developed for handling security requirements. For example, Haley and his colleagues [1] have developed a framework for security requirements engineering focusing on elicitation and analysis. The framework consists of four stages: Stage 1 - Identify Functional Requirements, Stage2 - Identify Security Goals security, Stage 3 - Identify Security and Stage 4 - Construct Satisfaction Arguments. However, the framework does not give any modeling information and it is a complicated method for software engineers to follow.

Myagmar et al. [18] investigated how threat modelling can be used as foundations for the specification of security requirements and they presented three case studies of threat modelling. Focusing on the requirements engineering process, they proposed that the ability of attackers to misuse or abuse the system can be lessen through appropriate identification of threats and a correct choice of counter measures. The threat modeling process set out by these authors is made up of three high-level steps: Characterizing the system; Identifying assets and access points; and Identifying threats. However, this model has yet to define security requirements and implement the security mechanisms.

Viega [19] proposed a method to build security requirements in a structured manner that is conducive to iterative refinement. Termed this method as the CLASP (Comprehensive, Lightweight Application Security Process), Veiga claimed that if this structured manner is correctly followed to the metrics for evaluation, an obvious improvement of traditional methods without having to consider the security can be achieved. An example of using a simple three-tiered architecture is also provided. The basic idea behind the ways that CLASP handles security requirements are firstly, the performance of a structured walkthrough of resources, and secondly the determination of how they address each core security service throughout the lifetime of that resource. While it is obviously a far more effective methodology than an ad-hoc treatment of security requirements, this methodology is still new and they do not have enough data to identify areas that can be improved.

Shin and Gomaa [20] studied approach models of the evolution of a non-secure application to a secure application in terms of the requirements model and the software architecture model (use case, static and dynamic models) by using distributed software architectures based on components and connectors. They proposed separating the security and application requirements. Security requirements are captured in security use cases and encapsulated in security objects. Additionally, the security services are encapsulated in connectors separately from the components that provide the functionality of the application. However, aspect-oriented modeling for separating business and security concerns are still in need of further investigation to implement security requirements and security software architecture model.

Grady [12] presented FURPS+ that consists of five categories: functional, usability, reliability, performance, and supportability. The symbol + indicates four conditions, namely 1: design constraints, 2: implementation requirements, 3: interface requirements, and 4: physical requirements. This approach classifies the requirements in general and avoids direct classification of the requirements into two public categories. However, it considers the first category as functional and the others as non-functional requirements. Further, it treats security like functional ones by putting it into the first category that consists of system properties, capabilities, and security.

Hussein and Zulkernine [21] proposed a framework for developing components with intrusion detection capabilities. The first stage of this framework is the requirement elicitation, in which developers identify services and intrusions. In this respect, they capture usersô requirements regarding the services and functionalities provided by the components, and identify the unwanted or illegal usage of components by intruders. Intrusion scenarios are elicited through the use of misuse-cases of a UML profile called UMLintr. Yet, their proposed framework still needs an extension scope of UMLintr to other security requirements. UMLintr can be extended by exploring how to specify and handle other security requirements like authentication, authorization, integrity, etc. Their framework is also considered as complex intrusion scenarios.

While Agile Security Requirements Engineering proposes the extension of agile practices to deal with security in an informal way, the communicative and assurance-driven spirit has its own limitation where it only partially supports consistency and does not support correctness and validation checking between the security requirements and business requirements. Many organizations have also realized that security requirements need to be addressed early in the lifecycle process and they feel that attention to this area will pay off in supporting their business goals. For example, Microsoft has a security requirements engineering method which is incorporated into their life-cycle processes. However, at present, there is still no consensus on a single best approach for security requirements engineering [7].In addition, I* frame, which is a modeling and analysis frame for organizational environments and their software systems, is based on the intentionality relations between agents only. Another methods-Tropos, which adopts i*modeling framework, is an agent-oriented software system development methodology. It focuses on describing both organizational environment of a system and the system itself [10]. A secure Tropos framework to model and analyze the security requirements is then built using Secure Tropose methodology [22]. It especially addresses the problem of modeling security requirements through ownership, permission and delegation among the actors or agents involved in the software system. The proposed framework needs are fining to the point where security services and mechanisms are derived and comparable to the standards: ISO-7498-2 and ISO-17799.

As for the usage of EUCs, Kaindl et al. [15] have proposed to use the EUCs for integrating the requirements and design phase due to its usage-centered attributes. Kamalrudin et al. [5] have also used the EUCs patterns to manage the consistency and validation requirements. They also used Essential Use Cases (EUC) patterns to capture multi-lingual requirements in English and Malay languages [16.17]. Here, they develop a tool support to automatically generate and validate EUCs requirements. As a result, they have proven that EUCs is useful to capture requirements, but this research is still open for enhancement, especially in exploring its usage in security requirements.

Overall, there are a number of works done to capture security requirements, but most of the works are still immature and face a lot of limitations and difficulties. It is also found that most of the works use modelling approach to capture security requirements, but none of them uses EUCs model [2].

IV. Our Proposed Research

A. Our Approach

The identified problems related to security requirements such as the lack of tool support and experience among engineers in security requirements, in addition to the lack of integration with other modelling approaches have motivated us to come up with a lightweight approach and tool support. As shown in Figure 1, our proposed research approach is divided into four stages and they are described below.

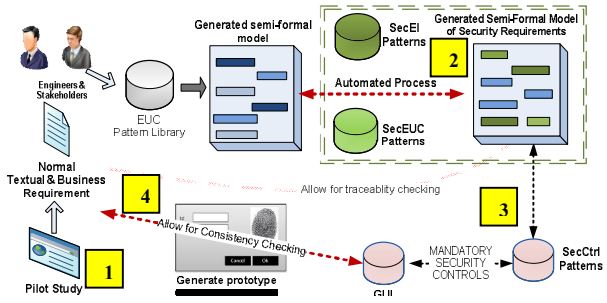

Fig. 1.Our propose approach

1. The first stage involves a pilot study and a generation of semi-formal models. After the pilot study, normal textual and business requirements are collected from selected engineers and stakeholders (clients) particularly from IT/software provider companies. The requirements are then analysed to develop EUC pattern library which is subsequently used to generate semi-formal model. In order to develop the semi-formal model, we have implemented [16][17] works.
2. The second stage is the process of generating semi-formal model of security requirements based on two library security patterns. After the requirements are successfully transformed in a form of semi-formal model, we enable our prototype tools to automate the process of capturing the security requirements. The development of two important library patterns, named as the security essential use cases pattern library (SecEUC) and the security essential interactions

(SecEI) pattern library are the key for the automated process. Our prototype tool allows for traceability checking (represented by the dotted yellow line as shown in Figure 1) between the captured security model and the normal textual requirements. This utility is crucial in order to keep and maintain the consistency of the captured requirements until the end of the process.

3. At the third stage, we exercise the mandatory security controls. This is done by developing a third library called security control patterns (SecCTRL).The mandatory security controls are derived from the captured security requirements.
4. Subsequently, we support for consistency checking between the normal requirements and the captured security requirements using a graphical user interface (GUI) library (represented by the red dotted arrow as shown in Figure1). This library allows users to perform consistency checking provided that they (requirement engineers) accurately capture the required security requirements from the normal business requirements. It also provides recommendations of suitable GUI, whereby towards the end of the process it enables users to check whether the clientôs request is fulfilled.

We hypothesize that our approach can provide better support for decision making and precise communications for users and developers to work with informal and semi-formal security requirements as well as to keep them consistent.

We have developed an automated prototype tool that provides authoring facilities for security requirements and for checking the consistency of these requirements. This tool assists requirement engineers and business analysts to better capture the security requirements and make verification whether their requirements (in written form or collected from natural language) are consistent with the captured security requirements and the design representations.

B. Research Methodology

This part summarizes the activities conducted in this research. The aim of this research is to develop a new security library patterns and then, investigate its accuracy towards capturing the security requirements.

Our research methodology is shown in Figure 2. The method employed in this research consists of three main stages; Analysis, Design and Development as well as Testing and Evaluation. All of the three stages are planned in the direction towards achieving our research objectives. The method is iterative where some tasks can be performed intermeshing with one another depending on their needs.

Fig. 2. Research Approach

Our research activities are shown in Figure 3 below. The boxes in grey indicate the process that we have carried out, while the bottom end, consisting of shapes in red depicts the research activities that have yet to be completed.

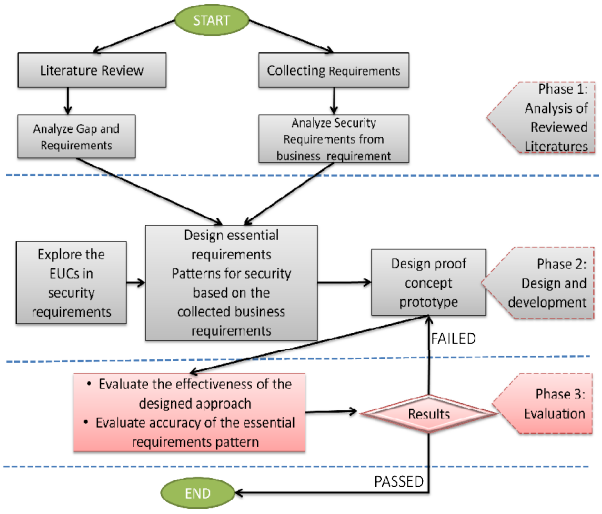

Fig.3. Research Activities

1) Phase 1: Analysis of Reviewed Literatures

A systematic review has been conducted. It began with reading and gathering relevant information on security requirement capturing process as well as understanding research background and identifying gaps of current research approaches. Examples of requirements were also collected from selected software companies in Malaysia to study the type of requirements used and problems faced in the industry. Additionally, essential interactions for the security requirements were collected. Here, the characteristics of the intended security requirements were identified, leading to the solution for the problems found in the research gaps.

2) Phase 2: Design and Development

Here, the findings in the earlier phase were reviewed and the EUCs were explored to capture the security requirements from the collected business requirements. Then, the patterns for the library of essential requirements were defined, and a prototype tool has been developed to realize the approach. At this stage, the design of questionnaires for a qualitative study is discussed for the purpose of investigating the effectiveness of the EUCs usage in capturing security requirements. Prior to that, a study on the current approach in capturing security requirements has been conducted.

3) Phase 3: Testing and evaluation

A qualitative and quantitative study will be conducted to evaluate the effectiveness of the usage of EUCs together with the essential requirements patterns to capture security requirements. The library patterns for essential requirements will also be evaluated for its efficacy and scalability across different domains of business requirements. Finally, the results gained from the latter study will be compared to the results

gained from the earlier study of the approach used to capture the security requirements.

V. Research Progress And Results

Based on our exploration and characterization of essential interaction, a security requirements patterns have been developed. Two library patterns, called as the security essentials use cases (SecEUC) and the security essential interactions (SecEI) have been developed. The design of the library patterns have been validated by a group of security experts from IBM India, Austria and France. After the validation of the library patterns, improved versions of the library will be developed. At present, the first versions of the library are released and applied to the developed prototype tools.

An initial prototype tool has been developed too. As shown in Figure 4, this tool is an enhanced version from MEReq [17]. We conducted a preliminary testing on our developed prototype. For this purpose, 40 students were recruited to carry out the test. The selected students were those who have sufficient level of understanding of the concept and methodology of essential use cases. Overall, the usability results show that our prototype tools are useful, user-friendly and simple (easily understandable). However, we believe that there are still rooms for improvement on this prototype tool. In future, the tools will evolve in consistent with the evolution of the security library patterns.

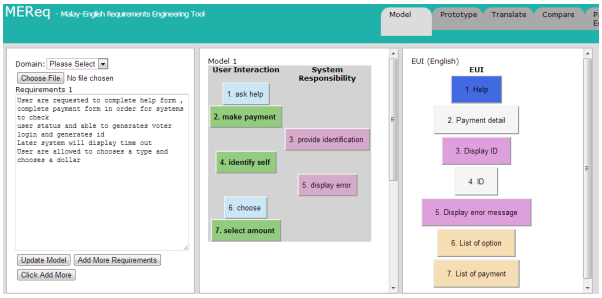

Fig. 4. The initial prototype tools

VI. Research Implication

Requirements Engineering research has been established for many years and requirements are the core part of any software or system development. It is found that requirements written in natural language are error-prone, incorrect and incomplete. Due to these deficiencies, semi-formal models such as Essential Use Cases (EUCs) have been developed to improve the quality of the requirements as well as to minimize the time taken and to ease the requirements capturing process. However, almost none of the work explores the usage of EUCs for capturing security requirements. This study attempts to investigate the usage of EUCs to improve the process of capturing the security requirements. The study also explores the development of appropriate library patterns of essential requirements for security requirements.

EUC is a semi-formal representation which is developed to capture and improve the results of the captured requirements. Thus, we believe this study will provide positive results to the requirements engineering study, particularly in security requirements capturing process which will lead to better quality security requirements.

The library patterns of security essential requirements will also benefit the quality of security requirements. It can be used to check for the consistency between the security requirements and the business requirements. Results of this research will give positive impacts to the software business as it contributes to the improvement of the quality of the software security requirements of a software projects, hence leading to the software industry to produce sustainable software that are correct, consistent and secure.

The outcomes of this research are believed will be a novel finding which contributes to wider requirements engineering study and practice. Specifically, this study contributes to improving the process of capturing the security requirements for developing accurate secure software. It is also believed that the finding fits with the demand of software industry in developing secure software for many application domains such as critical system, mobile application, and finance application.

Acknowledgment

The authors also would like to acknowledge Universiti Teknikal Malaysia Melaka and the Ministry of Education Malaysia of the scholarship Mybrain15. We also would like to thank the funding of this ERGS research grant: **ERGS/2013/FTMK/ICT01/UTEM/E00026** for funding this research. We also would like to thank Noorezam Yusop for his help in the development of the tool support.

References

[1] C. B. Haley, R. Laney, J. D. Moffett, and B. Nuseibeh, ñSecurity Requirements Engineering: A Framework for Representation and Analysis,ò IEEE Transactions on Software Engineering, vol. 34, no. 1, pp. 1ï 48, 2008.

[2] S.Yahya, M.Kamalrudin, S.Sidek, ñA Review on Tool Supports for Security Requirements Engineeringò Manuscript submitted for ICOS IEEE Malaysia Computer Chapter, 2013.

[3] P. Salini and S. Kanmani, ñA Model based Security Requirements Engineering Framework applied for Online Trading System,ò IEEE Conference Publications, pp. 1195ï 1202, 2011.

[4] D. Mellado, E. Fernández-Medina, and M. Piattini, ñA common criteria based security requirements engineering process for the development of secure information systems,ò Computer Standards & Interfaces, vol. 29, no. 2, pp.244ï 253, Feb. 2007.

[5] M.Kamalrudin J. Hosking, John Grundy, "Improving Requirements Quality using Essential Use Case Interaction Patterns," in ICSEô11, Honolulu, Hawaii, USA, 2011.

[6] Rodríguez, B. Bio, E. Fernández-medina, M. Piattini, and C. Real, ñSecurity Requirement with a UML 2 . 0 Profile,ò in Availability, Reliability and Security, 2006. ARES 2006. The First International Conference on, 2006.

[7] P. Salini and S. Kanmani,òSurvey and analysis on Security Requirements Engineering. Computers & Electrical Engineeringò Available at: http://linkinghub.elsevier.com/retrieve/pii/S0045790612001644 [Accessed October 1, 2012].

[8] T. R. Farkhani and M. R. Razzazi, ñExamination and Classification of Security Requirements of Software Systems, ñin Information and Communication Technologies, 2006. ICTTA õ06. 2nd, 2006, pp. 2ï 7.

[9] Rodríguez, A., Fernández-medina, E., & Piattini, M. (2006). Capturing Security Requirements in Business Processes Through a UML 2. 0 Activity Diagrams Profile, 32ï 42.

[10] Du, J., Yang, Y., & Wang, Q. (2009). An Analysis for Understanding Software Security Requirement Methodologies. 2009, Third IEEE International Conference on Secure Software Integration and Reliability Improvement, 141ï 149.

[11] Liu, L., Li, T., & Peng, F. (2010). Why Requirements Engineering Fails: A Survey Report from China. 2010 18th IEEE International Requirements Engineering Conference, 317ï 322

[12] Grady, Robert B. Practical Software Metrics for Project Management and Process Improvement, Englewood Cliffs, New Jersey: Prentice-Hall, 1992.

[13] L. L. Constantine, "Essential modeling: use cases for user interfaces," interactions, vol. 2, pp. 34-46, 1995.

[14] L. L.Constantine, A. D. L. Lockwood, "Structure and style in use cases for user interface design," in Object modeling and user interface design: designing interactive systems, ed: Addison-Wesley Longman Publishing Co., Inc., 2001, pp. 245-279.

[15] H. Kaindl, L. Constantine, O. Pastor, A. Sutcliffe, and D. Zowghi, "How to Combine Requirements Engineering and Interaction Design?ò 1 6th IEEE International Requirements Engineering, 2008. RE '08.Barcelona, Catalunya, Spain, 2008, pp. 299-301.

[16] M. Kamalrudin, J.Grundy, J.Hosking,"MaramaAI: Tool Support for Capturing and Managing Consistency of Multilingual Requirements", in Proc. 27th Automated Software Engineering 2012 (ASE 2012), ACM: Essen, Germany.

[17] M.Kamalrudin, J.Grundy, J.Hosking, "Supporting Requirements Modelling in the Malay Language using Essential Use Cases", in Proc. IEEE Symposium on Visual Languages and Human-Centric Computing, IEEE, 2012.Innsbruck,Austria

[18] Myagmar, ñThreat Modeling as a Basis for Security Requirements,ò in Storage SS õ05: Proceedings of the 2005 ACM workshop on Storage security and survivability, 2005, pp. 94ï 102.

[19] J. Viega, ñBuilding Security Requirements with CLASP,ò in SESS õ05 Proceedings of the 2005 workshop on Software engineering for secure systemsð building trustworthy applications Pages 1 - 7, 2010, pp. 1ï 7.

[20] M. E. Shin and H. Gomaa, ñSoftware requirements and architecture modeling for evolving non-secure applications into secure applications,ò Special Issue on the 5th International Workshop on System/Software Architectures (IWSSAõ06) Science of Computer Programming, vol. 66, no. 1, pp. 60ï 70, Apr. 2007.

[21] M. Hussein and M. Zulkernine, ñIntrusion detection aware component-based systems: A specification-based framework,ò Journal of Systems and Software, vol. 80, no. 5, pp. 700ï 710, 2007.

[22] P. Giorgini, F. Massacci, J. Mylopoulos, and N. Zannone (2005) ñModeling security requirements through ownership, permission and delegation,ò 13th IEEE International Conference on Requirements Engineering (REõ05), pp. 167ï 176.

[23] Software Assurance Pocket Guide Series: Development, ñRequirements Analysis for Secure Softwareò, Volume IV, Version 2.1, May18, 2012, https://buildsecurityin.us-cert.gov/swa.

Metaheuristic Retaining New Case in ASEAN Tele-Diagnose Framework

Daniel Hartono Sutanto, Nanna Suryana Herman, Mohd. Khanapi Abd. Ghani
Faculty of Information & Communication Technology
Universiti Teknikal Malaysia Melaka
Melaka, Malaysia
P031220009@student.utem.edu.my;{nsuryana;khanapi}@utem.edu.my

***Abstract*—In 2015, ASEAN Free Trade Area (AFTA) has focus in Health Care Community; the cross border trade needs Tele-Diagnose framework to exchange medical information between medical tourism or patient mobility such as Malaysia, Indonesia, and Thailand. This framework used Case-Based Reasoning (CBR) in order to diagnose biggest prevalence chronic disease like diabetes mellitus, so on in Indonesia 58 % from prevalence diabetes is undiagnosed and only 0.7% of people living with diabetes has achieved treatment targets. Meanwhile, CBR cycles have obstacle equally fault diagnose and knowledge elicitation. The performance parameter is size, time, and accuracy in retaining new case. Meanwhile, the problem is less accuracy rate of classifiers consequent of impaired index similar case in CBR phases due noisy in dataset. Hence, we purpose metaheuristic technique case based reasoning in diagnose chronic disease to optimize its framework which is using ICD-10 (International Classification Disease).**

***Index Terms*—AFTA, ASEAN Health Care, Tele-Diagnose, Chronic Disease, Case-Based Reasoning, ICD-10.**

I. INTRODUCTION

ASEAN moves towards the ASEAN Economic Community (AEC) in 2015, which is AEC has goal in 2015: "no restriction to ASEAN services suppliers in providing services and in establishing companies across national borders within the region, subject to domestic regulations" [1]. Meanwhile, healthcare remains a high priority sector to be integrated across the region and vital to successful integration of the sector is the harmonization and efficiency of its regulations [2]. ASEAN has agreement to regulate ASEAN Free Trade Area (AFTA) and it is resulting ASEAN Framework Agreement on Services (AFAS) [3]. In ASEAN Strategic Framework on Health Development has objective to ensure access to adequate and affordable healthcare, medical services and medicine, and promote healthy lifestyles for the peoples of ASEAN. Trade in Health Care Services, two types: (1) medical professional services and (2) health services by hospitals and other types of facilities e.g. laboratory, ambulance. There are 4 modes of service trade in professional and hospital service such as (i) cross border trade e.g. telemedicine or remote diagnostic service; (ii) consumption abroad e.g. medical tourists; (iii) commercial presence e.g. entry and operation of foreign operators; and (iv) presence of natural persons e.g. foreign doctors and nurses [4]. Hence, need of ASEAN health care community development to provide semantic exchange medical information or tele-diagnose with ICD-10 (International Classification Disease) which is issued by WHO as reliable and proven world standard classification disease code which is diabetes mellitus code (E10–E14) should not be accepted as "due to" any other disease except for conditions causing damage to the pancreas, and for a list of the conditions that can cause diabetes [5]. Code to unspecified diabetes mellitus with renal complications (E14.2 and N08.3). Code to insulin-dependent diabetes mellitus with multiple complications (E10.7). Codes E10.2 and N08.3 (Insulin-dependent diabetes with nephropathy), E10.5 (Insulin-dependent diabetes with peripheral circulatory complications) and E10.3 and H28.0 (Insulin-dependent diabetes with cataract) may be added as optional additional codes to identify the individual complications. The proper diagnose chronic disease can be adjust the best treatment especially for prevalence diabetes to emphasize fault diagnose.

The lives of many people in the world have been somber and curtail by chronic disease. Chronic disease is defined as disease that persists over a long condition which progress slowly and generally it can be controlled but not cure. In Malaysia, have 2 epidemic disease likely Diabetes Mellitus [6]. Total world prevalence of diabetes is around 371 million and ASEAN has 3 biggest countries who is prevalence of diabetes such as Indonesia 7,4 million, Thailand 3,4 million, Malaysia 2 million [7]. Nevertheless in Indonesia 58 % from prevalence diabetes is undiagnosed and only 0.7% of people living with diabetes has achieved treatment targets [8]. WHO predicts prevalence of diabetes in 2030 around 21,257,000 Indonesian and 2,479,000 Malaysian peoples [9]. Currently the predicted prevalence rate is around 15% to 22% of the adult population with more than 4 million Malaysians suffering from the disease, already surpassing the W.H.O.s prediction of 807,000 for the year 2025 [10].

In Southeast Asia, the health sector is expanding rapidly, attributable to rapid growth of the private sector and notably, medical tourism or patient mobility, which is emerging as a lucrative business opportunity. Countries here are capitalizing on their popularity as tourist destinations by combining high quality medical services at competitive prices with tourist packages. Some countries are establishing comparative advantages in service provision based on their health system's organizational structure describe on Table 1 [11].

Recently, diagnosis chronic disease can be applied in Expert System using Case-Based Reasoning [12]. Watson stated that CBR is a case-based reasoning solves problems by using or adapting solutions to old problems [13]. CBR is able to utilize the specific knowledge of previously experienced, concrete problem situations (cases). A new problem is solved by finding a similar past case, and reusing it in the new problem situation. A second important difference is that CBR has also been an approach to incremental, sustained learning, since a new experience is retained each time a problem has been solved, making it immediately available for future problems [14]. At Figure 1. the highest level of classification and generalist, a general CBR cycle may be described by the following: (i) RETRIEVE the most similar case or cases; (ii) REUSE the information and knowledge in that case to solve the problem; (iii) REVISE the proposed solution; (iv) RETAIN the parts of this experience likely to be useful for future problem solving. The systems of CBR have been used for diverse purposes likely classification, diagnostic, planning, and tutoring in the field of medical [15].

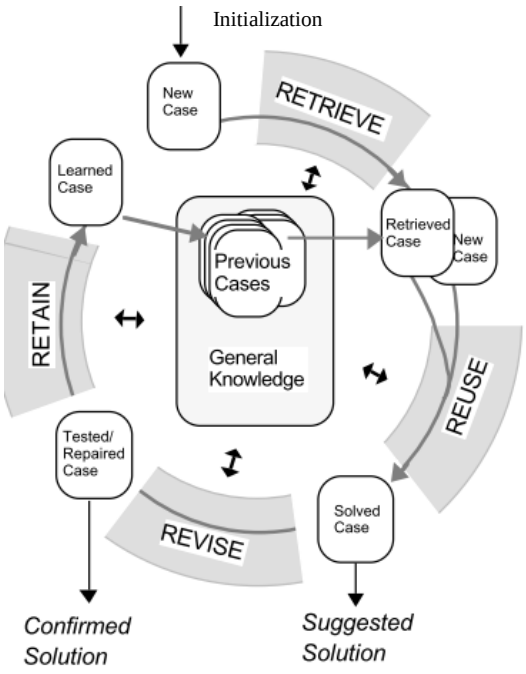

Fig 1. The CBR Cycle [14]

In some domains like fault diagnosis, the cases are heterogeneous (different cases have different features). How to deal with this heterogeneity for CBR different steps is still an open problem. In the retention part of the CBR cycle generally deletion, addition and editing of the case base is still an open problem. Retention estimation of a case base and identifying missing or redundant cases are important areas of research. Maintenance of the case base is done to decrease the utility problem, but itself adds a burden to the system. It has to be examined if the maintenance always increases the performance of the overall CBR system while decreasing the utility problem of the system.

In clinical decision support, error classification divides in three cases. Errors in classification may be in one of three cases [16].Type-I error (false-positive) occurs when the system erroneously classifies a case as positive when in fact, it is not. Type-II error (false-negative), on the other hand, describes missing an existent positive. For example, a patient who is affected by a certain disease is diagnosed as disease-free. Usually, improving one type of error comes at the expense of the other [17]. In practice, the significance of these error costs varies with the application itself. For example, in life threatening medical conditions that require prompt intervention, missing a diagnosis (a false-negative) might result in a waste of time that may lead to life losses or at least cases that are not treated properly. On the other hand, a false-positive may result in unnecessary procedures, anxiety and financial costs [16]. The last type of error is the unclassifiable error. In this case, the system is unable to classify a case, possibly due to the lack of historic data.

II. Related Work

On recent paper, Jha found that CBR can be applied in Diabetes Detection and Care uses 4 Diabetes Support System TM (4DSS). The Intelligence System has aimed to: (a) automatically Detect problems in BG control; (b) propose solutions to detect problems; and (c) remember which solutions are effective or ineffective for individual patients [18]. Previous research on case liver disease, Lin used model Neural Network – Back Propagation Network and Case-Based Reasoning to determine the type of liver disease. The best accuracy rate is 98.04% while its using learning rate 0,1, momentum rate 39-9-1 and the architecture 39-9-1 [19]. Other research in cardiovascular disease, Hsu purposed hypertension detection used Genetic Algorithm and Case-Based Reasoning. The dataset has false positive 172 subjects whereas 44 false negative cases and the accurate rate close to 94.8% [20]. A CBR system has been developed to effectively screen hyperactive attention deficit disorder in both children and adults [21]. In another study, historical cases have been used to make clinical recommendation [22]. Data-mining techniques, such as rule mining combined with medical cases, have been used to develop a CBR system to evaluate the prognostic issues of chronic diseases [12]. Missing values in a CBR system are a serious issue. It has been effectively handled by filter imputation and the performance of the tool has been tested in the treatment of prostate cancer [23]. In another study, the shortfall of CBR system, that is, its inability in assigning optimum weights to the attributes has been tackled by combining it with the genetic algorithm [24]. In PMS case

using Case-Based Reasoning with K-NN algorithm to diagnose which is proper criteria based on the American Psychiatric Association [25].

Metaheuristics can be effective and efficient tools. They are well known for solving various optimization problems, for their adaptation ability and for their ability to produce good solutions in reasonable time and memory requirements [16]. There are many potential benefits of variable and feature selection: facilitating data visualization and data understanding, reducing the measurement and storage requirements, reducing training and utilization times, defying the curse of dimensionality to improve prediction performance [26].

III. Research Methodology

This part the purpose research methodology such as data collection and experiment. Data collection is consist of interview, observation and survey. In this case, it will collecting the need and problem private hospital while do treatment to patient mobility of chronic disease. The experiment needs requirement analysis which framework is needed to design and develop. This framework diagnose chronic disease to classify diabetes mellitus based on ICD-10 code. Finally, evaluation and testing will verify the framework has been developed.

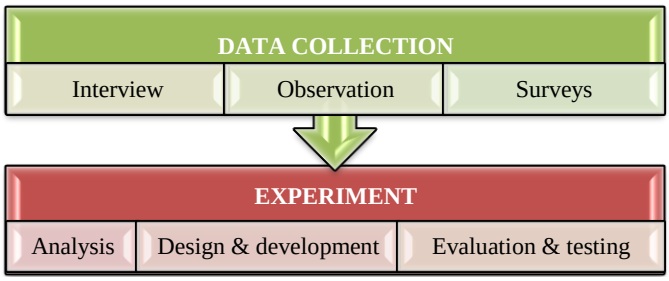

Fig 2. Research Methodology

IV. The Purpose Framework

In this section will explain what the tele-diagnose framework will develop. ASEAN patient mobility through cross border trade, which is exchange medical record include diagnose, report, and patient data. It will be transferred to all members ASEAN medical tourism countries, and it will be standardized using ICD-10. Tele-diagnose framework provides medical record, diagnose, document and patient's application.

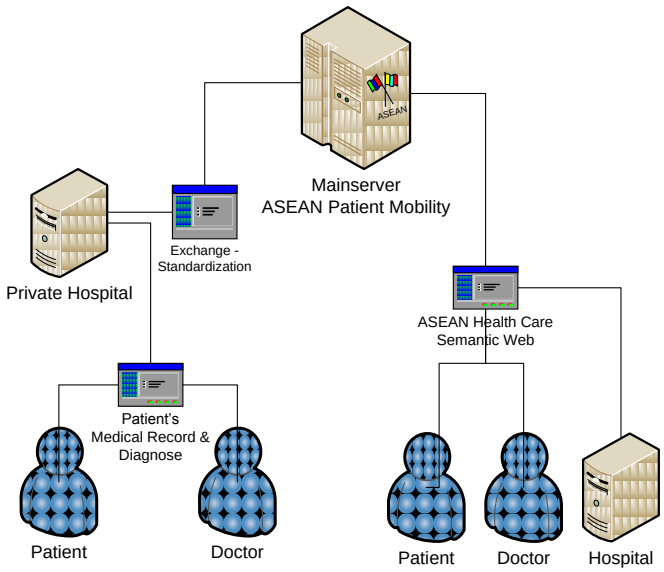

Fig 3. ASEAN Tele-Diagnose Framework

V. Expected Result

The system will be developed to fulfill the objective, hence there are points of result:

1. Medical record can be translate and transferred to all member ASEAN Health Care Community using ICD-10 standard.
2. Integrated patient mobility using semantic web through all policy about ASEAN Framework Agreement on Services.
3. Optimize case based reasoning in retaining cycle to produce new case with minimize fault diagnose.
4. In knowledge elicitation, metaheuristic CBR technique can be presented a new solution with high accuracy and scalability.

VI. Research Contribution

This research is developing framework in diagnose chronic disease using case-based reasoning and the research work is contribute:

1. To optimize the size of the case base and reduce the retrieval time using case based reasoning to get high scalability.
2. To eliminate useless cases and irrelevant and redundant instances that render the case base inconsistent in order to increase the accuracy of the CBR system in diagnose chronic disease.
3. To develop ASEAN Health Care Community in purpose to fulfill ASEAN patient mobility in tele-diagnose chronic disease.

VII. Conclusion And Future Work

First glance, CBR can be applied to diagnose medical domain, nevertheless it has obstacle to define new case with high accuracy and scalability. ASEAN tele-diagnose framework has to provide precise and secure exchange medical information between patient, doctor, hospital. So on, there are so many opportunity to develop in diagnose framework especially in patient mobility.

Acknowledgment

This research paper is made possible through the help and support from everyone, including: parents, supervisor, family, friends, ASEAN countries, hospital, and in essence, all sentient beings.

References

[1] ASEAN Secretariat, *Roadmap for an ASEAN Community 2009-2015*. ASEAN Secretariat, 2009, pp. 1–112.

[2] J. Ke, "ASEAN economic and political integration: legal and institutional enquires into prospects for deep integration," 2013.

[3] D. V. T. Thanh and P. Bartlett, "Ten Years of ASEAN Framework Agreement on Services (AFAS): An Assessment," 2006.

[4] O. Cattaneo, "Trade in Health Services What s in it for Developing Countries ?" 2009.

[5] WHO, "ICD 10 - International Statistical Classification of Diseases and Related Health Problems," vol. 2. pp. 1–195, 2010.

[6] W. Nazaimoon, W. Mohamud, K. I. Musa, A. Sharifuddin, K. Frcp, A. Ismail, I. S. Ismail, K. A. Kadir, N. Azmi, K. Mmed, N. Azwany, Y. Mcm, N. Mustafa, O. Ali, S. Harnida, and I. Mrcp, "Short Communication Prevalence of overweight and obesity among adult Malaysians : an update," vol. 20, no. May 2010, pp. 35–41, 2011.

[7] S. Asia, "IDF DIABETES ATLAS 5," pp. 3–4, 2012.

[8] N. Nordisk, "Changing Diabetes in Indonesia," in in *The Blueprint for Change Programme*, no. March, 2013, pp. 1–28.

[9] World Health Organization, "Country and regional data on diabetes," *WHO*, 2013. .

[10] NADI, "Diabetes 4 Laypersons Healthcare Fair & Exhibition and Launching of First," *Bernama Library and Infolink Services*, 2011. [Online]. Available: http://blis.bernama.com/getArticle.do?id=68735&tid=95&cid=2.

[11] N. S. Pocock and K. H. Phua, "Medical tourism and policy implications for health systems: a conceptual framework from a comparative study of Thailand, Singapore and Malaysia.," *Global. Health*, vol. 7, no. table 2, p. 12, Jan. 2011.

[12] M.-J. Huang, M.-Y. Chen, and S.-C. Lee, "Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis," *Expert Syst. Appl.*, vol. 32, no. 3, pp. 856–867, Apr. 2007.

[13] I. Watson, "Case-based reasoning is a methodology not a technology," *Knowledge-Based Syst.*, vol. 12, no. December 1998, pp. 303–308, 1999.

[14] A. Aamodt, "Case-Based Reasoning : Foundational Issues , Methodological Variations , and System Approaches," vol. 7, no. 1, pp. 39–59, 1994.

[15] M. Nilsson and M. Sollenborn, "Advancements and Trends in Medical Case-Based Reasoning : An Overview of Systems and System Development," 2004.

[16] S. Al-muhaideb, M. El, and B. Menai, "Hybrid Metaheuristics," vol. 434, pp. 187–217, 2013.

[17] P. A. Castillo, M. G. Arenas, J. J. Merelo, V. M. Rivas, and G. Romero, "Multiobjective Optimization of Ensembles of Multilayer Perceptrons for Pattern Classification," pp. 453–462, 2006.

[18] M. K. Jha, D. Pakhira, and B. Chakraborty, "Diabetes Detection and Care Applying CBR Techniques," no. 6, pp. 132–137, 2013.

[19] R.-H. Lin and C.-L. Chuang, "A hybrid diagnosis model for determining the types of the liver disease.," *Comput. Biol. Med.*, vol. 40, no. 7, pp. 665–70, Jul. 2010.

[20] K.-H. Hsu, C. Chiu, N.-H. Chiu, P.-C. Lee, W.-K. Chiu, T.-H. Liu, and C.-J. Hwang, "A case-based classifier for hypertension detection," *Knowledge-Based Syst.*, vol. 24, no. 1, pp. 33–39, Feb. 2011.

[21] D. Brien, J. Glasgow, and D. Munoz, *The Application of a Case-Based Reasoning System to Attention-Deficit Hyperactivity Disorder*, 6th Intern. Chicago, IL, USA: Springer Berlin Heidelberg, 2005, pp. 122–136.

[22] D. Doyle, P. Cunningham, and P. Walsh, "An Evaluation of The Usefulness of Explanation in A Case Based Reasoning System for Decision Support in Bronchiolitis Treatment," *Comput. Intell.*, vol. 22, no. 3–4, pp. 269–281, 2006.

[23] R. Jagannathan and S. Petrovic, "Dealing with missing values in a clinical case-based reasoning system," in *2009 2nd IEEE International Conference on Computer Science and Information Technology*, 2009, pp. 120–124.

[24] H. Ahn and K. Kim, "Global optimization of case-based reasoning for breast cytology diagnosis," *Expert Syst. Appl.*, vol. 36, no. 1, pp. 724–734, Jan. 2009.

[25] S. Chattopadhyay, S. Banerjee, F. a. Rabhi, and U. R. Acharya, "A Case-Based Reasoning system for complex medical diagnosis," *Expert Syst.*, vol. 30, no. 1, pp. 12–20, Feb. 2013.

[26] I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," *J. ofMachine Learn. Res.*, vol. 3, pp. 1157–1182, 2003.

A Framework for Accessing Patient Health Records Through Multi Channel of Devices

Noorayisahbe Mohd Yaacob
Biomedical Computing and Engineering Technologies
Research Group (BIOCORE), C-ACT
Faculty of Information & Communication Technology,
Universiti Teknikal Malaysia Melaka,
Melaka, Malaysia
syayaacob88@gmail.com

Mohd Khanapi Abd Ghani, Abd. Samad Hasan Basari
Biomedical Computing and Engineering Technologies
Research Group (BIOCORE), C-ACT
Faculty of Information & Communication Technology,
Universiti Teknikal Malaysia Melaka,
Melaka, Malaysia
{khanapi;abdsamad}@utem.edu.my

***Abstract*—Producing a new framework is important in healthcare framework environment for accessing patient health records through multi-channel devices. This is especially important, as users of future health care systems will be increasingly characterized by diversity. By relying only on highly experienced and technology-prone user groups, which might have been typical users in the last decades, is not sufficient anymore. A design of framework for pervasive healthcare is essential to provide different types of medical services and to support users individually (according to user profiles), adaptively (according to the course of disease) and sensitively (according to living conditions) in order to successfully design new medical technologies.**

***Index Terms*—Healthcare Framework, Personal Health Electronic Health Records, Pervasive Healthcare Systems.**

I. Introduction

While pervasive health care systems bear the potential to provide patients with a new quality of medical homecare, the complexity of such systems raises fundamental questions of behavior, communication and technology acceptance. This is especially important, as users of future health care systems will be increasingly characterized by diversity. Relying only on highly experienced and technology-prone user groups, which might have been typical users in the last decades, is not sufficient anymore. Rather, elderly users with a completely different upbringing and domain knowledge and ill or handicapped people will have to use the systems. Today, the understanding in which way physical, emotional and cognitive abilities caused by individual learning histories and health states, may impact the usage and acceptance of pervasive healthcare technologies is restricted[1]. Prototyping systems will develop and used for demonstration application in a number of domains[2]. However, realistic pervasive environments of the future need to integrate with enterprise systems, including software, hardware, databases, standards and life cycle. The aim of d a framework for ppervasive healthcare is to provide different types of medical services and to support users individually (according to user profiles), adaptively (according to the course of disease) and sensitively (according to living conditions) with considered in order to successfully design with new medical technologies.

II. Related Work

Personal health records (PHRs) are darting broaden area of medical informatics' due to the belief that they may improve healthcare delivery and control costs of care. The PHRs in use or in development today support a variety of different functions and consequently offer different value proposition. PHRs are defined as "a set of computer based tools that allow people to access and coordinate their lifelong health information and make appropriate parts of it available to those who need it[3]. Computer-based records, currently referred to as electronic health records (EHRs), have been proposed as a means for improving availability, legibility, and completeness of patient information. EHRs have been commercially available since 1970s, but their adoption has become widespread only recently, driven in part by financial encouragement ($27 billion worth) from the Office of the National Coordinator for Health Information Technology. According to [4] prompt access to a patient's electronic medical history or lifetime health record (LHR) is fundamental for providing seamless and continuous care. This can be achieved through the convergence of ICT, medical content and health knowledge. Through this convergence, the patient's medical history can be shared among healthcare professionals and across healthcare facilities regardless where the previous visit was made. However, the fragmented and paper-based medical records have significant limitations such as illegibility, unavailability, sheer physical volume (over time), difficult transferability and integration between providers and institutions, and the necessity to record the same data many times (duplication) on different documents [1,2,3]. Today, almost 3 of 4 physicians report using EHRs. Furthermore, with this increase in adoption, the medical community is now beginning to appreciate both the promise and perils of "going electronic." The framework is technology

based and concentrates on healthcare network infrastructures that focus on the telecommunication network, devices and data transmission. The framework defines a flexible gateway to receive and transmit patient physiological data (for example, vital signs) from where the patients might be able to access data at the central system. The framework can be described through the following methods concerning each key technology and system process [5]. Apply a more suitable of data gathering technologies to help better understand the many facets of patients in their daily lives and then modify therapies to the individual. They clearly describe how these technologies capture important aspects of better care that tuned not only the situation but also patients history, monitoring and control various vital functions under a very stressful conditions and between operating room to augment human ability to detect pattern of concern that could require immediate action. Besides facilities valuation, this easily help to reduce costs by getting appropriate care to the people who need it much faster. It also create expert care improvement accessible to more people and making healthcare more personalized, prompting individuals to take more responsibility in order to maintaining human health[2]. In [1], they analyze the contribution factors on acceptance pattern in different solutions for the implementation of medical technologies such smart home environment, smart clothes, mobile devices and also assessed each implementation concepts with the different usage of motive. The authors applied several method in order to collect the data and execute the ANOVA formula to get the better result of analysis[1]. S are described which examine and reacts to an individual's changing context interaction between people's with devices, computers and others in order to help navigate unfamiliar places. Clearly explained about four categories of context – aware application and developed prototype system for each categories to obtain the better result and data, focus on the hardware design and implementation of badge and network of infrared receivers[6]. Detailed Electronic Medical records (EMR) transitioning from desktop systems to mobile devices is also explained. This innovation presents challenges to medical practitioners in term of doctor – patient interaction, patient record integrity and continuing reliance on paper – based annotation schemas. The describe findings from a pilot study of EMR use in medical clinic and proposes guidelines for the design of EMR systems[7]. Personal Health Records and Personally Controlled Electronic Health Record (PCEHR) system are discussed and making the comparison between country such USA, Scotland, England, and others. Explain about international experiences for PCEHR. In [7], this paper discuss about quality use of medicine[8]. The role of clinical decision support systems in medicine is an important category of which expert system which considered about user requirement in the health care domain[9].

III. Research Methodology

This section outlines the activities that involved in this research. The aim of this research is to produce a framework for accessing patient health records through multi-channel of devices and subsequently will investigate the proposed framework through proto-type system to validate its interoperability, portability and standard in order to improve the framework.

Figure 1 shows the flowchart of the research methodology

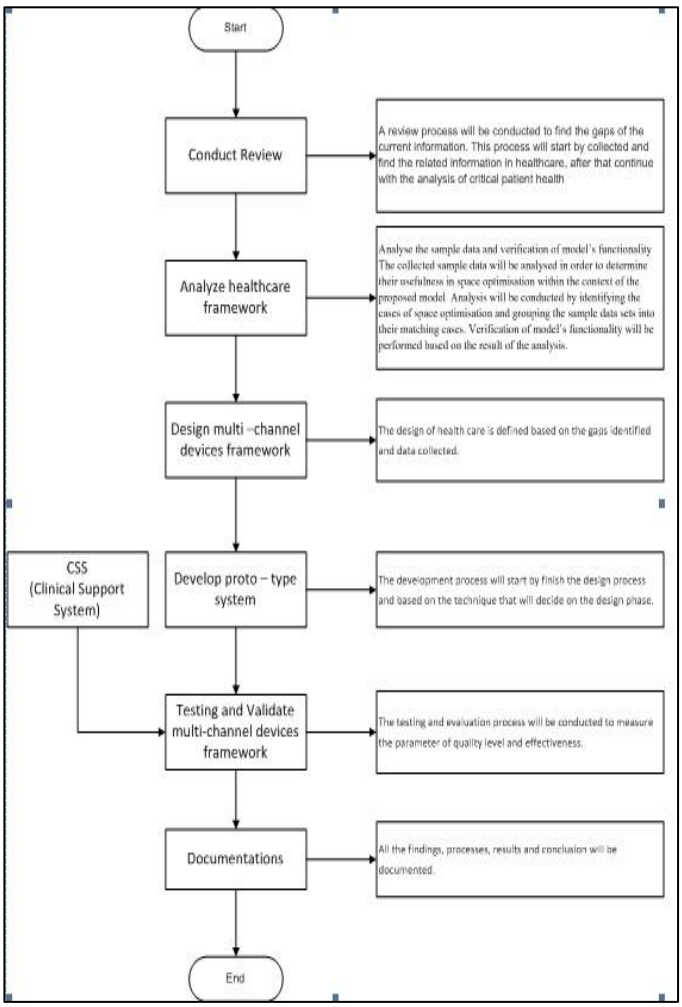

Fig 1: Flowchart of Methodology [Appendix 1]

The steps involved in this research depicted in figure 1 are as follows:

(a) Conduct Review

The conduct review process involve activities which started with a literature review and data collection of healthcare framework. A review process will be conducted to find the gaps of the current information. This process will start by collecting and finding the related information in healthcare, after that continue with the analysis of critical patient health.

(b) Analyze healthcare framework

The sample of data will be analyzed and verified according to model's functionality. The collected sample data will be

analyzed in order to determine their usefulness in space optimization within the context of the proposed model. Analysis will be conducted by identifying the cases of space optimization and grouping the sample data sets into their matching cases and compare the results. Verification of the model's functionality will be performed based on the result of the analysis.

(c) Design multi-channel devices framework

The design of health care is defined based on the gaps identified and data collected.

(d) Develop proto- type system

The development process will start after completion of the design process and based on the technique that are decided during the design phase.

(e) Test and Evaluate multi-channel devices framework

The testing and evaluation process will be conducted with Clinical Support System (CSS) to measure the parameter of interoperability, portability, standard and effectiveness.

(f) Documentation

All the findings, processes, results and conclusion will be documented.

The result of this analysis will be reported and published accordingly in journals or conferences.

For each methodology steps, a schedule and milestones have planned as the following table:

Table 1: Schedule and Milestones

Milestone	Duration (Month)	Date Started
Review and identify the current framework and pervasive method in healthcare.	18	September 2013
Analyze healthcare framework	4	January 2014
Design framework for multi-channel of devices	3	March 2014
Implementation and Developed system proto-type	3	May 2014
Validate proposed framework based on patient data.	2	June 2014
Completion of Documentation for analysis, results and publish the result.	6	30/11/2014

As shown in the table 1, the research schedule and milestone has been explained detailed including planned and duration to ensure the research progress as planned.

IV. Research Progress

As shown in Figure 1, the research is at conducting review process. Currently, the research is in the literature review stages in order to find more related and suitable works.

V. Conclusion and Future Work

Designing A framework for accessing patient health record through multi – channel of devices is aimed to provide different types of medical services and to support users individually (according to user profiles), adaptively (according to the course of disease) and sensitively (according to living conditions) in order to successfully design new medical technologies. The objective for this research is to produce a framework for accessing patient health records through multi-channel of devices.

The development of a framework for accessing patient health record through multi – channel of devices is to provide a different types of medical services and to support a myriad of different function and consequently offer different value proposition.

Acknowledgment

The authors would like to thanks Universiti Teknikal Malaysia Melaka (UTeM) for their support. This research is part of Master of Information and Communication Technology and funded by Universiti Teknikal Malaysa Melaka under grant PJP/2012/C-ACT/Y00003.

REFERENCES

[1] M. Ziefle and C. Rocker, "Acceptance of pervasive healthcare systems: A comparison of different implementation concepts," *Proceedings of the 4th International ICST Conference on Pervasive Computing Technologies for Healthcare*, 2010.

[2] J. P. Black, W. Segmuller, N. Cohen, B. Leiba, A. Misra, M. R. Ebling, and E. Stern, "Pervasive Computing in Health Care : Smart Spaces and Enterprise Information Systems."

[3] D. Johnston, D. Kaelber, E. C. Pan, D. Bu, S. Shah, J. M. Hook, B. Middleton, and P. Page, "A Framework and Approach for Assessing the Value of Personal Health Records (PHRs) Center for Information Technology Leadership (CITL), Partners HealthCare System , AMIA 2007 Symposium," pp. 374–378, 2007.

[4] M. K. A. Ghani, R. K. Bali, R. N. G. Naguib, I. M. Marshall, and a. S. Shibghatullah, "The design of flexible front end framework for accessing patient health records through short message service," *2007 Asia-Pacific Conference on Applied Electromagnetics*, pp. 1–5, Dec. 2007.

[5] M. E. Gordon, "A Framework for the Protection of Privacy in an Electronic Health Environment by A Framework for the Protection of Privacy in an Electronic Health," 2010.

[6] B. Schilit, N. Adams, and R. Want, "Context-aware computing applications," *Workshop on Mobile Computing Systems and Applications*, pp. 85–90.

[7] K. Larkin and A. Kelliher, "Designing flexible EMR systems for recording and summarizing doctor-patient interactions," *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems - CHI EA '11*, p. 1609, 2011.

[8] J. Reeve and R. Hosking, "Personal electronic health records : the start of a journey," vol. 36, no. 3, pp. 70–73, 2013.

[9] K. W. Darlington, "Designing for Explanation in Health Care Applications of Expert Systems," *SAGE Open*, vol. 1, no. 1, Apr. 2011.

A Framework for Gathering Electronic Medical Record through Telehomecare

Nor Afirdaus binti Zainal Abidin
Faculty of Information and Communication Technology,
Universiti Teknikal Malaysia Melaka,
Hang Tuah Jaya, 76100, Durian Tunggal, Melaka,
Malaysia.

Dr. Mohd Khanapi Abd Ghani
Faculty of Information and Communication Technology,
Universiti Teknikal Malaysia Melaka,
Hang Tuah Jaya, 76100, Durian Tunggal, Melaka,
Malaysia.

***Abstract*— Chronic disease among people in Malaysian is increasing nowadays. The patients need to be monitored regularly to ensure their health level is safe. Due to these factors, the Telehomecare system should be introduced in Malaysia even though the system is widely used in developed countries. This Telehomecare system is focus to provide a monitoring service for patient with chronic diseases. The patients is based at home will be monitored by nurses and doctors from remote hospitals and carried out using CCTV, bio sensors, and the Internet as a medium for communication and transmission of biological data from home to hospital. Before the system can be used at home, several factors need to identify especially to develop a framework in gathering patient health record in Telehomecare environment.**

Keywords: chronic disease; Telehomecare; framework; Malaysia.

I. INTRODUCTION

Telehomecare technology development is increasingly growing in the international medical technology. The term 'telehomecare' show that patients care is monitored by health workers through the use of ICT. Telehomecare is expected to improve the existing medical system towards efficiency of patient care, particularly chronic patients such as diabetes, heart disease and high blood pressure.

Another goal of telehomecare is to enhance the quality of life for the patients that still being monitored by healthcare professionals at home. The implementation of telehomecare technologies across sectors poses a challenge because it is not only in terms of technical aspects, but also in terms of merging work flows, coordinating tasks and the patient care process. This challenge is rooted in the implementation of telehomecare technology in an inter-organisational (network) and inter-professional landscape involving various healthcare professionals and their objectives and working culture.

Due to the above scenario, the overall aim of the research is to build a framework for gathering patient medical record through telehomecare.

A. Background Of Chronic Disease

The disease profile of the world is changing at an alarming rate. Chronic diseases include heart disease, stroke, cancers, chronic respiratory diseases, diabetes and mental disorders are now accounted for 47% of the global burden of disease and 60% of all deaths. World Health Organisation (WHO) in its groundbreaking report *Preventing Chronic Diseases: a Vital Investment* shows that 80% percent of chronic disease deaths occurred in the low and middle income countries and 50% of these deaths occurred prematurely in people under 70 years of age, disputing the long-held notions that chronic diseases are mainly affecting affluent countries and older generations. In Malaysia, WHO estimated that chronic diseases accounted for 71% of all deaths in 2002, using calculation based on the standard methods to maximise cross-country comparability. The first Malaysian Burden of Disease and Injury Study showed that the top two leading causes of deaths in the year 2000 for both sexes were ischaemic heart disease and cerebrovascular disease; and the total years of life lost for the Malaysian population in 2001 was 1.7 million years with almost two-thirds of this burden of premature deaths resulted from chronic diseases. Of the total burden of disease measured using Disability Adjusted Life Years (DALYs), approximately 69% was contributed by chronic diseases.

Globalisation, urbanisation, population ageing, as well as general policy environment has been identified as the underlying determinants responsible for the changing pattern of diseases in developing countries. Malaysia is of no exception in terms of the effect of urbanisation, as an estimated 63% Malaysians now live in urban areas. Although only 6.6% of Malaysians are aged 60 years and above, the ageing population is projected to gradually rise as the life expectancy at birth for both males and females has increased over the years to 70.6 years for males and 76.4 for females in 2005. Preliminary data from Malaysian Non-Communicable Disease (NCD) Surveillance 2005/06 estimated that approximately 11.6 million Malaysian adults aged 25-64 years were having at least one risk factor for chronic diseases and only about 3% did not have any risk factor. A recent national survey which sampled more than 16,000 Malaysians, showed that the prevalence of hypertension amongst those aged 30 years and above has increased from 32.9% in 1996 to

40.5% in 2004, while the overall prevalence of obesity amongst those aged 15 years and above was 11.7% in 2006, which is a staggering increase by 280% since the last National Health and Morbidity Survey II (NHMS II) in 1996.

B. Chronic Disease In Europe

Chronic diseases account for seven out of ten deaths in the United States, and consume 75 cents of every dollar spent on health care. Nearly half of all people in the US of all ages, race, and socio-economic background that live with a chronic condition, such as high blood pressure, diabetes, or asthma. More than two-thirds of all deaths are caused by one or more of five chronic diseases such as heart disease, cancer, stroke, chronic obstructive pulmonary disease and diabetes. Many chronic diseases are lifelong conditions, and their impact lessens the quality of life, not only of those suffering from the diseases, but also of their family members, caregivers and others. Chronic diseases not only affect health and quality of life and reduce economic productivity by contributing to increased absenteeism and poor at work performance, but they also drive rising health care costs and threaten health care affordability.

In 2007, the US spent over $2.2 trillion on health care, and three-quarters of this went to treat patients with one or more chronic diseases.

C. Chronic Disease In South East Asia

Chronic disease could kill up to 4.2 million people annually in Southeast Asia by the year 2030, according to research published in the UK medical journal, The Lancet.

In low- and middle income countries, about 29% of deaths occurred before the age of 60 in 2008, and about 48% are estimated to have occurred under the age of 70, compared with 26% in high income countries. Worldwide, approximately 44% of all noncommunicable disease deaths occurred before the age of 70.

The figure below shows the proportion of total noncommunicable disease deaths under age 70 in 2008 by their cause. Cardiovascular diseases were responsible for the largest proportion of noncommunicable disease deaths under age 70 (39%), followed by cancers (27%), chronic respiratory diseases, digestive diseases and other noncommunicable diseases were together responsible for approximately 30% of noncommunicable disease deaths, and diabetes was responsible for 4%.

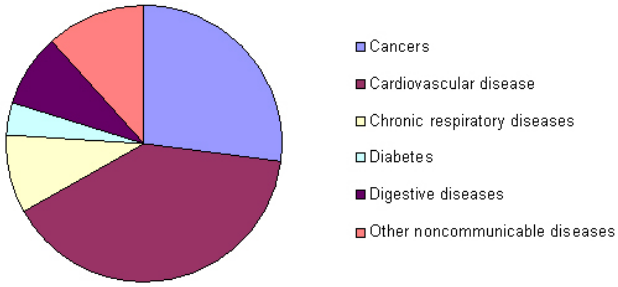

Figure I-1 % total chronic disease death, WHO, 2008

D. Chronic Disease In Malaysia

According to Department of Statistics, Malaysia that prepared based on death records from National Registration Department, diabetes mellitus, cardiovascular diseases, stroke, cancer, stroke, hypertension and others are among the top chronic diseases that cause death in Malaysia.

Chronic diseases are the major cause of death and disability in Malaysia, accounted for 71% of all deaths and 69% of the total burden of disease. The number of Malaysians with undiagnosed NCD risk factors had increased alarmingly, such as for diabetes, there may be one undiagnosed patient for every patient diagnosed, for hypertension, for every patient diagnosed, there may be another three still undiagnosed and for high cholesterol, there may be three undiagnosed patient for every patient diagnosed.

#	Medically certified	%	#	Not medically certified	%
1.	Ischaemic heart disease	12.9	1.	Old age 65+ years	59.4
2.	Pneumonia	7.0	2.	Cancer	6.9
3.	Cardiovascular disease	6.6	3.	Asthma	6.6
4.	Septicaemia	6.0	4.	Heart disease	5.6
5.	Transport accident	5.8	5.	Diabetes	3.3
6.	Chronic lower respiratory disease	2.3	6.	Fever	1.7
7.	Malignant neoplasm of trachea, bronchus and lung	2.2	7.	Stroke	1.7
8.	Diabetes mellitus	1.7	8.	Hypertension	1.6
9.	Certain condition originating in the perinatal period	1.6	9.	Kidney disease	1.0
10.	Diseases of the liver	1.4	10.	Notifiable disease	1.0
All causes		76,016		All causes	48,841

Table I-2 10 Principal Causes of Death, Department of Statistics, Malaysia 2008

E. Telehomecare Trend

Telehomecare system is focus to provide a monitoring service for patient with chronic diseases. The patients is based at home will be monitored by nurses and doctors from remote hospitals and carried out using CCTV, bio sensors, and the Internet as a medium for communication and transmission of biological data from home to hospital.

Telehomecare is being adopted across many jurisdictions in the US through 2007 to 2009. The slow pace of adoption is partly due to a lack of central data for proof of efficacy, economic benefit and partly a function of innovation speed in the health community and it can take 10 to 20 years for a new practice to go from development to mainstream use (Bohmer, 2010).

Telehomecare trend is still new in Malaysia but it started giving positive developments in the medical system in Malaysia. This will take time as well as technological, economic, social factors should be considered especially public awareness in health care and monitoring by patients itself at home.

F. The Importance Of Telehomecare For Malaysia

The effect of Malaysia's Telehomecare system solution on the nation will serve as a model for other nations to follow. The benefits include:

- Reducing cost and time - Cost of travel and time of patient can be reducing.
- Reducing the demand for healthcare services through health information and education - Many studies have shown that appropriate and well-designed health promotion can be done to raise the awareness of and to reduce the risk factors leading to these chronic diseases through information and education. Telehomecare system will help patients to improve their knowledge and concern about their health.
- Consequence of a healthy population on economic development and productivity - In addition to reducing healthcare cost, improved health as a result of preventive and promotional programs through information and education has important and substantial effect on a nation's output.
- Invaluable data collection - The data collected will be invaluable to policy makers, researchers, healthcare providers and many others. For example, it can be used for more efficient allocation of scarce resources making an impact where it is needed the most.

II. RESEARCH PROBLEM

Telehomecare growing in developed countries like Europe, but in Malaysia is still new system to be introduced. Procedures health workers in taking patient data must be identified before the Telehomecare system is used at home. Apart from the constraints of the economic, social and technology, can we assure the data and procedures that will be used at home can help the patient? These are some issues and challenges that need to be addressed before Telehomecare system is used at home.

A. What is the best technique to collect patient medical record?

This question is related to the nursing process on how their collect patient data at the hospital. At the end of the study, we can learn the best techniques in obtaining patient data and to develop a framework in gathering patient health record through Telehomecare.

B. What is the crucial data set to be collect for supporting the framework?

This question is important in determining the relevant patient data to be used. This set of data will be used to develop a framework and for helping health workers to diagnose patient problem at home.

III. RESEARCH OBJECTIVE

The successful introduction of telehomecare with tangible outputs requires in depth understanding of the existing health care system of the country and its challenges, strongly expressed 'genuine need' for the service by the all the domains (patients, health workers, organization and the public), the actual status of ICT infrastructure in the country and costs. Rigorous research should be carried out with an appropriate methodology before implementing new systems which all contribute towards sustainability of the project (Lama 2006).

In this study the focus is on a framework for gathering electronic medical record through telehomecare. This research is to fulfill the following objectives:

A. Provide a literature review

The first objective in this research is for analyzing and comparing existing framework for gathering patient medical record through Telehomecare. Aim of this objective is to develop a framework in gathering patient in a different way and to review the current status of framework in Malaysia.

B. Explore nursing process through field case study in public and private hospitals in Malaysia.

The second objective is to do case studies of nursing process in public and private hospitals. It will help to provide a way to share knowledge and gain more information on their experiences. At the end of this level, clinical patient data should be identified and defined to help gathering patient medical record.

C. Develop a framework

This objective is to develop a framework that achieves research objective. At this level the basic development of system has be identified and constructed. Collaboration with the hospital is required to perform the validation. Any deficiencies in the identification of appropriate techniques and consequently to improve work processes.

D. Validate the framework

The purpose of validation is to detect and then verify any gaps that may not represent the framework from achieving the aim of the research.

IV. RESEARCH METHODOLOGY

A. Introduction

The way in which research is conducted may be conceived of in terms of the research strategy employed and the research instruments utilised in the pursuit of a aim of research. The purpose of this chapter is to expound research strategy, including the research methodologies adopted and to introduce the research instruments that been developed and utilised in the pursuit of research are successfully achieved.

B. Survey Research

According to research objectives, me intend to investigate the use of ICT (Telehomecare) among health workers/patients in managing patient at home. The survey will involve health workers and patients in Johor Bahru and Kuala Lumpur.

C. Case Study Research

There are a number of important articles describing the case study approach to research that we refer to. Key among these is Benbasat et al.'s (1987) paper with its comprehensive definitions and suggestions for the conduct of case research. The case study is considered by Benbasat et al. (1987, p.370) to be viable for three reasons:

i. It is necessary to study the phenomenon in its natural setting;
ii. The researcher can ask "how" and "why" questions, so as to understand the nature and complexity of the processes taking place;
iii. Research is being conducted in an area where few, if any, previous studies have been undertaken.

From this research point of view, it became clearer that the case study research method was particularly well-suited to this research.

V. EVALUATION CRITERIA

The data necessary will be collected from some selected government and private hospitals in Malaysia. By this, the basic development system will be tested, modified, if necessary and validated thoroughly.

ACKNOWLEDGEMENT

This work was fully supported by Faculty of Information & Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). The authors would like to her supervisor for support, advice and anonymous input, which assisted in improving this paper.

REFERENCES

[1] Castle N, Ferguson JC, "What is nursing home quality and how is it measured?" Gerontologist 2010, 50(4):426–442.
[2] Centers for Disease Control and Prevention. Chronic Disease Overview. Atlanta: CDC, 2009.
[3] Hoffman D., An Urgent Reality, "The Need to Prevent and Control Chronic Disease", Atlanta National Association of Chronic Disease Directors, 2008.
[4] Lama, T., "Is telemedicine a viable strategy for healthcare provision in the rural areas of developing countries like Nepal?" Northumbria University, 2006.
[5] Levenson S, "The basis for improving and reforming long-term care, Part 3 Essential elements for quality care", J AmMed Dir Assoc., 2009, pp.597–606.
[6] Mor V, Gruneir A, Feng Z, Grabowski DC, Intrator O and Zinn J, "The effects of state polices on nursing home resident outcomes", 2011, pp.3–9.
[7] Nolte E and Mckee M., "A review caring for people with chronic conditions", A Health System Perspective, Open University Press, 2008, pp.64–91.
[8] WHO, The World Health Report 2003, "Shaping the future", Geneva, World Health Organisation, 2003.
[9] WHO, "Preventing chronic diseases: a vital investment" Geneva, World Health Organisation, 2005. *(online)*
[10] WHO, "The impact of chronic disease in Malaysia", Geneva: World Health Organisation, 2005. *(online*
[11] Yusoff AF, Kaur G, Omar MA and Mustafa AN, "Malaysian burden of disease and injury study 2005", Institute for Public Health, National Institute of Health, Malaysia.
[12] Department of Statistics Malaysia, 2008. *(online).*
[13] MOH, NCD Surveillance, "Report on NCD risk factor in Malaysia", Ministry of Health, Malaysia, 2005/2006.

Software Architecture

Conceptual Model for Managing Variability in Product Line Engineering **Azizah Binti Rahmat and Mohd Hasan Bin Selamat**
Hybrid Mobile Application Framework for Pervasive Health Record (HMAF4PHR) **Farah Aris and Mohd Khanapi Abd Ghani**
Designing Hybrid Software Architecture Framework for Capturing Clinical Findings during Doctor-Patient Consultation **Omar Mukhtar Hambaran and Mohd Khanapi Abd Ghani**
Health Informatics Framework for Postnatal Care: An integration between Modern Medicine and Traditional Malay Medicine **Raja Rina Raja Ikram, Abd Ghani, M.K. , and Abdullah, A.**
Enhancing Generic Pipeline Model in Preventing Code Clone during Software Development **Al-Fahim Mubarak Ali and Shahida Sulaiman**
Developing a Maturity Model to Measure Service-oriented Enterprise Architecture (SoEA) Adoption in an Organisation **Nur Azaliah Abu Bakar, Harihodin Selamat and Nazri Kama**
Ontology for Evaluation of Semantic Web Services Ontology **Nurhamizah Mohd Hamka and Radziah Mohamad**

Conceptual Model for Managing Variability in Product Line Engineering

Azizah Binti Rahmat
Software Engineering Section of MIIT
Universiti Kuala Lumpur
Kuala Lumpur, Malaysia
azizah@miit.unikl.edu.my

Mohd Hasan Bin Selamat
Department of Computer Science
Universiti Putra Malaysia
Serdang, Malaysia
mhasan@upm.edu.my

***Abstract*—Product Line Engineering (PLE) and Reference architecture (RA) are one of the important key elements in reused based software engineering. Several variability management researches had been conducted within the context of single and multi-domain. However, managing variability in cross domain requires extensive and complex mapping compared to single domain. This is due to the involvement of several product line architectures (PLA). Therefore, there is a need to create a mechanism or an approach to map the RA to system architecture (SA) of different domains and capable of resolving any existing mapping problem. This paper focuses on the GRACES conceptual model of intermediate architecture that enables the reference architecture to be mapped to product within the same domain as well as from the different domains in a systematic way. GRACES consists of goal (actor-oriented approach is the main goal), group (group of actors), role, relationship, responsibility, affordances, context, enabler and service. The intermediate architecture acts as an intermediary between the RA and SA. This GRACES model will ease the mapping between RA to CA within the intermediate architecture because actor is always natural regardless of domain. Ontology alignment and isomorphic graph theory is used to measure the consistency of ontology mapping between ontoRA, ontoIA and ontoSA and the mapping for RA, IA and SA.**

***Index Terms*—Reference architecture, Cross domain, Product Line Engineering, Variability management, Conceptual model, Actor-oriented approach**

I. Introduction

Product line engineering (PLE) is a reuse-based software development approach of family software within the same domain. PLE enables developers to develop multiple products, and each has its own cycles of release and version [4][12]. Domain engineering represents Producer (production of common artifacts) and the application engineering represents Consumer (consumes these artifacts to develop products in the consumer family – product A, B and C). Reference architecture (RA) is "an abstract architecture" while concrete architecture (CA) is an application architecture/ system application (SA) that has the common features of reference architecture [1]. Product line architecture (PLA) is considered as domain engineering and it is a subset of RA since it involving the commonality and variability and producing the reusable core asset [1][12][15]. Each concrete application in the application engineering will consume and maintain it. The motivation of this research is the systematic design with reuse. Figure 1 shows the relationship between the RA and SA. Intermediate architecture is introduced as a solution for mapping program between RA and SA as depicted in Figure 2. This research will focus at the functional layer of software architecture as shown at figure 3. Figure 2 shows the relationship between RA and SA via IA.

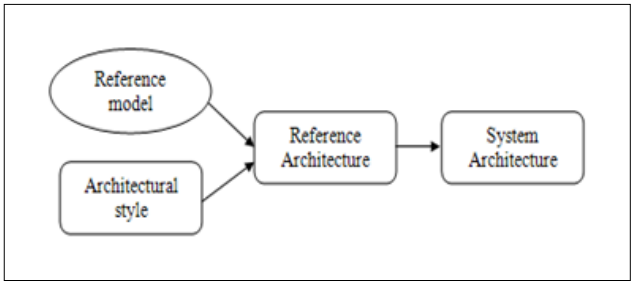

Fig. 1. The role of RA in PLE [1]

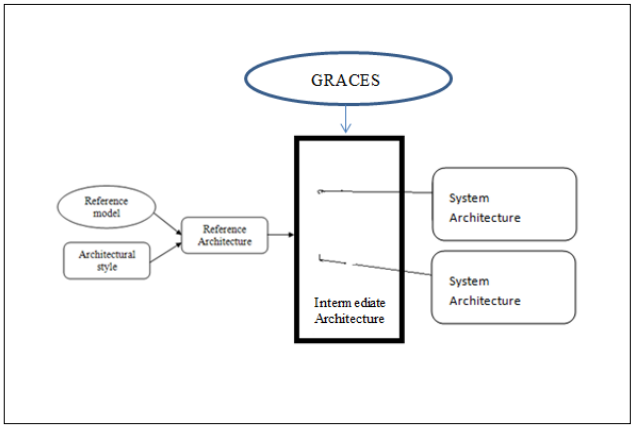

Fig. 2. The role of IA and GRACES in PLE

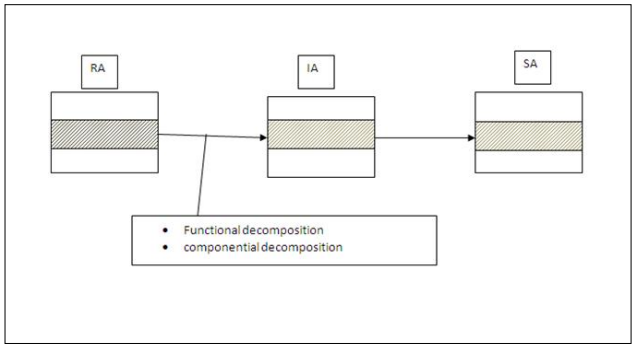

Fig. 3. The layer of this approach

II. RELATED WORK

PLA is a subset of RA [1][15]. Hence, RA is specialized to PLA where each PLA is instantiated to instance of the architecture. However, there is lack of detail specification on how the mapping is done. Hence, our approach will fill the gap between the RA and SA, and proposed the actor oriented approach that will map it systematically. Numerous researches had been conducted involving the domain engineering and application engineering by combining the software product line and service-oriented approach to enable cross domain mapping. Guidelines and steps on core asset development by Meideros [13] focus on the design of domain specific architecture. While Nguyen [16], Lee [10], Karam [7], Capilla [5], Kotonya [8] and Asadi [2] combined both domain engineering and application engineering steps with service-orientation approach. The objective is to produce a family of service-oriented product lines and allow cross domain mapping. However, the mapping is still not specific and without an intermediary that will link the RA to SA systematically.

Somehow rather, Lee [10] and Kotonya[8] applied the featured oriented approach and stored the reusable service component in reusable asset repository. It is more to reusable asset repository rather than systematic mapping of RA to SA, The domain engineering and application engineering that is based on SOA layer is introduced by Moosavi [14]. There are four types of variability which are workflow, composition, interface and logic variability. However, it is more to application engineering and not the mapping. Context-oriented programming by Salvaneschi [8] support dynamic adaptation that allows certain language-level abstraction .However, the focus is within the domain and application engineering of a single domain. Context is one of the element of GRACES and it will be applied in this research because the actor will act depends on specific context or stakeholder.

III. RESEARCH METHODOLOGY

The research methodology is conducted as showed at figure 4. The problem arises in determining the accuracy of the graph theory and the mathematical model during the evaluation of model and mapping using ontology alignment.

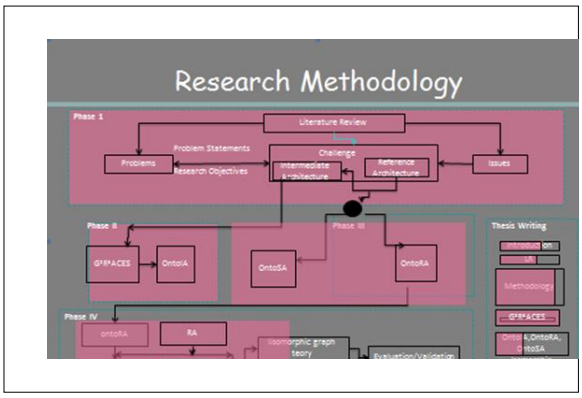

Fig. 4. Research methodology

IV. FINDING

A. Intermediate Architecture and GRACES conceptual model

Intermediate Architecture (IA) is the intermediary between RA and SA in the form of conceptual model called GRACES. It is the conceptual model of intermediate architecture in the domain of functional layer that allow cross domain between RA to SA in a systematic way. It consists of goal (G), group (G), role (R), responsibility(R), relationship(R), action/affordances (A), context (C), enabler (E) and services (S). Actor theory agent is the goal, the group the team of actors, while the role, relationship, and responsibility is what the actor must do, the affordances is what actor can do, the context is the scenario, environment or stakeholder, the enabler is the tools used and service is the supervision as depicted at figure 5.

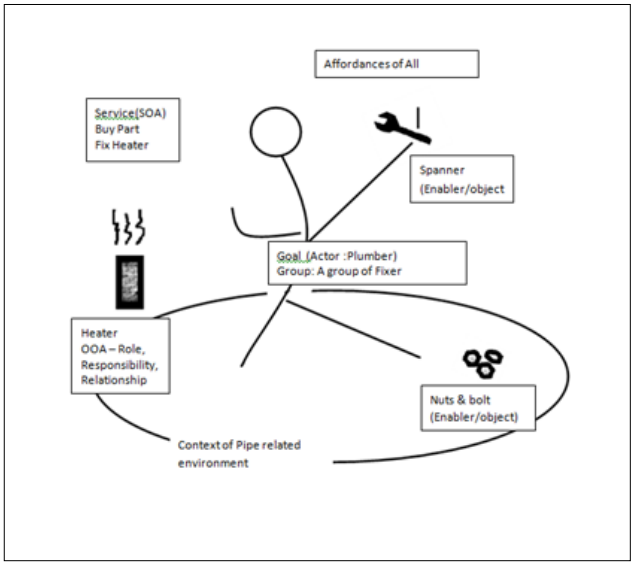

Fig. 5. GRACES conceptual model

Actor is the main goal of GRACES model. The idea of actor-oriented approach is inspired by actor theory in concurrent programming [3]. Actor could be human or non-human because it is defined as any entity capable to act changing in a certain state of affairs [17]. Actor theory in [6] is producing smart environment software reference architecture. Component call actors execute and communicate with other actors in concurrent programming [11][9]. Actor has goal that resembles the actor. Goal is the key element to differentiate the role of an actor for a specific domain or environment. For example, plumber is a fixer of pipe related things, while mechanic is a fixer of automotive related matters. Both plumber and mechanic share the same goal of fixing things. The way to achieve the goal depends on the equipment and tools because the objects to be fixed are not the same. However, the goal of a common actor is the same. This actor will act differently in each specific environment or domain he is in. A fixer actor will be a plumber when he fixes a sink in a house. Otherwise, he will be a mechanic who fixes a car in the garage. Table 1 shows the relationship of goal, actor and domain.

TABLE I.

Goal	Actor	Domain	Role
To fix something	Fixer	Automotive	Mechanic
		Pipe	Plumber
To convert something	Converter	E-commerce	Currency converter
		Medical	Measurement converter

B. X2Y actor

Actor-oriented designs promote systematic reuse in different domains. Systematic reuse is very important in developing software because the design and architecture is planned purposely for reuse in the application engineering. The Converter actor should have the same structure flow for bank (e-commerce) and medical domain. Assume that Converter X2Y is an actor to convert X to Y Where; Y=F(X) =kX

k is a constant in which the value is obtained from the domain being applied (type of system architecture). Converter X2Y is an actor because it plays different roles in different domains such as in banking or medical systems that achieves the goal of that particular converter function. The converter X2Y can communicate with different actor such as presenter. Presenter actor should then present the result of that Convertor's function. Below is the example where Converter X2Y can cross different domains.

Let;

Domain: bank	Domain: Medical
X=USD	X= L(iter)
Y=RM	Y=ML

Converter USD2RM	Converter L2ML
Y= 0.3 x X	Y=1000 x X
k=0.	k=1000

The relationship between RA and PLA initiated by Nakagawa [15] and Angelov [1] intended to cross domain mapping between RA and SA. The actor-oriented approach will act as the intermediary between the RA and SA as depicted at figure 6. This research uses functional layer to depict the relationship along RA to SA. ConverterX2Y actor is natural regardless of domain. However, the X and Y will be map to USD, L and RM, ML respectively. System architecture will implements it accordingly

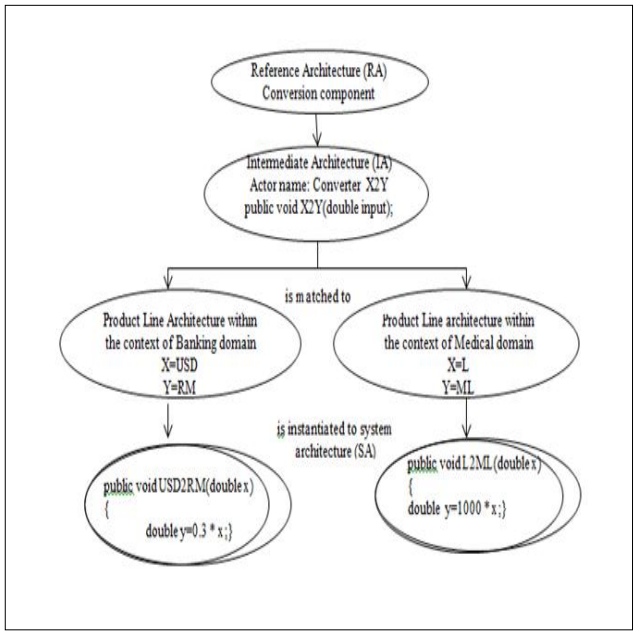

Fig. 6. Relationship between RA, IA and SA using actor.

C. Validation

Ontology alignment and isomorphic graph theory will be used to validate the model and the mapping between RA, IA and SA. The validation model is depicted at figure 7.

For example:
ConverterX2Y
Y=0.5 x X
CX2Y
Y=7 x X ➔ consistency difference in constant only
CX2Y
Y=0.5 x X2 ➔consistensy difference in exponent

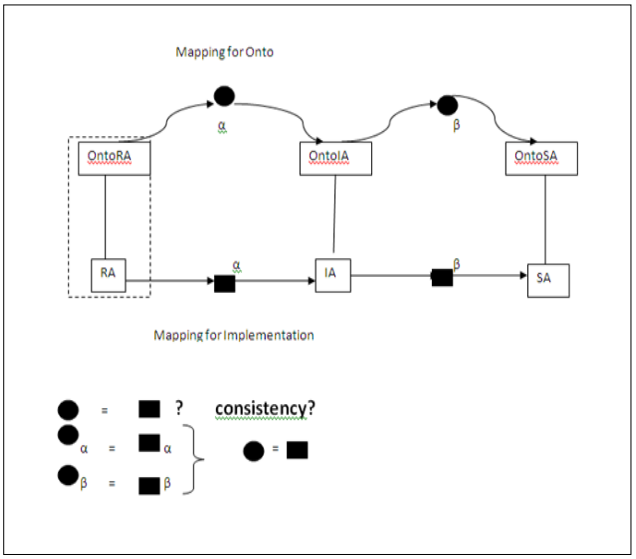

Fig. 7. Validation model

V. CONCLUSION

This research effort is to introduce GRACES conceptual model for intermediate architecture and adapt the actor theory in the field of software product line and software architecture in particular. This paper concludes that actor-oriented approach enables mapping between the RA to SA to resolve the extensive and complex mapping of cross domain. Actor-oriented approach is chosen because object-oriented and service-oriented approach alone will change based on the domain used. For example, a student object will be used by different domain in a different way. Banking domain will use student object for study loan while Education domain will treat student as customer. Another example is converter service, currency converter services involves the currency conversion, while measurement converter involves metric to imperial conversion and vice versa. Unlike other approaches, the actor-oriented approach seems to remain stable. For example, Fixer actor can be mechanic or plumber. Plumber will buy parts (services provided by other party) that related to his job scope (domain) or context while part like bolts and nuts is the object Mechanic and plumber will use different kind of bolts and nuts. The contribution of this research is GRACES Conceptual model of intermediate architecture (IA).

REFERENCES

[1] Angelov, S. et al, 2012. A framework for analysis and design of software reference architectures. Information and Software Technology, 54(4), 417–431. doi:10.1016/j.infsof.2011.11.009.

[2] Asadi, M. et al 2009. Model-driven development of families of service-oriented architectures. Proceedings of the First International Workshop on Feature-Oriented Software Development , ACM New York, USA. pp. 95-102.

[3] Baker, H. G. 1978. Actor systems for real-time computation. Technical report. Massachusetts Institute of Technology Cambridge, MA, USA.

[4] Bass, L. et al, 2003. Software architecture in practice (second ed.). Addison Wesley, MA, USA.

[5] Capilla, R., and Topaloglu, N. Y. 2005. Product lines for supporting the composition and evolution of service oriented applications. Proceedings of the 2005 Eight International Workshop on Principles of Software Evolution (IWPSE' 05), pp. 53-56.

[6] Fernandez-Montes, A. et al, 2009. Smart Environment Software Reference Architecture. Proceedings of Fifth International Joint Conference on INC, IMS and IDC, pp. 397–403. doi:10.1109/NCM.2009.115

[7] Karam, M. et al 2008. A product-line architecture for web service-based visual composition of web applications. Journal of System and Software., pp. 855-867.

[8] Kotonya, G. et al, 2009. A consumer-centred approach for service-oriented product line development. Proceedings of Second International Symposium on Computational Intelligence and Design, pp. 211-220.

[9] Lee, E., and Neuendorffer, S. (2004). Classes and Subclasses in Actor-Oriented Design Extended Abstract.Proceedings of Second ACM and IEEE International Conference. pp. 161-168.

[10] Lee, J. et al 2010. A feature-oriented approach for developing reusable product line assests of service-based systems. Journal of System and Software, pp. 1123-1136.

[11] Leung, M., and Lee, E. A. 2007. An Extensible Software Synthesis Framework for Heterogeneous Actor Models. Model-driven High-level Programming of Embedded Systems.

[12] Mili, H. et al, 2002. Reused-Based Software Engineering. Techniques, Organization, and Control, John Wiley and Sons, New York, USA.

[13] Medeiros,F.M et al, 2010. SOPLE: An Approach to design service-oriented product line architecture. Proceedings of Software Product Lines Conference, pp. 456-460.

[14] Moosavi, S. et al. 2009. A Method for Service Oriented Design. Proceedings of Sixth International Conference on Information Technology: New Generations /IEEE , pp. 290-295.

[15] Nakagawa, E. Y. et al, 2011. Exploring the Use of Reference Architectures in the Development of Product Line Artifacts. In Software Product Lines Conference 2011. Munich: ACM.

[16] Nguyen, T. and Colman, A. 2010. A feature-oriented approach for web service customization. Proceedings of IEEE International Conference on web services, pp. 383-400.

[17] Prado, A. B. and Baranauskas, M. C. C. 2012. Representing scientific associations through the lens of Actor-Network Theory. 2012 Fourth International Conference on Computational Aspects of Social Networks (CASoN), 87–92. doi:10.1109/CASoN.2012.6412383

[18] Salvaneschi, G. et al, 2012. Context-oriented programming: A software engineering perspective. Journal of Systems and Software, 85(8), 1801–1817. doi:10.1016/j.jss.2012.03.024

Hybrid Mobile Application Framework for Pervasive Health Record (HMAF4PHR)

Farah Aris
Faculty of Information & Communication Technology
Universiti Teknikal Malaysia Melaka,
Melaka, Malaysia
faraharis90@gmail.com

Mohd Khanapi Abd Ghani
Faculty of Information & Communication Technology
Universiti Teknikal Malaysia Melaka,
Melaka, Malaysia
khanapi@utem.edu.my

***Abstract*— Patient health records is important information for doctors to predict patient problems, advice and provide accurate medication to the patients. It is impossible that all healthcare facilities is equipped with EMR system and install with adequate network infrastructure that could link together all rural clinics, health centers, polyclinics and hospitals in integrated one. Pervasive healthcare could provide seamless access to patient health records where crucial patient health data could be store and retrieve to/from the pervasive devices. Once the health records could be stored and retrieved easily from person's gadget around his/her possession an environment where the connectivity is unobtrusive and always available can be created. The aim on this research is to develop and propose a validated hybrid mobile application framework for maintaining pervasive health record using smart phone. The research also identifies the best hybrid model and crucial patient health records datasets to be adopted in the proposed framework.**

***Index Terms*—Pervasive Healthcare, Mobile Application, Hybrid Framework**

I. INTRODUCTION

The handheld devices have been blooming and becoming a new trend and become advancement for the healthcare domain. The healthcare providers or healthcare professionals can utilize the technology for making the healthcare services one step forward to the advancement.

The healthcare records is the most crucial and important data that will be used by the medical practitioner for aiding their investigation on patient problems and providing accurate treatment to the patients. A pervasive health record is another step in the advancement of the healthcare domain and with a pervasive health record; the data can be accessed anytime and anywhere. From the research [1] it was stated that the pervasive health monitoring, intelligent emergency management system, pervasive healthcare data access and ubiquitous mobile telemedicine is some of the pervasive healthcare applications.

From the research [2] it was stated that the pervasive healthcare systems are in need due to the facts that the healthcare professionals experience a high level of mobility and they need to share resources and information with the staff, faculty and colleagues in a real time.

HMAF4PHR is a framework that will use a wireless connection to access data on the server. According to the research [3] it was stated that the wireless healthcare application allows user to healthcare workers to access health data and information that in turns can be useful for them to assist in a temporary prescription and emergency treatment when providing immediate access to health information

The research on the pervasive healthcare is beneficial as according to [4] it was stated that a number of impacts to clinical workflows that may be expected from a pervasive access to images, patient data, and analysis function becomes available via mobile and wireless device such as nomadic healthcare worker, radiologist and data explosion crisis and mission critical systems and the medical IT specialist

The aim on this research is to develop and propose a validated hybrid mobile application framework for maintaining pervasive health records using smart phones. The research also identifies the best hybrid model and crucial patient health records datasets to be adopted in the proposed framework.

II. HEALTH RECORDS PROBLEM AND ITS CONTEXT

This research tries to overcome some problem that associate to its contexts as follows:

a) Inadequate access to patient's health records - patient's health record is crucial information that will be ease the doctors to determine the conditions of the patients and to prescribe a suitable medicine and treatment for the patients. Without an adequate access to the patient's health record, it may lead to wrong treatment and may endanger the patient's life.

b) Homogenous medical records - electronic medical records (EMR) consist all type of data from texts, images, wave, audio, video and etc. It is impossible that these data could be stored in mobile devices and provide meaningful to healthcare professionals. An analysis of important EMR datasets needs to be conducted for identifying the related and meaningful data to be viewed by the doctor.

III. RESEARCH OBJECTIVE

The objective of this research is to (1) overcome the inadequate access to patient health care record that may lead to

wrong treatment for the patients; (2) expanding the technologies in the healthcare industry; (3) increase the knowledge in pervasive healthcare record; (4) facilitate the patient to access their own healthcare records; (5) propose and develop a hybrid architectural model that can access the pervasive health record by using a mobile application.

1) to analyze appropriate mobile application framework that could be incorporated in the proposed framework.

2) to analyse and identify important pervasive patient medical record.

3) to develop hybrid mobile application framework for pervasive health records.

4) to validate hybrid mobile application framework through case study approach at UTeM clinic.

IV. Related Work

According to [5] ubiquitous computing is enhancing the computer by making multiple computers available in the physical environment while making them invisible to the user. According to [6] pervasive healthcare can be support a wide range of services and application for example mobile telemedicine, patient monitoring, location-based medical services, emergency response and management, pervasive access to medical data, personalized monitoring and lifestyle incentive management.

The pervasive healthcare is aim to combing the pervasive computing for the healthcare purposes. The pervasive healthcare is aim to making it easier for patients to bring the medical records anywhere and can be access easily by the medical staff.

The previous research [7]has contributed many insights on what is user acceptance on different implementation concept of pervasive healthcare systems.

The research [7] is focusing on three type of implementation of medical technology. The first one is the technology implement on the mobile devices, second is the technology implements on the smart environments and third the technology implement on that smart clothing.

The outcome of the research [7] has concluded that the age, gender and health status did not impact the overall willingness of using the devices. The results have shown that the technology implement on the smart homes having more negative feedback while the technologies implement on the mobile devices having more positive feedback. The conclusion from the research [7] shows that most of the participants most likely to use technology implement on the mobile devices for the medical purposes.

From the research [7] it obviously show that most users are most comfortable on using the healthcare system that is implement on the mobile devices. Respond from this research the “Hybrid Mobile Application Framework for Pervasive Health Record (HMAF4PHR) is proposes to achieve the pervasive healthcare objective.

The research [8] proposed a Pervasive Mobile Healthcare that will accessing Electronic Medical Records (EMR) by using Elliptical Curve Cryptography (ECC) algorithm that will encrypt the data when transmitted to mobile application. The disadvantage on this research is the Pervasive Mobile Healthcare only work when the device is online and only works on Android operating system. Therefore the HMAF4PHR will be an improvement from this research in terms of proposes a hybrid framework that will work on most of the handheld device's operating system.

V. Research Methodology

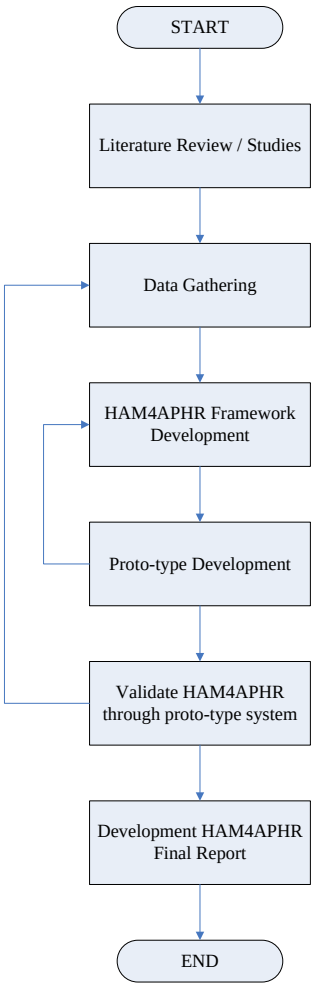

Figure 1: Research Methodology Flowchart

From the figure 1, it shows the research methodology flowchart. To conduct this research (1) an extensive literature review and studies need to be done. Afterwards, (2) data gathering that will be used in the research need to be done. After the data have been gathered the (3) HMAF4PHR framework will be developed. Then (4) based on the HMAF4PHR framework, the prototype will be developed. (5) The HMAF4PHR framework will be validated through the prototype system that has been developed. (6) To conclude the research a thesis will be developed.

VI. Milestones

TABLE I. Research Milestones

Milestone	Completion Dates
Studies in the literature review	December 2013
Data gathering	February 2013
HMAF4PHR framework	Jun 2013
Prototype development	November 2013
Validate HMAF4PHR through the prototype system	February 2014
Development HMAF4PHR final report	August 2014
Milestone	Completion Dates
Studies in the literature review	December 2013
Data gathering	February 2013
HAM4APHR framework	Jun 2013
Prototype development	November 2013
Validate HAM4APHR through the proto - type system	February 2014
Development HAM4APHR final report	August 2014

Table 1 shows the research milestones that is studies on the literature review, data gathering, HMAF4PHR framework, proto-type development, validate HMAF4PHR through proto-type system and the development HMAF4PHR final report. The duration of the research is two years.

Currently the research progress is on the studies in the literature review.

VII. Research Output

The HMAF4PHR development will facilitate patients to gain access to their own healthcare records by using a handheld device that is the Smartphone. An adequate access to patient's healthcare records can provide the medical practitioners to provide the suitable a suitable treatment to the patients.

The HMAF4PHR can bring a new technology to the healthcare domain.

Acknowledgment

This research is funded by Universiti Teknikal Malaysia Melaka under university short grant number PJP/2012/FTMK (540)/S01061.

References

[1] U. Varshney, "Pervasive Healthcare and Wireless Health Monitoring," *Mob. Networks Appl.*, vol. 12, no. 2–3, pp. 113–127, Jul. 2007.

[2] S. I. Ahamed, N. Talukder, and C. Science, "Towards Privacy Protection in Pervasive Healthcare," in *3rd IET International Conference on Intelligent Environments (IE 07)*, 2007, pp. 296–303.

[3] K. Wang, Y. Sui, X. Zou, A. Durresi, and S. Fang, "Pervasive and Trustworthy Healthcare," in *22nd International Conference on Advanced Information Networking and Applications - Workshops (aina workshops 2008)*, 2008, pp. 750–755.

[4] H. Pohjonen, P. Ross, J. G. Blickman, and R. Kamman, "Pervasive access to images and data--the use of computing grids and mobile/wireless devices across healthcare enterprises.," in *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, 2007, vol. 11, no. 1, pp. 81–86.

[5] M. West and C. College, "Ubiquitous Computing," in *ACM Conference on Computer Science*, 2011, pp. 175–181.

[6] U. Varshney, "Pervasive healthcare," *IEEE*, vol. 36, no. 12, pp. 138–140, Dec-2003.

[7] D. Kang, H. Lee, E. Ko, K. Kang, and J. Lee, "A Wearable Context Aware System for Ubiquitous Healthcare," in *Proceedings of the 28th IEEE EMBS Annual International Conference*, 2006, pp. 5192–5195.

[8] G. Sudha and R. Ganesan, "Secure transmission medical data for pervasive healthcare system using android," in *2013 International Conference on Communication and Signal Processing*, 2013, pp. 433–436.

Designing Hybrid Software Architecture Framework for Capturing Clinical Findings during Doctor-Patient Consultation

Omar Mukhtar Hambaran
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka
Melaka, Malaysia
kidzeclipes@yahoo.com

Mohd Khanapi Abd Ghani
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka
Melaka, Malaysia
khanapi@utem.edu.my

***Abstract*—The purpose of this paper is to purpose a hybrid software architecture framework for capturing clinical findings during doctor-patient consultation. Doctor-patient consultation plays a very important role in diagnosis of many clinical findings. Currently, most doctor-patient consultation used conventional ways of capturing clinical findings using paper's note, note book, manually entered digital records, and so on. With these conventional ways, the number of patient to be treated properly in the consultation process is less than the number of patients that had been registered per day. Framework in this paper use hybrid technique including of Fast Track Search that using SNOMED Clinical Terminology Code for doctors to search their common symptoms or medication, and technique that use seamless network accessing to upkeep a patient's health records continuously and seamlessly regardless what the state of the network downtime. Application of this proposed hybrid framework dramatically reduce the time taken for doctor in capturing patient's clinical record during the consultation process.**

Keywords: Hybrid Software Architecture Framework, Patient Health Record, SNOMED Clinical Terminology, Network Downtime.

I. Introduction

Doctor-patient consultation plays a very important role in diagnosis of many clinical findings. Currently, most doctor-patient consultation used conventional ways of capturing clinical findings using paper's note, note book, manually entered digital records, and so on. With these conventional ways, the number of patient to be treated properly in the consultation process is less than the number of patients that had been registered per day. This research will provide a better way to fasten the process of doctor-patient consultation using this purposed hybrid software architecture framework.

Currently, most patients always need to go to hospitals for consultation. Thus, because of many patients await their doctor; they need to wait till their name being called. Plus, patients who lived in rural area will suffer a lot from travelling frequently especially to have repeating consultation compare to those in urban area although those in urban area affected by other factor like traffic jam that caused the travel period became long. For the additional knowledge, the use of normal narrowband in a telemedicine system for transporting the patients' health records across healthcare facilities nationwide does not make sense due to issues of telecommunication reliability, speed and responsiveness. This research will provide a seamlessly solution for the doctor-patient consultation process in any condition of network downtime.

II. Aims of Investigation

The aim of this research is to design a hybrid software architecture that could be used for capturing patient health records during doctor-patient consultation.

III. Problem Statements

- The current system that has been computerized in most Hospital used high end bandwidth. Their web-based system depends hundred percent on the network communication. Adequate and stable bandwidth is important to facilitate this kind of application. With regards of the Health Care System, it use large amount of data that to be received or transferred into the central database. When the data is required by the doctor, the data will takes time to be downloaded since the network bandwidth size is low.
- The current application that used by the doctor produce slow respond time for the doctor to capture patient's health records.
- Due to the unstable network bandwidth, the data that currently been captured during in the middle of the consultation process will be lost.
- Non-standard clinical terminology used by the doctors that cause difficulty in generating EMR (Electronic Medical Records).
- The doctors' hand-writing of clinical notes need to be converted into coding system, and these task normally took long process to be completed.

IV. Research Objectives

- To conduct review on existing clinical terminology and existing software architecture framework in capturing clinical records, and compare and contrast between them.
- To conduct data collection on clinical terminology usage among medical professional in public hospital and data collection on clinical records that had been captured by a few hospital and clinic.
- To design and develop an application that used this purposed hybrid software architecture framework.
- To validate this purposed hybrid software architecture framework through pilot implementation in UTeM's clinic.

V. Findings

Within industrialized countries healthcare systems currently change to cope with the upcoming consequences of the demographic change. One of the most serious challenges is the maintenance of the area-wide supply chain of medical care despite the threatening shortage of physicians. In this context, telemedical services for communication between doctors and patients gain in importance. Crucial for the success of such electronic services is the choice of the medium, which must be appropriate for this special purpose of use and, finally, accepted by its users. A telephone consultation was similarly well accepted than face-to-face communication. As the most comfortable service a face-to-face-consultation was corroborated, followed by the videophone consultation [1].

Healthcare has been slow in catching up with the cutting edge technology for patient treatment and monitoring [2]. However, according to [3], the industry has hardly made use of Information and Communications Technologies (ICT) to support these practices. This is changing rapidly as the industry begins to realize the potential of ICT and increase opportunities for vendors, system integrators, service providers and resellers. In general, ICT can play following roles in healthcare:

- Provide the right knowledge, at the right place at the right time to support decisions.
- Support person to person, machine to machine and person to machine communications.
- Support group communication.
- Support data transfer chain by transforming data into information and then become knowledge [3].

Prompt access to a patient's electronic medical history or lifetime health record (LHR) is fundamental for providing seamless and continuous care. This can be achieved through the convergence of ICT, medical content and health knowledge. Through this convergence, the patient's medical history can be shared among healthcare professionals and across healthcare facilities regardless where the previous visit was made. However, the fragmented and paper-based medical records have significant limitations such as illegibility, unavailability, sheer physical volume (over time), difficult transferability and integration between providers and institutions, and the necessity to record the same data many times (duplication) on different documents [4] [5].

Electronic medical records (EMR) system such as clinical information system is used to capture, store and access patient health records (PHR) electronically. Electronic PHR would enhance healthcare services by providing prompt access to patient's longitudinal health records. The comprehensive information of longitudinal health records would help physician in providing accurate treatment and suggesting the right medication to patient. Hence, doctor-patient encounter could be enhanced and the doctor could improve their interaction with the patient effectively [6].

Medical imaging plays a very important role in diagnosis of diseases. Tele-consultation between doctors would become more meaningful if a tool is available that not only supports verbal communication but displays medical images at their consoles with enriched features such as marking, text annotating, zooming etc. Implementing such a tool over the Internet is a challenging task as it involves transfer of different types of contents of large volume of data between the clients. Any such exchange is also fraught with the problems of data security [7].

A modern paper version of a clinical case history very often contains an extensive collection of diverse documents, which provide information on clinical findings from medical experts, a vast array of results of laboratory and other tests, and the documentation of clinical course and outcome. This compendium of information generally represents the accumulation of data from a large number of health care specialists involved in a patient's care, often without the benefit of any integration or mutual communication. The end result can be a large folder of disjointed information which is difficult to process even by a primary care provider, much less by a consultant specialist or other physician less familiar with the individual case [8].

The importance of doctor–patient interactions and their potential influence on patients' behavior and well-being has become widely recognized. In studying the influence of such interactions, research has focused primarily on physicians' communicative behaviors, whereas so called ''patient outcomes'' have been used to assess the extent of these influences [9].

Currently the Internet is now accessible using a mobile smart phone. According to research conducted by Insight MarkPlus of 2161, mentioned that the number of internet users in Indonesia in 2011 has reached 55 million people, and showed an increase in 2010 which reaches 42 million. Mark Plus Insight Research shows that average Internet users in Indonesia access by using the smart phone and notebook. Mobile Internet penetration in Indonesia is at 57 percent. In terms of population data from Insight Mark Plus mentioned

that currently there are 29 million mobile Internet users in Indonesia [10].

VI. Methodology

A. Provide Findings

- Collect all journals/articles/books that related to this research.
- Summarize all the information from journals/articles/books that related to this research.

B. Conduct Preliminary Data Collection

- Conduct primary data collection of clinical findings during the process of doctor-patient consultation.

C. Design Prototyping

- Planning and designing the conceptual Framework for this research.
- The purposed Framework is illustrated in Fig. 1.

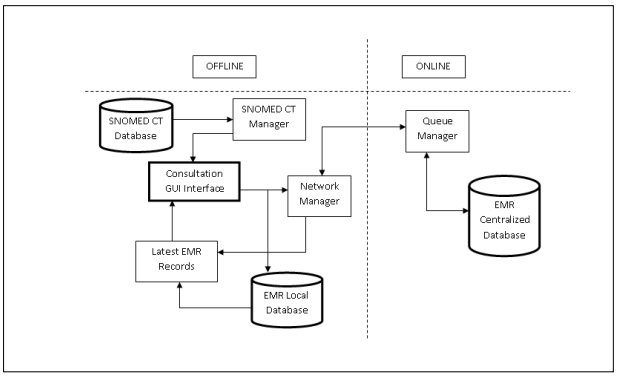

Fig. 1. Purposed Framework of Software Architecture in Capturing Clinical Findings in Doctor-Patient Consultation

1) SNOMED CT Manager: This manager will provide SNOMED Clinical Terminology code for the purposed system. When the doctor want to search symptoms or diseases, this system will provide the common result that had been searched by other doctor or from his/her previous consultation process.

2) Latest EMR Records: This function will provide the system patient's latest electronic medical record to the system. This record will be used by the doctor to ease his/her consultation process with the patient.

3) Network Manager: This manager will check the connection between client and server side of databases. This will determine which databases need to be used when network downtime is occurs. It also will sync all the data from the server databases into local databases when needed.

D. Develop a System based on the Purposed Framework

- Develop a system based on the purposed Framework of software architecture that related to this research.
- Debug and do software testing to the system.
- In this phase, the system will be developed using Java Programming Language.
- The online database is using MySQL database, and it will be centralized in one server.
- The offline database is using HSQL database, and it will be stored in each client's workstation.

E. Validate and Evaluating the System

- The evaluation is undertaken to understand system performance, improve clinical performance and patient outcomes, and understand the unpredictable characteristics of network downtime in any general condition.
- After the system has been computerized, it will be tested in the a few hospitals or clinics that have outpatient consultation process.
- Several conditions of patient's medical records will be captured.

F. Analyzing the Data

- The result and outcome data from this research will be analyzed for a comparison study.
- The data will be compared based on the time taken in the consultation process, and the quantity data that can be captured by the system.

VII. Expected Outcome

The outcome of this research is a software/system that use this purposed hybrid software architecture framework to reduce the usage of time taken by the doctor and amount of patients' electronic medical records been captured during doctor-patient consultation.

VIII. References

[1] S. Beul, M. Ziefle, and E.-M. Jakobs, "Users' Preferences for Telemedical Consultations: Comparing users' attitude towards different media in technology-mediated doctor-patient-communication," *Proceedings of the 5th International ICST Conference on Pervasive Computing Technologies for Healthcare*, p. 7, 2011.

[2] T. K. Chuan, N. Zakaria, and M. A. Osman, "User Requirements for Online Consultation in Cancer Patient Portal," *2010 IEEE EMBS Conference on Biomedical Engineering & Sciences (IECBES 2010), Kuala Lumpur, Malaysia*, no. December, p. 6, 2010.

[3] B. G. Eastwood, "ICT Opportunities in Healthcare Table of Contents Gary Eastwood," p. 10, 2005.

[4] M. K. A. Ghani, R. K. Bali, R. N. G. Naguib, I. M. Marshall, and a. S. Shibghatullah, "The design of flexible front end framework for accessing patient health records through short message service," *2007 Asia-Pacific Conference on Applied Electromagnetics*, pp. 1–5, Dec. 2007.

[5] I. Román, L. M. Roa, J. Reina-Tosina, and G. Madinabeitia, "Demographic management in a federated healthcare environment.," *International journal of medical informatics*, vol. 75, no. 9, pp. 671–682, Sep. 2006.

[6] M. Khanapi and A. Ghani, "Comparing Speech Recognition and Text Writing in Recording Patient Health Records," *2012 IEEE EMBS International Conference on Biomedical Engineering and Sciences, Langkawi*, no. December, p. 6, 2012.

[7] D. D. Prasad, S. Ray, A. K. Majumdar, J. Mukherjee, B. Majumdar, S. Paul, and A. K. Verma, "Through Internet," *Journal of Healthcare Engineering*, vol. 1, no. 1, pp. 141–154, 2010.

[8] N. Shklovskiy-Kordi, J. Freidin, S. Goldberg, M. Krol, N. Rivkind, B. Zingerman, and S. Davis, "Standardization for telemedical consultation on a basis of multimedia case history," *Proceedings 14th IEEE Symposium on Computer-Based Medical Systems. CBMS 2001*, pp. 535–540, 2001.

[9] L. M. Ong, M. R. Visser, F. B. Lammes, and J. C. de Haes, "Doctor-patient communication and cancer patients' quality of life and satisfaction.," *Patient education and counseling*, vol. 41, no. 2, pp. 145–156, Sep. 2000.

[10] D. P. Irsan and J. Simarmata, "Design and Implementation of Mobile Health Consultation between Doctors and Diabetes Patient," p. 4, 2011.

Health Informatics Framework for Postnatal Care:

An integration between Modern Medicine and Traditional Malay Medicine

Raja Ikram, R. R., Abd Ghani, M.K. , and Abdullah, A.

Biomedical Computing and Engineering Technologies (BIOCORE) Applied Research Group,
University Teknikal Malaysia Melaka (UTeM),
Melaka, Malaysia.

rajarina@gmail.com, {khanapi;noraswaliza}@utem.edu.my

***Abstract*— The overall aim of this research is to produce a validated framework of the integration of Traditional Malay Medicine (TMM) with modern medicine in the field of postnatal care or famously known by the Malay community as Malay Confinement. TMM has been widely used within the Malaysian community for the past decades and has been recognised by the Ministry of Health, Malaysia as part of the traditional and complementary medicine services in Malaysia. Despite this, Malay Confinement has only been partially recognised as part of the official healthcare system, even though there are separate wellness centres or postnatal services provided by private organisations. This research shall initially examine the current traditional and complementary medicine implementation in the Malaysian healthcare system and other countries and highlight challenges and lessons learnt based on other country's experiences. The research shall than focus on the current implementation and requirements of modern and traditional medicine in the postnatal care in the Malay community. This research shall attempt to identify the issues related to the implementation and propose a framework to integrate current modern medicine with traditional Malay medicine in postnatal care. This framework is expected to provide a guideline to assist the healthcare community to promote and introduce Traditional Malay Medicine services, and thus recognizing it as part of an official medical treatment with proper health records as evidence. This framework is also expected to assist healthcare software developers in the software development lifecycle, particularly in the requirements phase.**

***Index Terms*— Traditional Malay Medicine, postnatal care**

I. INTRODUCTION

Traditional and Complementary Medicine (T&CM) has been gaining acknowledgement and acceptance all over the world. T&CM is an invaluable treasure and has been developed over the course of thousands of years in the quest for human wellbeing. It is a form of health- related practice designed to prevent, treat, or manage illness and preserve the mental and physical well-being of individuals. This practice includes traditional Malay medicine, traditional Chinese medicine, traditional Indian medicine, homeopathy and complementary therapies, and excludes medical or dental practices by registered medical or dental practitioner [3]. There are nine integrated public hospitals which are practicing T&CM in Malaysia. They are Kepala Batas Hospital in Pulau Pinang, Putrajaya Hospital in Putrajaya, Sultan Ismail Hospital in Johor Bharu, Duchess of Kent in Sabah, Sultanah Nur Zahirah Hospital in Kuala Terengganu, Sarawak General Hospital, Port Dickson Hospital in Negeri Sembilan, Sultanah Bahiyah Hospital in Alor Setar Kedah and Cameron Highlands Hospital [4]. These hospitals practice traditional Malay massage, acupuncture, herbal oncology and postnatal massage [3]. Traditional Malay massage and acupuncture are used for chronic pain and stroke; herbal oncology is used for complement treatment with allopathy therapy whereas postnatal massage is used to relieve muscle cramps and fatigue after labour [4].

Postnatal care is essential because newborn survivors are directly linked to the mother's health [6]. The postnatal period (the time just after delivery and through the first six weeks of life) is especially critical for newborns and mothers because global statistics show that more than two thirds of new born deaths occur within the first week after delivery [7] and two thirds of all maternal deaths occur in the postnatal period [8]. Thus, integrated maternal and newborn postnatal care during the postnatal period should be provided as a concerted strategy to improve survival of both.

II. RESEARCH PROBLEM

Even though Traditional Malay Medicine is practised in these hospitals, however majority of the TMM practice is limited to Traditional Malay massage. However, the Traditional & Complementary Medicine (T&CM) Unit of Putrajaya Hospital and Johor Baru's Sultan Ismail Hospital midwifery care practices postnatal massage, hot compress or *bertungku* and herbal body wrap or *bengkung* [1]. Traditional Malay massage is part of the processes involved in Malay Confinement. Malay Confinement is essentially an all-encompassing process that aims to preserve the health and femininity of Malay women. According to Datin Sharifah Anisah, founder of Nona Roguy (now NR) and author of *Ensiklopedia Perbidanan Melayu* (Encyclopaedia of Malay Midwifery), confinement practices stem from the belief that the womb is a woman's life force and affects her overall health [1]. There are also many Malay Confinement centres for mothers to rehabilitate and rest after childbirth, including Confinement lady services providing Malay Confinement services from home. Today, many urban new mothers in the

Malay community find themselves unable to fully observe the practices of the traditional Confinement, mainly because they lack the family and community support that made Malay Confinement possible. However, some practices are still observed by many Malay women. Some have been adapted to suit more urban lifestyles. Among the Malay Confinement commonly practiced are *bengkung* (the traditional wrap or girdle), postnatal massage, *jamu* (traditional made supplements), hot compress and herbal baths [1]. These complete Malay Confinement practises are not fully implemented by Malaysia's official healthcare system or government hospitals. However it is supported by many private and individual confinement care services that provide this alternative for new mothers. Thus, there is an isolation of Malay Confinement implementation in the government or official healthcare system and private hospitals or wellness centres. Private hospitals that implement midwives for home visits also would usually outsource the Malay Confinement practices to individual midwives without proper documentation of health records. This research shall attempt to propose a framework to integrate Malay Confinement practices in the Malaysian healthcare system, so its practices can be recognized as part of an official medical treatment with proper health records as evidence.

III. Research Objectives

This research shall examine the possible integration frameworks for the harmonisation of Traditional Malay Medicine and modern medicine in the field of Malay Confinement. The objectives of this research can be summarised as follows:

- To investigate requirements for modern medicine and traditional Malay medicine for postnatal care.
- To collect data for the current implementation of Malay Confinement in the Malaysian healthcare system
- To develop a framework to integrate modern medicine with Traditional Malay Medicine in the field of postnatal care.
- To validate the initial framework.
- To finalise the framework based on the validation findings.

A brief description of the research objectives are in the following subsections.

A. . To Investigate Requirements for Modern Medicine and Traditional Malay Medicine for Postnatal Care.

To analyse and define requirements for modern medicine and Traditional Malay Medicine for postnatal care in terms of diagnosis, prescription and treatment. According to World Health Organisation [9], postnatal care for newborns should include immediate and direct breastfeeding, warming of infant, hygienic care of umbilical cord, and timely identification of danger signs with referral treatment, especially in low birth weight babies. Other key elements for postnatal care include the optimum number and timing of postnatal care visits for first contact, follow up visits and extra visits for low birth weight babies [10]. The results of this objective contain vital information to obtain requirements for the critical data sets required to develop the framework. This objective shall be achieved through review of secondary sources.

B. To Collect Data of the Current Implementation of Malay Confinement in the Malaysian Healthcare System

Data collection shall include a few case studies of hospitals that implement Traditional Malay Medicine in Malaysia, and postnatal care wellness centres. Common Malay confinement practices include *bengkung* (the traditional wrap or girdle), postnatal massage, *jamu* (traditional made supplements), hot compress or *tungku* and herbal baths [1]. Hot compress or *bertungku* is also believed to help the womb contract, break down fat and help the body return to its pre-pregnancy state. Nutrition and dietary supplements (commonly known as *jamu*) is believed to boost energy levels, keep the body warm, remove excess fluids, dispel wind and slim down the tummy [1]. Critical issues to be established for Malay Confinement practices are what are the key elements of a routine postnatal care, number of optimum postnatal visits, where should postnatal care be provided, and who should provide it [10].The result of this data collection shall significantly contribute to the design and development of framework.

C. To Develop a Framework to Integrate Modern Medicine with Traditional Malay Medicine in the Field of Postnatal Care.

The findings from the data collection of primary and secondary sources shall be analysed and the requirements shall be refined to design and develop an initial framework. This framework shall take into account the critical data sets to support and standardize electronic health records as evidence of an official medical treatment. Thus, this framework indirectly acknowledges the significant contribution of Traditional Malay Medicine in the healthcare community by recognizing it as an official medical treatment with evidence of electronic health records.

D. To Validate the Initial Framework

The initial framework shall then be validated by developing a prototype application of the proposed framework and validating the application with subject matter experts. Feedback shall be obtained based on open and closed ended questionnaires. The results of the feedback questionnaire shall be tabulated and elaborated to ensure every element of the framework is verified.

E. To Finalise the Framework Based on the Validation Findings.

The results of the validation phase shall be discussed and analysed, and improvements of the initial framework shall be proposed where applicable.

IV. Research Questions

This study shall attempt to reveal the answers for the following issues.

A. *Key Research Question 1: What are the critical data sets required to support and standardize the health informatics framework for modern medicine in postnatal case studies?*

This research question shall be addressed by investigating the requirements for postnatal care in modern medicine in the analysis and requirements phase. The answer to this research question shall consider data collection from primary and secondary resources, as well as review and analysis of existing publications.

B. *Key Research Question 2: What are the critical data sets required to support and standardize the health informatics framework for Traditional Malay Medicine in postnatal case studies?*

This research question shall be addressed by investigating the requirements for postnatal care in Traditional Malay Medicine. The answer to this research question shall consider data collection from and not limited to health practitioners, implementation of TMM in government and private healthcare centres.

C. *Key Research Question 3: What is the best integration framework for modern medicine and Traditional Malay Medicine to be implemented in the postnatal care domain?*

The answer to this research question shall be addressed in the design and development of framework phase in this study.

V. Literature ReView

A literature review shall be conducted on a few issues related to the existing system. The main issues that shall be covered in the literature review are evaluation of implementation of traditional medicine frameworks, review and analysis of traditional and complementary medicine implementation in Malaysia and other countries, a review of the implementation of modern medicine in the postnatal domain and an analysis of the implementation of traditional Malay medicine in Malaysia in the postnatal domain.

An evaluation of traditional and complementary medicine implementation framework shall be investigated to identify existing framework issues, similarities and possibly propose framework improvements. These framework reviews shall cover the requirements to implement traditional and complementary medicine within a national healthcare system. Some issues that shall be explored are national policies, control of clinical trials, registration and quality control [2]. The outcome of this section is to identify a general framework and important data required for implementing traditional and complementary medicine. This outcome shall possibly contribute to the design of electronic health record (EHR) and other system integration requirements to integrate with the current modern medicine implementation.

A review of implementation of traditional and complementary medicine in Malaysia and other countries shall also be executed to assess the relevance of the framework identified. The purpose of this review is to highlight the reality of implementation of traditional and complementary medicine frameworks and its challenges. The results of this review shall also be compared with the implementation of traditional and complementary medicine in Malaysia. Amongst the traditional and complementary medicine practises recognised by the Ministry of Health are Malay Traditional Medicine, Chinese Traditional Medicine, Indian Traditional Medicine, Homeopathy, and Complementary Medicine [5]. The progress of establishment of traditional and complementary medicine began in 1992, with the registration of traditional and complementary medicine, followed by the award of Good Manufacturing Practice (GMP) licenses to traditional medicine manufacturers in 1997 and the formation of National Policy of Traditional and Complementary Medicine in 2001 by the Ministry of Health (MOH) [5]. The outcome of this review is to provide a comparison and highlight challenges and lessons learnt with relation to implementation of traditional and complementary medicine worldwide. The results shall then be evaluated to propose improvements in the current implementation of traditional and complementary medicine in the Malaysian healthcare system. The significance of this review shall contribute to the proposed framework for traditional Malay medicine in the postnatal domain.

The literature review shall then focus on the main scope of study by conducting a review and evaluation of the requirements of modern medicine in the postnatal domain. The expected outcome of this assessment shall contribute to the design of data sets and electronic healthcare record of a clinical system in the postnatal domain.

Finally, an analysis of implementation of traditional Malay Medicine in the postnatal domain in the Malaysian healthcare system shall be conducted. This analysis shall highlight the requirements of traditional Malay medicine in the postnatal domain, or more famously known as Malay Confinement. A review of the current implementation of postnatal care in Malaysia shall also be studied. The requirements identified are expected to contribute to the design of data set, electronic healthcare record (EHR) and other integration requirements with the existing Malaysian healthcare system. The gaps of the implementation of Malay Confinement in the Malay community and Malaysian healthcare system shall be highlighted, and improvements shall be proposed to comprehensively integrate both traditional and modern medicine.

The results for the findings from the literature review shall be utilised to design the questionnaire and semi structured interview questions in the design and development of framework phase of this research.

VI. Methodology

The methodology section shall highlight the methods and instruments used throughout the study. The methodology used shall be divided into four main sections. The first section is an analysis of existing system or literature review of the study. Second, is the design and development of proposed framework. Third, the validation phase and finally, the discussion and revision of the initial proposed framework.

A summary of the research methodology can be referred to in Figure 1.

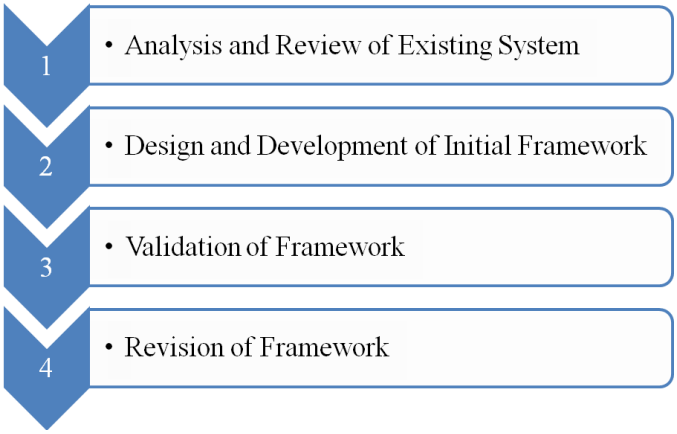

Fig. 1. Research Methodology Process

The literature review shall provide and analysis of existing system and shall be divided to four main research areas. First, an evaluation of implementation of traditional medicine frameworks in the national healthcare system. Second, a review and analysis of traditional and complementary medicine implementation in Malaysia and other countries. Third review of the implementation of modern medicine in the postnatal domain and the final research area shall be an analysis of the implementation of traditional Malay medicine in Malaysia in the postnatal domain.

The design and development of framework phase shall involve data collection from primary and secondary sources. The method for data collection shall be based on questionnaire and semi structured interviews. Data collection shall be conducted in three main areas. First, diagnosis, treatment and prescription for modern medicine in postnatal care. Data collection shall be conducted to understand the data sets and electronic health record required for any diagnosis, treatment and prescription of drugs. Second, diagnosis, treatment and prescription required for postnatal care in Malay Confinement. Data collection shall be conducted to understand the data sets and electronic health record required as evidence of medical treatment. It is crucial to understand the requirements for both modern and Traditional Malay Medicine before any framework integration can be successfully proposed. Data collection shall also be carried out to understand the implementation of traditional Malay medicine in Malaysian government and private hospitals and clinics to identify the current gaps in the implementation of Malay Confinement in the national healthcare system. The outcome of the design and development phase is expected to contribute to a framework to comprehensively integrate both modern and traditional Malay medicine in postnatal care and suggest improvements to its implementation in the national healthcare system.

The validation of framework phase shall involve a development of prototype application of the initial framework. This prototype application shall then be validated via open and closed ended questionnaires with subject matter experts in the software engineering and healthcare field.

The revision of framework phase shall utilize the validation results to revise the initial proposed framework to ensure its relevance and acceptance in the healthcare community.

VII. Expected Contribution

The outcome of this research is expected to contribute to the domain of traditional medicine in health informatics, and provide a framework to integrate Traditional Malay Medicine and modern medicine in the Malaysian healthcare system. This framework may provide a guideline to assist the healthcare community to promote and introduce Traditional Malay Medicine services, recognizing it as part of an official medical treatment with proper healthcare record as evidence. This framework is also expected to assist healthcare software developers in their software development lifecycle, particularly in the requirements phase.

VIII. Significance of Contribution

This research shall significantly promote Traditional Malay Medicine to the global health informatics community, indirectly highlighting the diverse traditional and complementary medicines available in Malaysia. This also encourages and promotes healthcare tourism and enables Malaysia to be recognised as a chosen location for complementary and alternative healthcare services to foreign tourists.

IX. Conclusion

The outcome of this research shall assist the implementation of integration of modern medicine and Traditional Malay Medicine from a technology perspective in the Malaysian healthcare system. Malay Confinement should be recognised as an official part of the healthcare system due to its extensive implementation among Malaysian mothers, immense availability of supplements in the market which have been registered with the Ministry of Health, and establishment of many private services and wellness centres.

References

[1] Babycenter, 2013, "Malay Pantang: An Overview". Available from http://www.babycenter.com.my/A1036820/malay-pantang-an-overview (Accessed 12 July 2013).

[2] Jayaraj P., 2010, Regulation of traditional and complementary medicinal products in Malaysia*", International Journal of Green Pharmaceuticals* ; vol 4:10-4.

[3] Ministry of Health Malaysia, 2013, "Traditional medicine or complementary medicine. What is it?". Available from: http://tcm.moh.gov.my (accessed 10 July 2013).

[4] Maihebureti A., Sharifa Ezat W. P., Syed Aljuni, 2011, "Role of Traditional And Complementary Medicine In Universal Coverage", *Malaysian Journal of Public Health Medicine*, Vol. 11(2): 1-5.

[5] Pillay, M.S., 2006, "Traditional and Complementary Medicine in Malaysia", *International Conclave on Traditional Medicine*, New Delhi, India, 16-17 November 2006.

[6] Sines W., Syed, U., Wall, S. and Worley, H., 2007, "Postnatal Care: A Critical Opportunity to Save Mothers and Newborns", Population Reference Bureau, Washington, USA.

[7] Joy E. L., 2006, "Analysis based on 38 DHS datasets (2000 to 2004) with 9,022 neonatal deaths", Save the Children-U.S., State of the World's Mothers 2006, Washington, USA.
[8] Ronsman, C. Graham, W. J., "Maternal Mortality: Who, When, Where, and Why? Maternal Survival,"The Lancet Maternal Survival Series (2006): 13-23.
[9] World Health Organization, 1998, Postpartum Care of the Mother and Newborn: A Practical Guide, World Health Organisation, Geneva.. Switzerland.
[10]Warren, C. Daly, P., Toure, L., Mong, P., 2010, "Postnatal care", Opportunities for Africa's Newborns, World Health Organisation, Geneva, Switzerland.

Enhancing Generic Pipeline Model in Preventing Code Clone during Software Development

Al-Fahim Mubarak Ali, Shahida Sulaiman
Department of Software Engineering
Faculty of Computing
Universiti Teknologi Malaysia
81310 UTM, Skudai, Johor, Malaysia.
fahimbhai86@gmail.com, shahidasulaiman@utm.my

***Abstract*—Code clone is an issue not only in maintenance but also development of a software. Code clone happens mainly due to copy and paste method applied while writing the source code during software development. There are works that been done in reducing code clone during software development. Therefore, this work looks into preventing code clone during software development by enhancing the generic pipeline model. The outcome of the enhancement is to have a preventive approach that can prevent code clone during software development and at the same time assisting software developers in writing code clone free software.**

***Index Terms*—Code Clone, Code Clone Prevention Approach, Software Development, Generic Pipeline Model.**

I. INTRODUCTION

With the world growing rapidly in knowledge and technology, software is developed to ease users in their daily life. As the change of time occurs, these software need to be maintained to sustain the quality and relevancy of software according to change of time technology. Maintenance of software system can be defined as a software product modification after implementation of the software to improve performance or to adapt the products to a modified environment [1]. Clones in software, or better known as code clone, is one of the factors that affect maintenance of a software [2].

It is essential to know that clones occur due to different scenarios. Stan and Jarzabek mentioned that there are three scenarios contribute to cloning [3]. The first scenario is clones caused by poor design. These clones can be removed by replacing it with functions or through the refactoring process but technically there will be risks that might cause the clone removal process difficult. The second scenario is between long-lived clones and temporary clones. Long-lived clones are clones that have existed in a program for a time while temporary clones only exist during the creation of the program. The third scenario is an essential clone that cannot be eliminated from the program. This scenario happens if the elimination of the clone affects quality of program, simplification as a prime reason for clone removal and complexity of the program is due to the restriction of programming language or design techniques approach.

Code clone detection is the most critical activity as this is the activity where code clones are detected by applying certain detection approaches and models. The detection approaches are distinguished based on information that can be obtained from the approach [2]. These approaches are applied as part of a process in code clone detection tools [4]. There are five major code clone detection approaches which are text based comparison approach, metric based comparison approach, tree based comparison approach, token based comparison approach and graph based comparison approach. Table I shows the description of each approaches.

TABLE I. MAJOR DETECTION APPROACHES AND TECHNIQUES [5] [6]

Approach	Description
Text based comparison	The code clones in a program are compared textually
Metric based comparison	Source codes are formed into vectors. These vectors are then compared to find clones.
Tree based comparison	Code clones are parsed into tree syntax using a parser. These syntax are then compared to find similarity between them.
Token based comparison	Code clone are transformed into tokens that are formed from lexical analysis. These tokens are compared and analyzed to find similar clone pattern.
Graph based comparison	This approach uses graph based techniques to find code clone. Example of graph technique is Program Dependency Graph.

Apart of the code clone detection approaches, there is also model used to detect code clone. Generic pipeline model is a code clone detection model that uses combination of processes to detect code clone. It is flexible yet extensible code clone detection process contains all required steps in a clone detection process. There are five processes involved in this

generic pipeline model [7, 8]. Figure I show the overall view of the generic pipeline model.

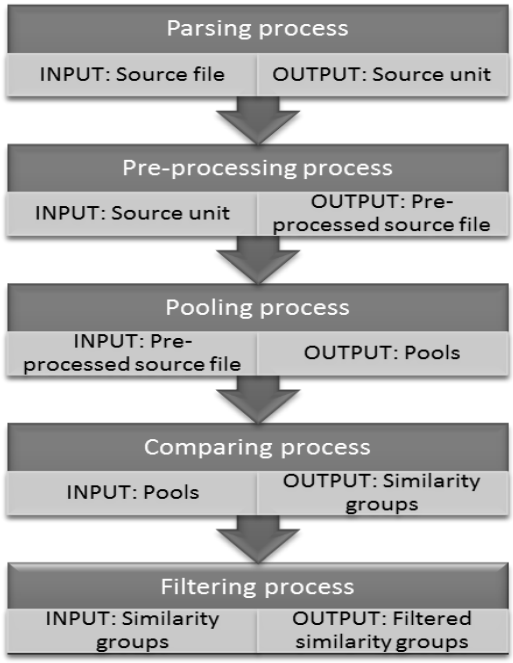

Fig. I. Generic Pipeline Model [8]

Parsing process is the first process in the generic pipeline model. A process transforms source code into source units. Source unit indicates start and end of a fragment in a source file. The source units then enter the pre-processing phase. Pre-processing process is the second phase in the generic pipeline model. It is a process to normalize source units and to add additional annotations to the source units. Normalization turns the source units into a regular form and makes different source units more similar.

Pooling process is the third process in the generic pipeline model. This process comes after the pre-processing process. It is a process of grouping pre-processed AST source units into sets of groups according to defined characteristics based on criteria set by the user. The sets are called as pool. The user defined criteria are characteristics that can be directly read from the source unit and its annotations without comparing it to another one. The pools then enter the fourth process which is comparing process.Comparing process is a process of recursive comparison of source units in all pools using a divide and conquer strategy. Pooling process previously allows the comparing step to apply a divide-and-conquer strategy. The final process which is the filtering process is then used to remove irrelevant clone candidate sets from the result set. This process is utilized in removing non-relevant candidate sets out of the result set.

Code clone detection approaches is not only used for clone detection purposes. It is also used to prevent code clone. Code clone prevention is an activity revolves in applying prevention control that is complemented with integration of problem mining approach in the prevention control [2]. It also can be said as a measure taken to continuously detect code clone while the code is being modified. It also encompasses action taken in avoiding new code clones [9].

There are two main characteristic involved in a code clone prevention approach which are preventive control and problem mining [5]. Preventive control is a continuous process that is aimed to control the introduction of newly added code clones in the code repository or during the system development. The new clones are added if it is deemed necessary. This process depends on the technology of code clone detection in detecting identical codes. Problem mining is a complementary process for the preventive control. It is a process aims to monitor the clone under modification with other clones that exist in system or code repository during development. The clone is searched throughout the system and if found, is deemed for modification.

Therefore this work looks into the aspect preventing code clone during software development. The goal of this research is to assist newbie programmers in reducing code clone occurrences during software development. The remaining five sections are as follows. Section II presents the motivation of this paper. Section III shows the related work of this study. Section IV shows the research procedure that is being used followed by Section V explains the proposed workand finally Section VI describes the expected results.

II. Background Of The Problem

Due to the rapid development in the world of computers today, complexity of the software development is different between one and another. As the complexity of the program increases, programmers or software developers tend to avoid these complexities by making multiple copies of codes or better said as code clones in their programs. Code clone happens due to the programmers' tendency to adopt copy paste method to solve the complexity of software [10, 11].The programmers usually follow the copy-paste technique instead of costly redesigning approach to avoid the complexity, hence causing code clones. Programmers tend to apply code clone as it is the most basic and widely used approach towards software reuse [10]. Some other reasons programmers tend to clone codes is to speed up development process. This occurs when a new requirement is not fully understood and a similar piece of code is present in the system that is not designed for reuse [11].

A preliminary study was done in understanding problems in the code clone detection phase and modification phase among programmers especially novice programmers [12]. The outcome of the study concluded that for novice programmers:

i. They may not be aware of the existence of code clones during the software maintenance phase.
ii. They tend to revise only defective codes they found first and not to search and revise their clones.
iii. It is difficult for the them to decide whether they should revise the files even if they are able to search for those files.
iv. It is difficult for a novice programmer to search for his target precisely with low cost. It is even more difficult to do it in large-scale legacy software

because the target is vast and the terms are not standardized.

Based on the outcome of the preliminary study done, it shows the newbie programmers have difficulty maintaining software due to the lack of knowledge and awareness regarding code clone. The practice that is adopted by them clearly does not reduce code clone but might even more code clones to occur. It is clear that it is hard for the newbie programmers to track and change code clones in large scale software systems due to non-standardization used in that large scale software.

III. Related Work

There are various approaches has been applied to prevent code clone during software development. The prevention approaches that had been used are not dependent only on the mentioned major approaches.

An experiment was done to prove the need of preventing clone during software development [9]. The experiment proved there is a need of embedding code clone preventive mechanism and problem mining during software development to avoid code clone during software development.

Linked Editing was used during software development to control code clone occurrence in development using a development tool called CodeLink [13]. It was implemented as plug-in for the tool. The plug-in was developed to detect clone from Java.

Clone Tracker [14] is a tool developed using tree based approach to track the change of codes during software development. This Eclipse plug-in tracks the change made by the developers when developing Java software.

Clone Detective [15] was introduced for Visual Studio environment development to detect code clones. This Visual Studio plug-in uses token based approach to is implemented using C++ and C#. It tracks and notifies users regarding code clones during the development of the software.

SHINOBI [12] is also another Visual Studio plug-in that is used to detect code clone. This tool detects and notifies developers of code clones during maintenance of a software. It uses token based approach to detect the code clones.

IV. Research Procedure

Figure II shows the research procedure that is being applied for this work. There are four phases involved in this work which are gather information on domain, design and develop the prototype, test and evaluate the prototype and report write up.

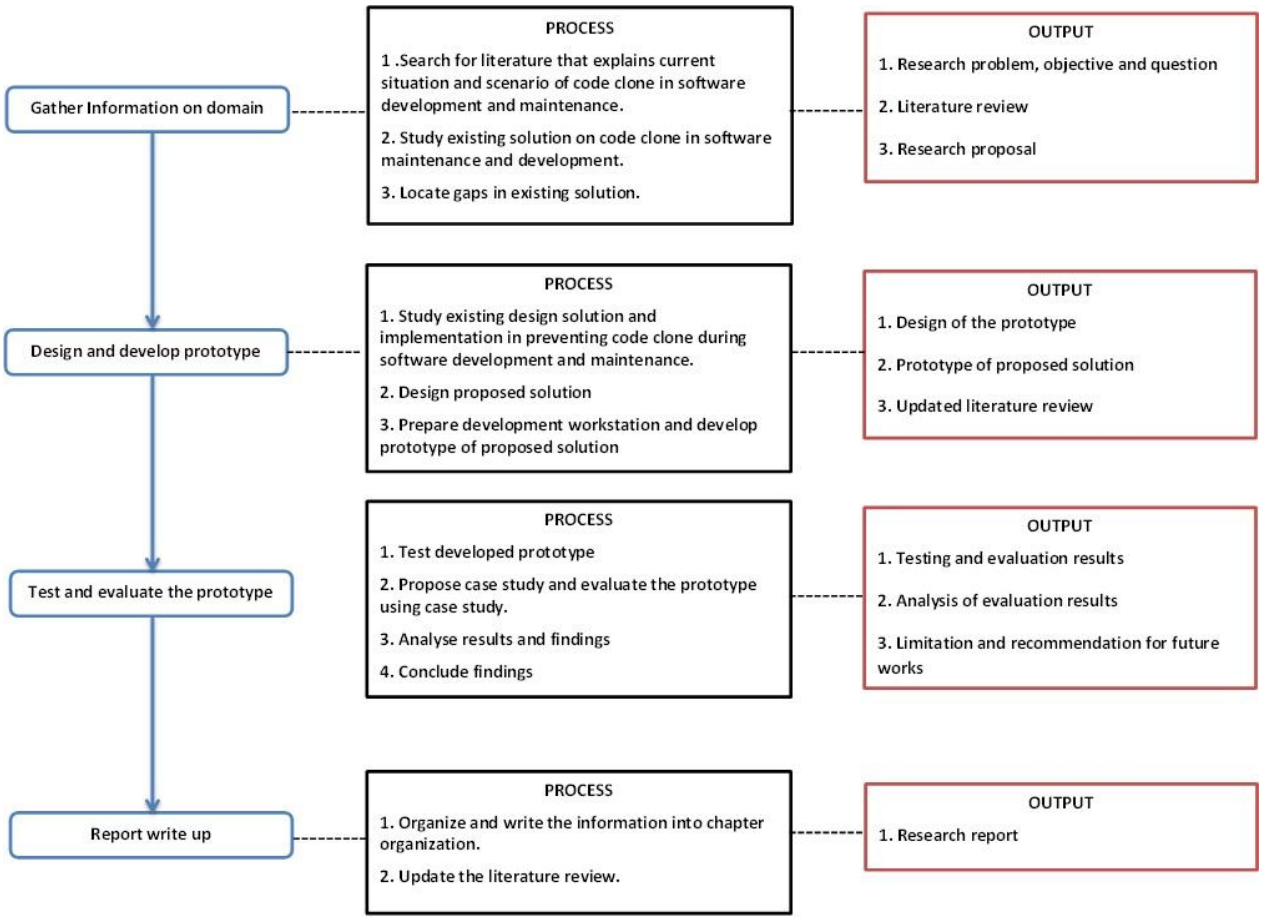

Fig. II. Research Procedure

The first phase is gather information on domain. The purpose of this phase is to create a solid foundation for the proposed research which is code clone in software development and software maintenance. Firstly, information that that explains current situation and scenario of code clone in software development and software maintenance will be gathered. This information can be obtained from sources such as journals, conference papers, books, websites and review papers from the internet and also electronic databases such as Thompson Reuters, Google Scholar, IEEEXplore, Springerlink and Science Direct. Secondly, based from the gathered literatures, the existing solution will be studied and critically reviewed. The review process is to search for issues and requirement, analysis of previous work and locate the gaps in existing solution that can be used in forming the research problem, research objective, research question, literature review and research proposal.

The second phase is design and develop prototype. The purpose of this phase is to provide the design and develop the flow and architecture of the proposed work. Firstly, the existing design solution and implementation in preventing code clone during software development and maintenance will be studied. The requirements that are been analyzed from the previous phase is also used in this phase. Secondly, the design and architecture of the prototype will be determined based on the gathered requirement. The architecture of the prototype will be drawn together with the flow of the prototype. Thirdly, a development workstation will be prepared and the prototype will be developed on the workstation. The prototype was based on the flow and architecture from the previous phase. The expected outputs that will be obtained from this phase are design of the prototype, prototype of proposed solution and updated literature review.

The third phase is test and evaluation. The purpose of this phase is to test and develop the developed prototype. Firstly the prototype will be tested using test from the Integrated Development Environment (IDE) of the developed prototype to clear out the bugs in the prototype. Secondly, case study will be proposed and the developed prototype will be evaluated through this case study. The evaluation was focused on the workability of the developed prototype. The focus of this evaluation is to evaluate the workability of the proposed approach in the developed prototype. The obtained result is then analyzed and the findings from the results are then concluded. The expected outputs from this phase are testing and evaluation results, analysis of evaluation results and limitation with recommendation for future work.

The last phase is report write up. All the information, data, design, results and analysis will be organized and written into chapters. The literature review also will be updated and written in the form of chapters. The expected output of this phase is research report.

Current phase of this work is at phase two which is the design and develops the prototype. The design of the prototype is explained in the next section.

V. PROPOSED WORK

Figure III shows the proposed work of this research. the goal of this research is to create a preventive approach that can assist newbie programmers in reducing code clone occurrences during software development and maintenance. Therefore, this work proposes a recommender-like process to assist newbie programmers in writing source code with reduced code clone occurrences. The enhancement is done by adding this process as a separate process into the generic pipeline model. This process will be supported by two processes which are clone categorization and suggestion process; and clone view process.

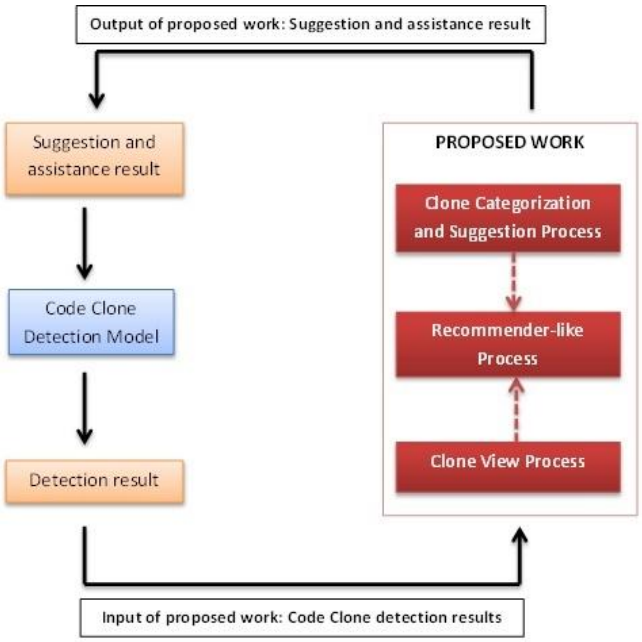

Fig. III. Proposed Work

The preventive approach begins when the software developer starts to write the source code of the software. The proposed work enhances an existing code clone detection model which is the generic pipeline model [7, 8]. In order to detect code clone, a model or detection is needed. Therefore, the generic pipeline model [7, 8] that is implemented in Java Code Clone Detector (JCCD) [8] is chosen. This model is chosen due to its extendibility and expendability of the model together with its tool.

The code clone detection model produces the code clone detection results. These results are then used as input for the proposed work which is the recommender-like process. The recommender-like process is supported by two processes which are Clone Categorization and Suggestion process; and Clone View process. The output of the proposed work is the suggestion and assistance results which will help software developers during software development. This process is recursive and will end once the software developer stops the development activity.

VI. Expected Result

This work is expected to produce a preventive approach that can reduce code clone during software development and at the same time can assist software developers during software development. The proposed work will be evaluated on two aspects. The first is the runtime performance and detected code clone of the proposed work. Secondly is the user perception upon using the proposed work. This will be known feedback of the user after using the proposed enhancement on the generic pipeline model.

It is believed thorough the proposed processes which are Recommender-like process, clone categorization and suggestion process; and clone view process will assist the software developers in reducing code clone during software development by assisting them writing clone free software.

References

[1] Y. Ueda, T. Kamiya, S. Kusumoto, and K. Inoue, "Code clone analysis environment for supporting software development and maintenance," *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, vol. 89, no. 11, pp. 10-18, 2006/11//, 2006.

[2] C. K. Roy, and J. R. Cordy, "A Survey on Software Clone Detection Research," *SCHOOL OF COMPUTING TR 2007-541, QUEEN'S UNIVERSITY,* vol. 115, September 26, 2007., 2007.

[3] S. Jarzabek, and Y. Xue, "Are clones harmful for maintenance?," in Proceedings of the 4th International Workshop on Software Clones, Cape Town, South Africa, 2010, pp. 73-74.

[4] S. Bellon, R. Koschke, G. Antoniol, J. Krinke, and E. Merlo, "Comparison and evaluation of clone detection tools," *IEEE Transactions on Software Engineering*, vol. 33, no. 9, pp. 577-591, 2007, 2007.

[5] R. Koschke, "Frontiers of software clone management," in Frontiers of Software Maintenance, 2008. FoSM 2008., Beijing, 2008, pp. 119-128.

[6] D. Rattan, R. Bhatia, and M. Singh, "Software clone detection: A systematic review," *Information and Software Technology*, vol. 55, no. 7, pp. 1165-1199, 2013.

[7] B. Biegel, and S. Diehl, "Highly configurable and extensible code clone detection," in 17th Working Conference on Reverse Engineering (WCRE), Beverly, MA, 2010, pp. 237-241.

[8] B. Biegel, and S. Diehl, "JCCD: a flexible and extensible API for implementing custom code clone detectors," in Proceedings of the IEEE/ACM International Conference on Automated Software Engineering (ASE '10) Antwerp, Belgium, 2010, pp. 167-168.

[9] B. Lague, D. Proulx, J. Mayrand, E. M. Merlo, and J. Hudepohl, "Assessing the benefits of incorporating function clone detection in a development process." pp. 314-321.

[10] H. A. Basit, and S. Jarzabek, "Detecting higher-level similarity patterns in programs," *SIGSOFT Softw. Eng. Notes*, vol. 30, no. 5, pp. 156-165, 2005.

[11] D. Hou, F. Jacob, and P. Jablonski, "Exploring the design space of proactive tool support for copy-and-paste programming." pp. 188-202.

[12] S. Kawaguchi, T. Yamashina, H. Uwano, K. Fushida, Y. Kamei, M. Nagura, and H. Iida, "SHINOBI: A Tool for Automatic Code Clone Detection in the IDE." pp. 313-314.

[13] M. Toomim, A. Begel, and S. L. Graham, "Managing Duplicated Code with Linked Editing," in Visual Languages and Human Centric Computing, 2004 IEEE Symposium on, 2004, pp. 173-180.

[14] E. Duala-Ekoko, and M. P. Robillard, "Clonetracker: tool support for code clone management," in Proceedings of the 30th international conference on Software engineering, Leipzig, Germany, 2008, pp. 843-846.

[15] E. Juergens, F. Deissenboeck, and B. Hummel, "CloneDetective - A workbench for clone detection research," in Proceedings of the 31st International Conference on Software Engineering, 2009, pp. 603-606.

Developing a Maturity Model to Measure Service-oriented Enterprise Architecture (SoEA) Adoption in an Organisation

Nur Azaliah Abu Bakar
Advanced Informatics School
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
nazaliah2@live.utm.my

Harihodin Selamat;Nazri Kama
Advanced Informatics School
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
{harihodin;nazrikama}@ic.utm.my

***Abstract*— A number of researchers has highlighted the need to explore on the Service-oriented Enterprise Architecture (SoEA) area and its benefits to the organisation. However, SoEA adoption is still at the early stage in many organisations and not well understood. To assess the level of SoEA adoption, a proper method needs to be defined. Therefore, a study will be conducted to investigate the area of SoEA adoption, followed by proposing a new SoEA maturity model as a method to measure SoEA adoption level in an organisation. This paper will describe the early phase of this study by explaining the research motivation, objectives and related works. In this paper, the research conceptual framework will be proposed and the significance of the study will be explained at end of this paper.**

Keywords— Service-oriented Enterprise Architecture (SoEA); Maturity Model;public sector organisation

I. INTRODUCTION

Service-oriented Architecture (SOA) has gained a significant momentum, industry support, and adoption. Indisputably, it has become a mainstream component of almost every Information Technology (IT) organisation. SOA has helped transform many businesses to make them more efficient and agile. The term "Service-oriented Architecture" or "SOA" was first coined by Gartner analyst [1] in the research paper published in 1996. SOA is specified as "a style of multi-tier computing that helps organisations share logic and data among multiple applications and usage modes" [2]. Since then, SOA has been described in many ways. SOA can be defined an approach focused on software development to build loosely-coupled distributed applications using a collection of services [3]. In an SOA, resources are made available to other participants in the network as independent services that are accessed in a standardised way. According to [4], SOA is a technology architectural model for service-oriented solutions with distinct characteristics in support of realising service-orientation and the strategic goals associated with service-oriented computing.

Meanwhile, Enterprise Architecture (EA) is a term used to explain the integration of strategy, business, information systems and technology towards a common goal and mastering organisational complexity through the development and usage of architectural descriptions [5]. EA has developed to bring the information system design and business requirements together. EA analyses an organisation all the way from its generic strategic components to its detailed IT infrastructure. Hence, EA is more than architecture because it encompasses governance as well as a roadmap for aligning IT investments with business needs. The concept of SOA has induced EA methodological changes. The combination of SOA and EA introducing the notion of Service-oriented Enterprise Architecture (SoEA) which has highlights their synergic relationship. This new approach allows EA and SOA to complete each other for better support of agile business needs [6].

The term Service-oriented Enterprise Architecture (SoEA) or Enterprise Service-oriented Architecture (ESOA) is both referring to the same new enterprise software architectural style which it is an abstraction of concrete SOA architectures in enterprises. SoEA combines SOA basic principles and constraints with specific EA environment and business requirements either its functional and non-functional. It focuses on service orientation and loose-coupled integration, and promotes interoperability, agility, performance, reliability, reusability and extensibility for enterprise business interests. Normally enterprise systems consist of complicate applications in heterogeneous environment. So having SoEA can help application integration because of its interoperability and relatively loose coupling service nature [7].

According to [8], SoEA easily integrates wide-spread technological approaches such as SOA because it uses service as structuring and governing paradigm. SoEA which is derived by SOA has always gained research interest. This is due to the great potential that has been demonstrated by the SOA itself. SOA has been heavily promoted by the analysts and IT vendors as architecture capable in addressing the business needs of the modern organisation in a cost-effective and timely manner. Perceived SOA benefits include improved flexibility and alignment between business processes and the supporting enterprise applications, lower integrations costs especially for legacy applications and numerous other advantages [9]. It is clear that SOA is having a substantial impact on the way in which software systems are developed. Based on Gartner Group report, 50 per cent of new mission-critical operational applications and business processes were designed in 2007

around SOA, and that number will be more than 80 per cent by 2010 [10]. Despite recent news that SOA adoption rates are falling and that "SOA is dead," Forrester Group recently reported that SOA adoption is increasing across all of its vertical-industry groups [11]. The reality is that SOA is currently the best option available for systems integration and leverage of legacy systems [12]. Therefore it is not surprising that SoEA also got the attention of researchers and leading IT technology. Existing SOA coupled with enterprise-wide capability would optimise the level of SOA adoption in large organisation.

However SoEA adoption is still at the early stage in many organisations and not well understood [13], [14]. Moreover, SoEA might be difficult to achieve because it assumes a willingness by units within the enterprise to share with other units whereas those services that were developed for their own needs [15].

II. Research Motivation

Although EA, SOA and SoEA are gaining wider acceptance, their current adoption rate and the characteristics of the adoptions in most organisations were not known. Very limited empirical research regarding the identification of adoption determinants, governance mechanisms and the evaluation of SOA's actual business value has been conducted [16]. In many cases some organisations do not know on which SOA based projects would bring value and show real benefits to the organisation [17]. Several studies that have been conducted reveals that the challenges in SOA adoption are on the security issues, incomplete or immature standards or solution of SOA, inability to integrate legacy applications and quality of service (QoS) issues [18]. Limited developer support and lack of skills or expertise in this area also become the main concern [19]. With the beliefs that SOA shall provide a viable business models, it make more difficult to identify the best services among many SOA service provider. Organisation also has difficulty determining where and how to start the SOA initiatives [20].

There have been many efforts to describe and rationalise the adoption of SOA. One of the approaches is to use maturity model. A maturity model can be viewed as a set of structured levels that describe how well the behaviours, practices and processes of an organisation can reliably and sustainably produce required outcomes [21]. A maturity model can be used as a benchmark for comparison and as an aid to understand the commonality of different organisations. Maturity model was first developed by Software Engineering Institute (SEI) to distinguish software engineering levels and encourage migration to better practices. It subsequently generalised the model for use in other domains including SOA. There are well known model such as SOA Maturity Model by Oracle and The Open Group, while there are also self-developed maturity model which are tailored to the specific used of the study itself. Based on the comparison of all selected maturity models, their maturity levels, characteristics and how are they defined, it can be concluded that there are some overlapping and joint characteristics.

In order for the organisation to move on with SOA they need to perform the 'as-is' situation analysis, which in this case the SOA maturity measurement tools play a big role. The maturity assessment is a way of examining the broad range of capabilities and assets that are required to support the successful adoption of SOA, and assessing their maturity [22]. Currently various maturity measurement tools existed but there is no standardisation and integration between the tools thus lead to the confusion to the organisation in choosing the suitable SOA measurement tools [23]. In the addition, most of the research on the enterprise SOA adoption and maturity model evolves among the industries and not the public service organisations. Therefore, this research will investigate the area of SoEA adoption and maturity model for public sector organisation.

III. Problem Statement

Organisations across many industries have moved or in the process of moving towards a service-oriented IT architecture. Many vendors have acclaimed SOA as a better way to design and build enterprise-wide software. While these claims may be true in the private sector, there is scant evidence in the literature for the same success in the public sector [15], [24]. The adoption of SOA as the dominant development paradigm within an organisation often requires significant organisational changes, which are difficult to achieve in the public sector, with its strict hierarchies and departmental independence. Indeed, many government departments are not ready to adopt this new paradigm [15].

From the research gap, it is important to have a clear understanding on the assimilation of SOA and EA towards the establishment of SoEA and its impact to the public sector organisation. Most public sector organisations have already embarked on SoEA technology but do not know to what extent its success and effectiveness to the organisation itself. Are their investments in SoEA worthwhile as stated by other organisations in different sectors? Is SoEA is suitable to be applied across the public sector? Many questions linger among the SoEA implementer in the public sector which until now difficult to answer because most technology providers or IT consultants will say that every SoEA journey is different and unique. However, it is believed that if there is a method to assess and identify the level of SOA adoption, at least the organisation can evaluate the how successful is the implementation and able to ensure they are on the right track from the start. Thus a new specific SOA maturity model need to be developed fulfill this need, so it will benefits the future planning of other SOA projects and supports various public sector IT initiatives.

IV. Research Objectives

This research aims to develop a Service-oriented Enterprise Architecture Maturity Model for public sector organisation. The research is designed to accomplish the following specific objectives:

1. To identify a set of characteristics and critical success factors of SoEA adoption in organisation.
2. To identify current measurement models in assessing SoEA adoption in organisation.

3. To identify a set of criteria in determining the maturity level of SoEA in public sector organisation.
4. To develop new SoEA maturity model for public sector organisation.

V. Related Works

Study by [3], stated that there are four pillars for a successful SOA adoption. The pillars are maturity, technology, governance and change management. When an organisation has decided to adopt SOA, it is necessary to evaluate the maturity of the organisation so SOA adoption path can be established. A case study conducted at a Dutch insurance company found that focus on reusability, focus on reducing complexity and engage in SOA governance are CSFs in adopting a SOA [25]. Meanwhile from the study on Korean firms adopting SOA [8] there are 20 CSFs identified for successful adoption of SOA. This study has investigated from both vendors and users viewpoints of SOA. Some of the CSFs are; deepening of enterprise-wide perception of SOA, strengthening perceptions of SOA by sharing success story, building strong support for enterprise-wide core human resources and have a clear goal-setting based on business value. Study of SOA adoption in German service industry has classified the CSFs according to Technology-Organisation-Environment (TOE) framework [16]. In technology, the factors are compatibility, relative advantage, complexity and cost; for organisation the related factors are organisation size, IT expertise, and top management support; lastly for environment category, the factors are competition and extent of adoption among competitors. The survey by [26] with IT professionals from 4,500 medium-sized and large companies from almost all industry sectors, operating in Switzerland, Germany, and Austria reported that the overall SOA adoption success is positively influenced by culture and communication, momentum, resources, and strategic importance. Other factors that also effect the adoption are architecture and service design as well as the service orientation infrastructure.

Meanwhile according to study conducted in Malaysian public sector organisation, the critical success factors in SoEA implementation can be grouped to six main factors (Awareness, Strategy, IT Governance, Infrastructure and Technology, Governance Process and Competency [15]. The study reported that SoEA awareness is critical for an agency and its top management, thus the establishment of high-level governance including governance structure, governance process or mechanism to steer the project and the behaviour of project members at the level of SoEA project governance is critical. From the study done by [27], through literature and empirical study, five CSFs were identified. The reported CSFs are Governance, Organizational Change, Business Process Modelling, Implementation Methodology and Re-use. According to [28], in the study of five different companies in Europe (consists of bank, furniture, engineering and airline companies) the identified CSFs are SOA governance, business process and budgeting, information architecture, dependency on commercial products, legacy architecture, close monitoring, testing and technical skills. Whereas the survey by [29] in South African organisations shows that the CSFs for SOA adoption and SOA project success are use of multiple standards and platforms, compatibility, top management support, good governance and strategy, adequate human and financial resources and vendor support for integration and development tools.

According to [30], in order to support an SOA implementation, maturity models should be used to investigate the readiness and maturity of service-oriented architectures. Therefore he has conducted the research in German banking industry to assess the maturity level and found that the maturity levels of SOAs in the observed German banks are very similar to each other: a basic technical platform is established, but a holistic and mature enterprise SOA including a well-established SOA Life Cycle Management could not be found. The only limitation of this model is, it only focuses on the maturity of the services and not on the entire SOA. This study [23] proposed that a SOA maturity model which is able to provide a parental framework for the SOA roadmaps. The model is known as Combined SOA Maturity Model (CSOAMM), the combination of SOA Maturity Model (SOAMM) and Service Integration Maturity Model (SIMM). By using the maturity model, companies can position themselves and their SOA adoption better and compare themselves to their collaborators.

From the studies in SOA adoption and maturity assessment, it can be concluded that there is gaining research interest in this area. The practitioner and the academician are aware on the great benefits that SOA especially on the EA platform can brings to the organisation. Instead of focusing only on the technical aspect of the SoEA and constantly introducing new SoEA technology, it is time now to re-assess and re-check the success of the SoEA implementation. Most of the existing studies focus on the SoEA/SOA adoption and implementation separately and very few relate it to the SoEA/SOA maturity. Therefore this study will combine the SoEA adoption, implementation and explain its relationship to the SOA maturity.

VI. Proposed Research Conceptual Framework

The study is focus to SoEA adoption and maturity model considerations. Adoption is understood in this context as the process of adapting and implementing SoEA principles and introducing the best practice recommendations prescribed by SoEA research community. Whereas maturity model in this context can be viewed as a measurement model of a set of structured levels that describe how well the behaviours, practices and processes of an organisation can reliably and sustainably produce required outcomes.

This topic was chosen when limited works were found on the SOA adoption and maturity assessment especially when it involves EA platform. In order to produce a new maturity model tailored for public sector organisation, an analysis will be conducted to clarify the relationship between the factors of SoEA adoption and how it can affect SoEA maturity level within the organisation. This maturity model shall be based on General Model Theory (GMT) by Stachowiak (1973) and guided by Capability Maturity Model Integration (CMMI®) by Software Engineering Institute (SEI). To strengthen the model, Critical Success Factors (CSFs) method, Institutional

Theory and Technology-Organization-Environment (TOE) Framework will be used in this study. Figure 1 shows the conceptual framework for this study as explained.

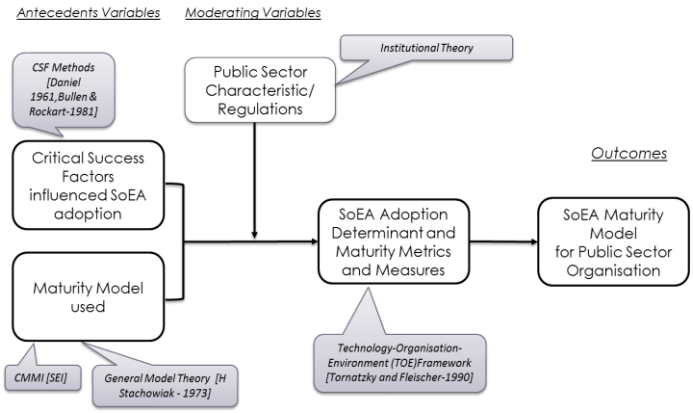

Fig. 1. Conceptual Framework for this study

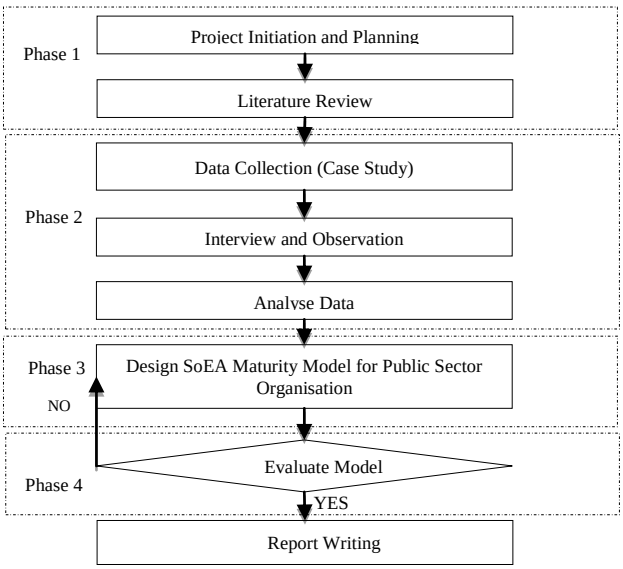

Fig. 2. Operational Framework for this study

VII. Proposed Research Methodology

This study will use a mixed method approach and one Malaysian public sector organisations will be selected as a case study. The selection is based on criteria such as the readiness and status of the organisation to implement EA and SoEA, the number of the IT services provided the size of organisation, and also the criticality of the organisation itself. In order to obtain the required information, there will be an interview session with the key person who is responsible for the SoEA projects in the organisation. Hence, based on the result, a new SoEA Maturity Model will be developed. The model then will be evaluated by the experts and tested in-depth using the previous selected test case study of that organisation. Figure 2 shows the operational framework of this study which consists of all the phases in this study. Therefore, the results of this study are expected to be tailored accordingly to the need of public sector organisation and cannot be generalised to private sector organisation without further research to justify whether this maturity model is complete of useful for such organisation. Since this study is focusing on SoEA adoption and maturity model only, any detailed, technical and low-level (operational) service-oriented characteristics is excluded from this study.

VIII. Current Stage of the Research

This research is currently in the proposal stage. The researcher is in the process of gathering the 'as-is' situation on SoEA in public sector organisation through extensive literature review analysis and preliminary interviews from the IT Managers in the Malaysian public sector organisation

IX. Significance of Research

This study involves the development of a maturity model to measure SoEA adoption in an organisation, with the focus of public sector organisation. Thus, this study is significant firstly from theoretical aspect because various possible relevant theories and methods will be used in this study, in which this approach is rarely done by any previous research in developing a maturity model for SOA or SoEA. Secondly, this study contributes in practical aspect by development of SoEA maturity model for the public sector organisation. This model shall help any organisation, especially those in public sector to measure their maturity level of SoEA adoption accordingly. It is hoped that by knowing their current state SoEA implementation, the organisation will be able to strategically plan for the future direction of their IT project or investment.

X. Conclusion

In this paper, primary issues related to research have been discussed. The goal of this research was defined as developing a new maturity model for SoEA in public sector organisations. The questions of this research being raised by background of research, literature review and researcher's experiences. In addition, the research objectives and scopes are defined with regard to the research motivation and research questions. This research attempts to contribute to practice by developing a new SoEA maturity model for public sector organisation in order to assess the level of SoEA adoption. This model will help the organisation to strategically plan the future direction of their IT project or investment. In academia, this research shall unveil how the Critical Success Factors (CSFs) methods combined with Technology-Organization-Environment (TOE) Framework and Institutional Theory become the essence in developing the maturity model based on General Model Theory (GMT) and Capability Maturity Model® Integration (CMMI®)

ACKNOWLEDGEMENT

The research is financially supported by Ministry of Higher Education Malaysia and Universiti Teknologi Malaysia under Vot No: 07J87.

REFERENCES

[1] R. W. Schulte and Y. V Natis, "'Service oriented'architectures, part 1," Gartner, SSA Res. Note SPA-401-068, 1996.

[2] J. L. Zhao, H. J. Wang, M. Goul, S. Purao, and P. Vitharana, "Impact of service-centric computing on business and education," Commun. Assoc. Inf. Syst., vol. 22, no. 16, pp. 295–310, 2008.

[3] L. Bastida, A. Berreteaga, and I. Cañadas, "Adopting service oriented architectures made simple," in in Enterprise Interoperability III, Springer, 2008, pp. 221–232.

[4] T. Erl, M. L. Taub, K. Hart, J. Mcfarland, and T. Young, SOA Design Patterns, 1st ed. Boston, USA: PRENTICE HALL, 2009.

[5] E. Niemi and S. Pekkola, "Enterprise Architecture Quality Attributes: A Case Study," 46th Hawaii Int. Conf. Syst. Sci., pp. 3878–3887, Jan. 2013.

[6] M. K. Haki and M. W. Forte, "Service Oriented Enterprise Architecture Framework," 2010 6th World Congr. Serv., pp. 391–398, Jul. 2010.

[7] L. Tang, J. Dong, Y. Zhao, and W.-T. Tsai, "A Classification of Enterprise Service-Oriented Architecture," 2010 Fifth IEEE Int. Symp. Serv. Oriented Syst. Eng., pp. 74–81, Jun. 2010.

[8] S S. Nurcan, "Service oriented Enterprise Architecture for Enterprise Engineering," in The Fifth Workshop on SoEA4EE, 2012, pp. 1–4.

[9] G. Feuerlicht and S. Govardhan, "SOA: Trends and directions," Syst. Integr., vol. 149, 2009.

[10] C. Abrams and R. W. Schulte, "Service-oriented architecture overview and guide to SOA research," G00154463, Stamford Gart. Res., no. January, 2008.

[11] R. Heffner, "SOA Adoption: Technology Diversity Accelerators," Forrester Res., 2008.

[12] G. A. Lewis, "Is SOA Being Pushed Beyond Its Limits ?," vol. 2, no. 1, pp. 17–23, 2013.

[13] S. Gerić and N. Vrček, "Prerequisites for successful implementation of Service-Oriented Architecture," in Proceedings of the ITI 2009 (ITI'09) 31st International Conference on Information Technology Interfaces, 2009, pp. 175–180.

[14] J. H. Lee, H.-J. Shim, and K. K. Kim, "Critical Success Factors in SOA Implementation: An Exploratory Study," Inf. Syst. Manag., vol. 27, no. 2, pp. 123–145, Apr. 2010.

[15] M. A. Manan and P. Hyland, "Enterprise SOA Implementation Readiness : a Case Study in Malaysia," in 22nd Australasian Conference on Information Systems, 2011.

[16] N. Joachim, D. Beimborn, and T. Weitzel, "Investigating Adoption Determinants of Service-Oriented Architectures (SOA)," in Proceedings of the Special Interest Group on Services (SIG SVC) (Pre-ICIS Workshop), 2010, vol. 10, no. 2010, pp. 10–13.

[17] L. O'Brien, J. Gibson, and J. Gray, "A method for selecting SOA pilot projects including a pilot metrics framework," in 33rd International Conference on Software Engineering (ICSE '11), 2011, p. 653.

[18] A. Becker, T. Widjaja, and P. Buxmann, "Value Potentials and Challenges of Service-Oriented Architectures," Bus. Inf. Syst. Eng., vol. 3, no. 4, pp. 199–210, Jul. 2011.

[19] R. K. Das and M. R. Patra, "SOA for e-Governance in India : Potentials and Pitfalls," in 3rd international conference on Theory and practice of electronic governance, 2009, pp. 36–42.

[20] S. Geriü, "The Potential of Service-Oriented Architectures," in 32nd International Conference on Information Technology Interfaces (ITI 2010), 2010, pp. 471–476.

[21] M. C. Paulk, "A History of the Capability Maturity Model for Software," vol. 1, no. 1, 2008.

[22] A. Pugsley, "Assessing your SOA Program," 2006.

[23] E. Söderström and F. Meier, "Combined SOA Maturity Model (CSOAMM): Towards a Guide for SOA Adoption," in in Enterprise Interoperability II SE - 43, R. Gonçalves, J. Müller, K. Mertins, and M. Zelm, Eds. Springer London, 2007, pp. 389–400.

[24] E. MacLennan and J.-P. Belle, "Factors affecting the organizational adoption of service-oriented architecture (SOA)," Inf. Syst. E-bus. Manag., Jan. 2013.

[25] W. Vegter, "Critical success factors for a SOA implementation A case study in the financial sector," in 11thTwente Student Conference on IT, Enschede, 2009.

[26] S. Aier, T. Bucher, and R. Winter, "Critical Success Factors of Service Orientation in Information Systems Engineering," Bus. Inf. Syst. Eng., vol. 3, no. 2, pp. 77–88, Feb. 2011.

[27] I. Owens and J. Cunningham, "The Identification of Service Oriented Architecture-Specific Critical Success Factors," in 6th European Conference on Information Management and Evaluation, 2012, p. 267.

[28] M. Galinium and N. Shahbaz, "Success factors model: Case studies in the migration of legacy systems to Service Oriented Architecture," 2012 Ninth Int. Conf. Comput. Sci. Softw. Eng., pp. 236–241, May 2012.

[29] E. MacLennan and J.-P. Belle, "Factors affecting the organizational adoption of service-oriented architecture (SOA)," Inf. Syst. E-bus. Manag., Jan. 2013.

[30] J. Eckert, M. Bachhuber, A. A. Miede, A. Papageorgiou, and R. Steinmetz, "Readiness and maturity of service-oriented architectures in the German banking industry," in 4th IEEE International Conference on Digital Ecosystems and Technologies, 2010, pp. 270–274.

Ontology for Evaluation of Semantic Web Services Ontology

Nurhamizah Mohd Hamka, Radziah Mohamad
Department of Software Engineering
Faculty of Computing
Universiti Teknologi Malaysia
81310, Skudai, Johor, Malaysia
mieza.hamka@gmail.com, radziahm@utm.my

***Abstract*— Ontology has being used in large area, from clinical use case towards Web services. Web services techniques has involved services provider and services consumer to interact via Web environment to exchange services information using XML and WSDL technologies. Ontology were used to support the semantic part of the services description that could not being supported by the WSDL format, which it only describe the syntactic part of the services. The selection of suitable ontology to describe services description were not in focus. The issues within the ontology component aspect might affect services discovery. Since semantic Web services uses ontology to support semantic description for ease of service allocation, the ontology would then first be evaluated. Current ontological evaluation were not yet focused on semantic Web services ontology evaluation and need formalization. The study propose on developing evaluation ontology that focus on this types of ontology, so that it enable the ontology to serve services discovery.**

***Index Terms*— Ontology, Ontology Evaluation, Structural, Semantic, Evaluation Aspect, Quality Criteria**

I. INTRODUCTION

Ontology is a conceptual model to describe knowledge within certain domain. The term ontology was abstracted from philosophical world and used in computer science by Tom Gruber in 1993 [1]. It act as mediators of agreement towards conceptual representation of body of knowledge [2]. The ontology component were introduced with different terms situated as the following. Conceptual of ontologies were visualized in figure 1 consist of the component mention in table 1 and the element were describe as in [3].

TABLE I. COMPONENT IN ONTOLOGY

OWL	Protégé	Description
Class	Class	Class or concept to describe domain
Properties	Slot	Relationship to connect between Class and Instances
Instances	Individual	Domain of discourse

In the semantic Web, the ontology has being used to describe the semantic description of the Web services. Semantic Web services technology discover Web services by its description handled in the level of ontology conceptualization. This technology grouped information data and services in various resources without losing its meaning [4].There are ontology to describe Web services, for example OWL-S [5], WSMO [6] and FLOWS [7]. These ontology referred to WSDL. There are two main types of ontology group related to semantic Web services ontology, known as the services ontology and Web services ontology.

However, the issue with WSDL is that, it does not support semantic aspect of the Web services description. Hence, it effect services discovery. Besides, ontology language dependency also becomes one of the services discovery effect factor [8]. The evaluation focused towards the semantic Web services, and also ontology had existed, but there are not take semantic Web services ontology as their types of domain of evaluation. Just several research focused on semantic Web services types of ontology, introduced in [9] where she also list down several requirements to be achieved by this types of ontology. The requirements were grouped by two different ontology, which is the generic Web services ontology and Web services domain ontology.

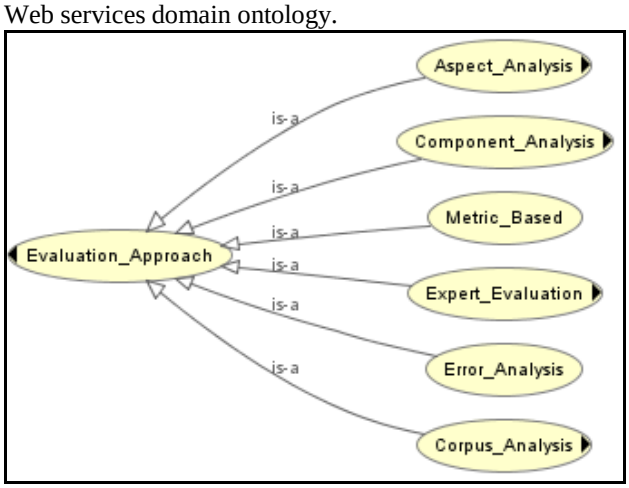

Fig. 1. Evaluation Approach Concept in Proposed Ontology

The discovery of Web services make use ontology technology to support the description of the Web services in

the Web. Hence, the description state in the ontology need to be understandable for ease of Web services allocation. Since there are no standardize requirements of ontology to describe Web services, there are aspect within the ontology that need refinement and evaluated so that the ontology used to referred semantic description support the discovery of Web services. The study of ontological aspect of evaluation had being grouped in [10]. Ontological element were considered for evaluation because the approach for describing the semantic of services and annotations depends on the ontology description language [8].

Our contribution is to enhanced an evaluation ontology to evaluate the semantic Web services ontology evaluation from current study made by [11]. The methodology used for the ontology development was based on competency questions by [1] and by ontology development 101 [3]. The study is structured by the following section situated the problem statement, section three as objectives, section four as related studies, section five as research methodology, section six as the expected result and conclusion of our research study.

II. Problem Background

Ontology enable reasoning in Semantic Web for services discovery. However, there is still no standardize requirements that had being gather for semantic Web services ontology [9], [12]. Besides, the construction of generic Web services ontology are also time consuming. Although we were not focusing on building an ontology for Web services ontology, the factors of timeframe for the ontology might affect the process of evaluating this types of ontology. The description of Web services does not stick to only one ontology description language. Survey made by [8] reviewed that the semantic description of Web services has being supported by different ontological language. It does mean that, different types of ontological language might differ in terms of its evaluation approach or techniques.

The services discovery issues depends towards the usage of the ontology. This types of ontology has little attention towards evaluation when it comes to fulfill its own quality [9], with the addition of the requirements are still not formalized. Since it being used to support services discovery, the ontology need to be evaluate because there might be some problem in the structural or semantic aspect within the ontology. The issues happen of semantic disappearance. during transition of OWL-S into WSDL format, the semantic information does not being carried out in the WSDL. There are also large amount of ontology attached during services discovery [13] thus ontology evaluation need to be done for suitable ontology identifications.

III. Objectives

Based from the problem being state above, the objectives of this research are situated as the following:

1. To analyse the requirements of evaluating semantic web services ontologies
2. To propose an ontology to evaluate ontologies for semantic web services
3. To enhance the ranking of semantic web services ontologies

IV. Related Studies

There are related study consist of search for ontology by keywords and evaluate the found ontology that used within . Semantic Web services provides task of discovery of services in Web environment. There are many ontology evaluation approached to date. The following table are some of the requirements findings related to semantic Web services ontology gather from literature survey.

TABLE II. REQUIREMENTS OF SEMANTIC WEB SERVICES ONTOLOGY

No	Source	Layers of Ontologies	Requirements
1	Generic Web Services Ontology [9]	Top-level	High modelling expressiveness
			Clear Semantic and rich formalization
			Adaptability
			Harmonization with other standards
			Usability
2	Web Services Domain Ontology [9]	Domain	Broad domain coverage
			Scalability
3	An Ontology for Semantic Web Services [14]	Top-level	Similarity
4	Application Ontologies for Semantic Web Services [12]	Application	Similarity
			Service retrieval accuracy
5	First Order Logic Ontology for Web Services (FLOWS) [7]	Top-level	Reusability
			Share ability
6	Concept-Based Readability of Web Services Descriptions [15]	Domain	Readability
7	Search and Evaluation of Web Services Ontology[16]	Domain	Coverage

There are several previous research works to identify the contribution towards ontology evaluation. There are development methodologies [17] and also metric of evaluation [18] in literature. Hence, the needs to come out with proposal of evaluation approach that will provide the ranks of ontology searched from web, make the selection of the ontologies become more easier. The related studies that being major focus for this research is based on the aspect of evaluation, metric used [18][19][20], searching [16], [21], ranking [22]–[25] and selection of the suitable ontology.

TABLE III. APPROACHES TOWARDS ONTOLOGY EVALUATION

Approached	Sources
Evaluation tools	[2]
Development evaluation methodologies	[17] [26]
Metric of evaluation	[18], [27][19][20]
Ontology Searching	[16], [21]
Ontology Ranking	[22]–[25][28]

V. RESEARCH METHODOLOGY

The research framework for this study were proposed to follow as Fig. 1. The research flow consist of requirement phase, analysis phase, development phase, implementation, evaluation and maintenance phase as the discipline introduced in Software Engineering software development. At current, this research study were in development phase. This phase consist of development of OntoUji, the evaluation ontology used to evaluate semantic Web services ontology after being retrieved from the ontology repositories. The motivation behind this evaluation is to select the suitable ontology before the services provider used it to create semantic description of the Web services. By then, it will support future services discovery.

Research framework consist of the exact activities and component to reach objectives of the research. Based on *Fig. 3*, the number indicate the flow happen during the framework that consist of:

1. Flow 1:Ontology searching based on keywords suggestion by user
2. Flow 2: Search ontology from repositories in the Web
3. Flow 3: Retrieve list of ontologies that match with the keyword suggested by the user
4. Flow 4:The retrieved ontology will then be evaluated by the criteria of others aspect intended during the evaluation process
5. Flow 5:The process evaluation output that list ontology and do the process of ranking to the list
6. Flow 6:The result of the ontology list will be return towards the centre process
7. Flow 7:Display list of ranked ontology to the interface of the tools
8. Flow 8:Display the interface to the user
9. Flow 9:User are able to view the ranked ontology in the interface
10. Flow 10:Selection of ontology from the list
11. Flow 11:The selected ontology were viewed and used for creation process
12. Flow 12:Process of converting the selected ontology to semantic description of the services that want to be presented
13. Flow 13:The retrieval of services will be evaluated when the discovery of services being made to search the description made based from the selected ontology

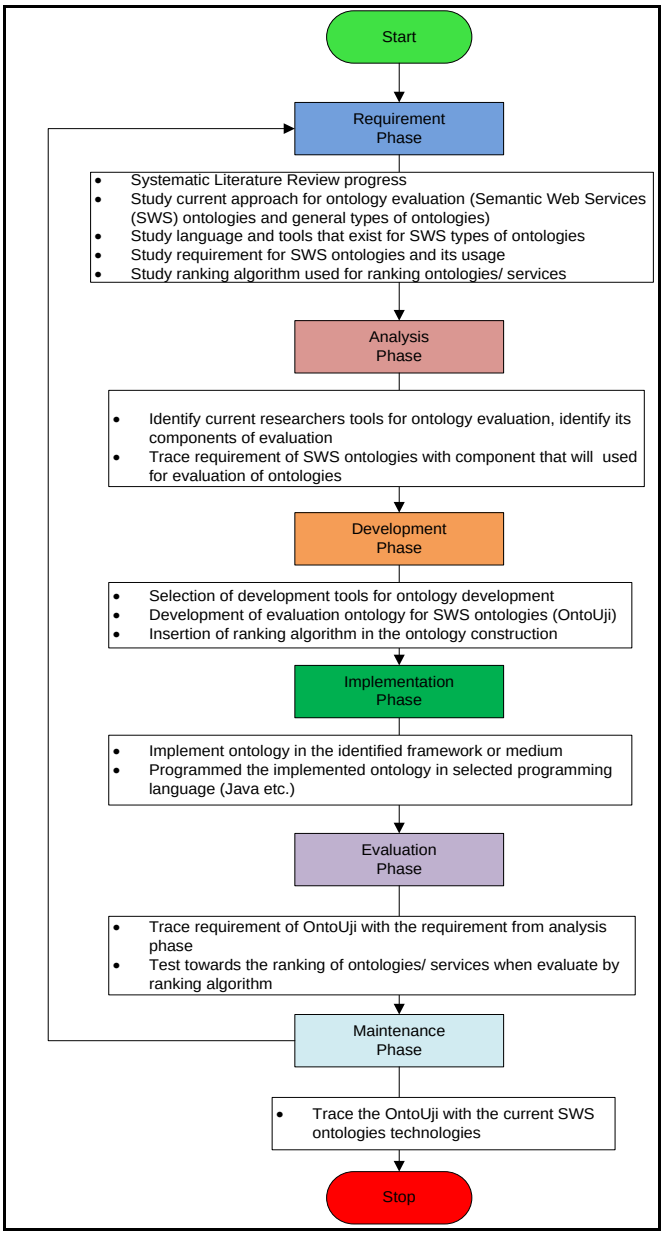

Fig. 2. Research Flow

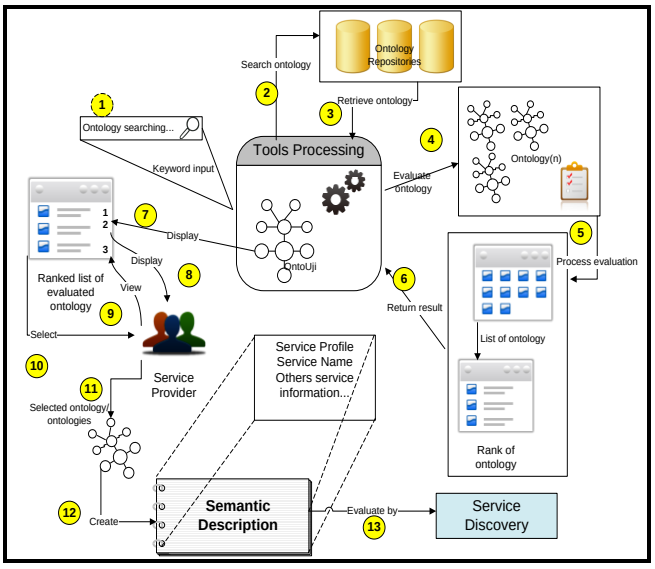

Fig. 3. Research Framework

VI. Expected Result

Expected result from the research findings is the ontological evaluation tools. Enable user to search for keywords of services, manipulate Web services description within the Web by calling it by OntoUji, the ontology of evaluation, and rank the found ontology with the result of correctness percentage. The selected ontology will be validated by user whether it helping in discovery of services. Below is the referred framework for ontology evaluation from [10]. It will then enhanced within oQual framework proposed by [29] and we will our connect with OntoUji to initiate semantic Web services ontology evaluation.

There are various proposed of evaluation output from previous experts, but there are no formal or standardize aspect for the ontology to be selected. The motives of proposing ontology of evaluation is to standardized the evaluation element that are needed to be focus when dealing with the semantic Web services types of ontology.

VII. Conclusion

Our study come out to proposed evaluation ontology that might be will look up across the limitation of the structure ontology that being the major gap on selection of the suitable ontology. Future works is to refer the oQual framework [29] of ontology evaluation. We will enhance the conceptual aspect to support structural needs of the ontology to be correct and consistent. From the study, the proposed evaluation ontology will be encrypt within ranking tools. It is to provide the best measurement of ontology and ease of use for user to select the suitable ontology from the interfaces of the ranked list of ontologies.

From the above study, standardize evaluation factors and element need to be grouped into one conceptualization to satisfy the objectives of rank suitable ontology to be selected by user. The proposed ontology will also undergo inspection of evaluation. This process are evolving, many information will be gather to unite quality aspect that need to be fulfill. The evaluation will focus on one aspect or combination of others aspect known in literature, so that the proposed work are relate with the current required element of ontology evaluation.

Acknowledgement

We would like to thank Ministry of Education Malaysia (MOE) for sponsoring the research through the FRGS grant with vote number 4F165 and Universiti Teknologi Malaysia for providing the facilities and support for the research. In addition, we would like to extend the gratitude to our lab members in EReTSEL Lab of Faculty of Computing, Universiti Teknologi Malaysia for their ideas and support throughout this study.

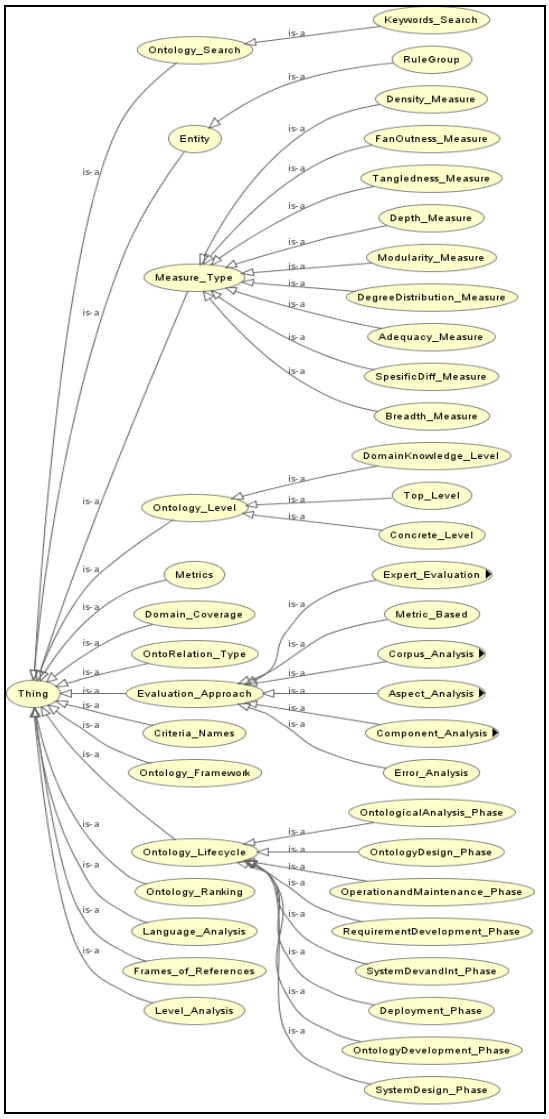

Fig. 4. Concepts in OntoUji

References

[1] T. R. Gruber, "Toward Principles for the Design of Ontologies Used for Knowledge Sharing," Int. J. Human-Computer Stud., pp. 907–928, 1993.

[2] T. Aruna, K. Saranya, and C. Bhandari, "A Survey on Ontology Evaluation Tools," in 2011 International Conference on Process Automation, Control and Computing, 2011, pp. 1–5.

[3] N. F. Noy and D. L. Mcguinness, "Ontology Development 101 : A Guide to Creating Your First Ontology," pp. 1–25, 2000.

[4] R. Mohamad and F. Zeshan, "Comparative Evaluation of Semantic Web Service," in in Software Engineering and Computer Systems, J.M. Zain., J. M. Zain, W. M. bt W. Mohd, and E. El-Qawasmeh, Eds. Springer-Verlag Berlin Heidelberg, 2011, pp. 283–290.

[5] D. Martin, M. Burstein, J. Hobbs, O. Lassila, D. McDermott, and D. Mcdermott, "OWL-S : Semantic

Markup for Web Services," no. November, pp. 1–34, 2004.

[6] M. Dimitrov and A. Simov, "WSMO Studio Users Guide v. 1.8," 2006.

[7] M. Gruninger, R. Hull, and S. A. Mcilraith, "A Short Overview of FLOWS : A First-Order Logic Ontology for Web Services ∗ ," Bull. IEEE Comput. Soc. Tech. Comm. Data Eng., vol. 31, no. 1, pp. 1–5, 2008.

[8] I. El Bitar, F.-Z. Belouadha, and O. Roudies, "Review of Web services description approaches," in 2013 8th International Conference on Intelligent Systems: Theories and Applications (SITA), 2013, pp. 1–5.

[9] M. Sabou, "Building Web Service Ontologies," Dutch Graduate School for Information and Knowledge Systems, 2006.

[10] D. Vrandečić, "Ontology Evaluation," Springer, 2010.

[11] A. Gangemi, C. Catenacci, M. Ciaramita, J. Lehmann, and R. Gil, "Ontology evaluation and validation An integrated formal model for the quality diagnostic task (Technical Paper)," Media, vol. 3. Springer, pp. 1–53, 2005.

[12] X. Wang and W. A. Halang, Discovery and Selection of Semantic Web Services, vol. 453. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 1–132.

[13] Z. Liu, H. Wang, and B. Zhou, "A Scalable Mechanism for Semantic Service Discovery in Multi-ontology Environment," in in Advances in Grid and Pervasive Computing, Springer Berlin Heidelberg, 2007, pp. 136–145.

[14] Q. Qiu and Q. Xiong, "An Ontology for Semantic Web Services," in in High Performance Computing and Communications, HPCC 2007,., R. Perrott, Ed. Springer-Verlag Berlin Heidelberg 2007, 2007, pp. 776–784.

[15] P. Sripairojthikoon and T. Senivongse, "Concept-Based Readability of Web Services Descriptions," 2013, pp. 853–858.

[16] L. Engel, M. C. Jaeger, and G. Mühl, "Search And Evaluation Of Ontologies For Semantic Web Services In The Internet," in IADIS International Conference on WWW/Internet 2005, 2005, pp. 255–260.

[17] M. Cristani and R. Cuel, "A Survey on Ontology Creation Methodologies," Int. J. Semant. Web Inf. Syst., vol. 1, no. 2, pp. 49–69, Jan. 2005.

[18] J. García, F. J. García-peñalvo, and R. Therón, "A Survey on Ontology Metrics," in in M.D. Lytras et al. (Eds.): WSKS 2010, Part I, CCIS 111, M.D. Lytra., Springer-Verlag Berlin Heidelberg 2010, 2010, pp. 22–27.

[19] S. Oh and J. Ahn, "Ontology Module Metrics," in 2009 IEEE International Conference on e-Business Engineering, 2009, pp. 11–18.

[20] S. Tartir and I. B. Arpinar, "Ontology Evaluation and Ranking using OntoQA," in International Conference on Semantic Computing (ICSC 2007), 2007, pp. 185–192.

[21] P. Buitelaar and T. Eigner, "Evaluating Ontology Search," in Proceedings of the Fifth International Workshop on Evaluation of Ontologies and Ontology-based Tools (EON 2007, 2007.

[22] Z. Ding and Z. Duan, "Improved Ontology Ranking Algorithm Based on Semantic Web," in Ubi-media Computing (U-Media), 2010 3rd IEEE International Conference on, 2010, pp. 103–107.

[23] C. Patel, K. Supekar, Y. Lee, and E. K. Park, "OntoKhoj : a semantic web portal for ontology searching, ranking and classification," in Proceedings of the fifth ACM international workshop on Web information and data management - WIDM '03, 2003, p. 58.

[24] R. Subhashini, J. Akilandeswari, and V. Sinthuja, "A Review on Ontology Ranking Algorithms," Int. J. Comput. Appl. (0975 – 8887), vol. 33, no. 4, pp. 6–11, 2011.

[25] E. Thomas, D. Sleeman, C. Brewster, and U. Sheffield, "Searching and Ranking Ontologies on the Semantic Web," 2005.

[26] S. Staab, A. Gómez-Pérez, W. Daelemana, M. L. Reinberger, and N. F. Noy, Why evaluate ontology technologies? Because it works!, vol. 19, no. 4. IEEE Educational Activities Department, 2004, pp. 74–81.

[27] D. Vrandečić, Y. Sure, and D. Vrandecic, "How to Design Better Ontology Metrics," in The Semantic Web Research and Applications, 2007, vol. 4519, pp. 311–325.

[28] Y. Wei, C. Jiaheng, and C. Junpeng, "A Novel Approach for Ranking Ontologies on the Semantic Web ISwoogls," in 2006 1 st International Symposium on Pervasive Computing and Applications, 2006, no. 60273072, pp. 608–612.

[29] A. Gangemi, C. Catenacci, M. Ciaramita, and J. Lehmann, "Modelling Ontology Evaluation and Validation," in Proceedings of ESWC2006, Springer, 2006, vol. 4011, pp. 140–154.

Software Design

Mobile Commerce Customer Relationships Management Model (MCCRMM) **Muhib Ullah and Siffat Ullah Khan**
Software Re-engineering Risks Management Model (SRE-RMM) **Nasir Rashid and Siffat Ullah Khan**
A Review on Web Design Processes: How They Incorporated in a Complex System **Siti Azreena Mubin and Azrul Hazri Jantan**
Predicting Hearing Loss Symptoms from Audiometry Data Using Machine Learning Algorithms **Nasir G. Noma and Mohd Khanapi Abd Ghani**
The Impact of EMR User Interface Design on Doctor Satisfaction **Shatilah Zakaria and Mohd Khanapi Abd Ghani**
Migrating Traditional Projects Framework to Enterprise Resource Planning System: An Empirical Study **Basem Zughoul and Burairah Hussain**

Mobile Commerce Customer Relationships Management Model

(MCCRMM)

Muhib Ullah
Software Engineering Research Group (SERG_UOM),
Department of Computer Science & IT University of Malakand, Pakistan.
muhib6839@yahoo.com

Siffat Ullah Khan
Software Engineering Research Group (SERG_UOM),
Department of Computer Sc. & IT, Deptt: of Software Engineering, University Of Malakand, Pakistan.
siffatullah@uom.edu.pk

***Abstract*— Mobile commerce (M-Commerce) is an emerging business paradigm in which the buying and selling of products and services is done through mobile devices. These devices are mobile phones, personal digital assistant (PDA) or other hand held wireless devices. M-commerce is the concept of covering any form of business transaction or information exchange using wireless and mobile technologies. The objective of our study is to develop Mobile Commerce Customer Relationships Management Model (MCCRMM) to assist vendor organizations in establishing, strengthening and maintenance of the relationships with mobile customers in an M-commerce/Mobile business situation. We will use empirical research methodology for this research project. Systematic literature review will be used for the identification of success factors and challenges. Systematic literature review (SLR) is based on a structured protocol, and is therefore, different from ordinary literature review. SLR gives more prescribe results than the ordinary review. Finding of SLR will be validated through questionnaire survey. The different success factors and challenges from the relevant published literature will be collected for recognition of the MCCRMM through SLR. The expected results of this research will be a catalog of challenges faced by M- commerce vendors in establishment, strengthening and safeguarding of relationships with consumers in M-commerce business paradigm. This study will also explore important success parameters for organization and maintenance of long term relationships with consumers in mobile commerce environment.**

***Keywords*- Mobile Commerce, M-Commerce, M-Commerce Customer relationships Management, M-commerce vendors.**

I. INTRODUCTION

Mobile commerce is probable to be the subsequently giant signal of business [1]. Mobile commerce, or M-commerce, has significantly reduced the distances between purchaser and customer. The customer is getting connected to the business world any time anywhere.

With the advancement of wireless and mobile communication technologies it is possible for customers to buy goods anywhere using mobile devices. This new strategy of performing business referred to mobile commerce [2].

The extraordinary development in the use of cellular phones and more recently, also of wirelessly enabled mobile phones, is giving rise to yet another revolution in the business world i.e. Mobile commerce or m-commerce. Due to recent development in wireless technology, the customer is able to access information and services from a packed sized device anywhere any time [3]. There has been a wonderful increase in wireless technology in the last few years. This advancement has change the way people do businesses i.e. people are using wireless technologies for business purposes. Using wireless technologies hand held wireless devices the customers are now able to access information and services any time anywhere.

Like traditional businesses in mobile business it is extremely important to have deep understating of the customers and have good relationship with customers. Customer relationships management (CRM) means that how customer should be managed by an organization. CRM is a commerce strategy that enhances organization's competitive aptitude, it makes a winning strategy for understanding the customer's necessities, promoting the rising rate of organization and uphold customer in a high competitive market environment. CRM is needed because it improves customer expectation, trustworthiness and satisfaction [4]. Customer trust is one of the important factor for the development of mobile commerce (M-commerce). However, it is hard to develop as it involves technology and commerce practices. Moreover, it is a continuous struggle from initial trust formation to continuous trust development [5]. From the literature review on CRM it has been cleared that it is durable to catch the attention of a new customer then to maintain the relationship with exiting one. Giving proper concentration and worth to customer has good impression on customer [6]. Prior to undertake the business activity, the customer decides either to go for traditional commerce; e-commerce or m-commerce, each of the business strategy has its own advantages and disadvantages.

Due to latest development in ICTs (Information and communication technologies) current e-commerce suppliers have started business through mobile devices. This provide some features like "always on", location centric,

convenience, customization, and identifiability which are not available in traditional e-commerce [7].

In future mobile commerce can have a significant effect on the society and business world. So it is necessary for the m-commerce practitioners and developers to understand the customers' perception of m-commerce application in order to effective design and deliver m-commerce services [8].

The use of mobile phone has been dramatically increased in developed countries as well as in developing countries. Due to which the purchase of products and services also increased using mobile devices. The rapid development of mobile applications has given rise to a new word: m-commerce. M-commerce is defined as the application of wireless communications [9]. Due to wireless and mobile communication the customers are now able to access information and services any time anywhere. i.e. (access banking financial statement and pay bills, receive store quotes and initiate buy/sell dealings, or receive special promotions) [10].

Although the scholarly study on Mobile Commerce Customer Relationships Management (MCCRM) is slowly developing, there is still a need of empirical research on this topic. Keeping in view the future of mobile commerce a variety of issues need to be explored. This study is a part of undertaking this exploration. A number of studies have shed light on the need and importance of mobile commerce [11-13].

There are two research questions [14] mentioned in section III of this paper that have provoked the effort reported in this study.

II. OBJECTIVE OF THE STUDY

The aim of this study is to develop Mobile Commerce Customer Relationships Management Model (MCCRMM) to assist vendor organizations in establishing, strengthening and maintenance of the relationships with customers in a mobile commerce business paradigm.

III. RESEARCH MOTIVATION QUESTIONS

Two research questions have motivated this research work.

RQ1.What are the challenges, as identified in literature, faced by mobile commerce vendors in building and strengthening of relationships with customers?

RQ2.What are the success factors (SFs), as identified in literature, to be adopted by mobile commerce vendors for establishment and maintenance of long term relationships with customers?

IV. RESEARCH METHODOLOGY

The methodology we used consists of three phases. In first phase the data will be collected through systematic literature review (SLR). Systematic literature review is an innovative research methodology for the data collection, identification, evaluation and compilation of data relevant to some research question or topic or phenomenon of concern [15]. In the second phase an empirical study will be conducted for validation of the findings from SLR. In the last phase case study will be conducted for the evaluation of the MCCRMM. Case study approach is an observational study used to investigate a single entity with in a particular time [16]. Details are given in the following subsections.

A. Data Acquisition and Investigation

For RQ1 and RQ2, the identification of Challenges and Critical Success Factors (CSFs), we will use systematic literature review (SLR). The first phase of SLR is the development of SLR protocol which has been developed, validated and has been published [14]. "Systematic literature review (SLR) has become an important research methodology in software engineering, since the introduction of evidence-based software engineering (EBSE) in 2004" [17]. We have used a similar methodology in our previous research [18, 19].

B. MCCRMM Development

In order to develop Mobile Commerce Customer Relationships Management Model (MCCRMM) there are six stages involve, as shown in the Figure 1.

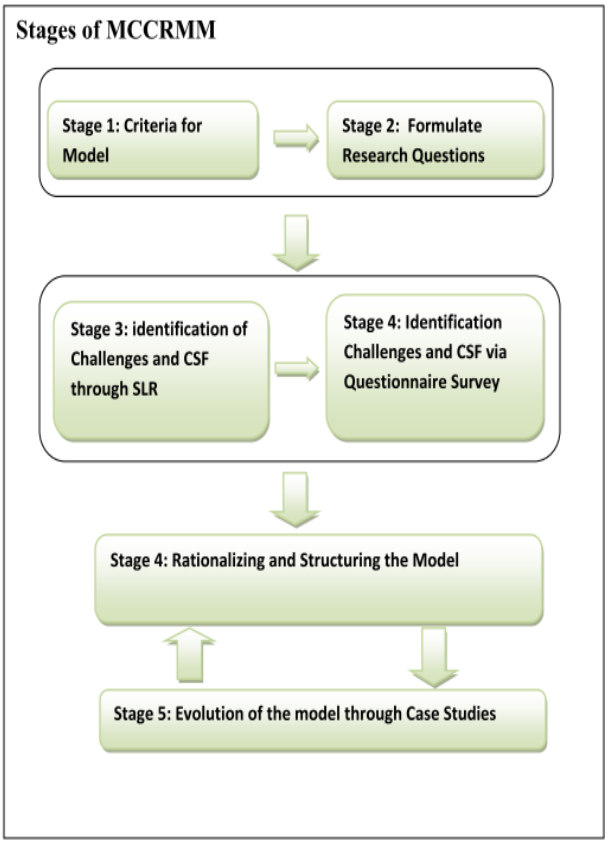

Fig. 1. Activities involved in building the MCCRMM

In the first stage we set the criteria for the development of MCCRMM. The criteria used for Mobile Commerce Customer Relationship Management Model is given bellow.

- Ease of use: The structure of MCCRMM must be easy to use, understandable and flexible.
- User satisfaction: The users should be satisfied with the result of MCCRMM.

The research questions are developed in stage 2. Data will be collected and analyzed in stage 3 and stage 4 through SLR and questionnaire survey. Rationalization and structuring of MCCRMM will be performed in stage 5. The evaluation of the model will be performed in the final stage through case studies.

Figure 2 show the proposed structure of MCCRM Model. A similar approach has been used in our previous study [18]. Similarly other researchers have also used the same approach [20]. It represents the relationship between

MCCRMM Levels, challenges and success factors in mobile commerce customer relationship management and different practices used to implement these challenges and success factors.

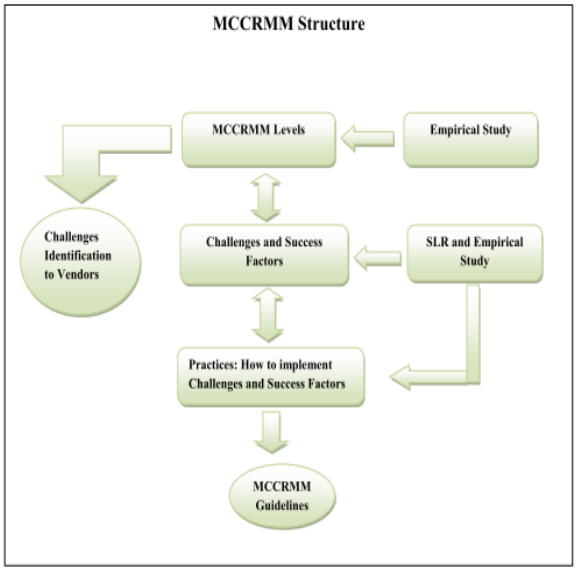

Fig. 2. MCCRMM Structure

C. *MCCRM Model Assessment*

Case study is most widely used methodology[21], therefore case study approach will be used in order to assess and review the proposed model i.e. (MCCRMM).

Four case studies will be conducted at mobile commerce vendor organizations to evaluate the Mobile Commerce Customer Relationships Management Model (MCCRMM). In order to get feedback about MCCRMM model from the participants, focus group sessions will be conducted at the end of these case studies. The focus group sessions is chosen for feedback the reason is that the data obtain from focus groups is mainly from the interaction between the members in the group without any suggestion or help from the researchers. Another reason for conducing focus groups sessions is that a group of peers allows them to be more open about issues discussed than they will be in personal interviews.

V. SUITABILITY AND NOVELTY

Different studies show that trust is critical for all business relationships [22]. As for as Mobile Commerce Customer Relationship Management (MCCRM) is concern, we know that many mobile commerce vendors are providing the same type of stuff and services. As the mobile commerce vendors are working in cyber world so it means that a competitive environment exists among the mobile commerce vendors. So for mobile commerce vendors who are providing services to mobile customers, it is extremely important to have sound and long lasting relationships with customers. In the literature many researchers have tried to deal with some of the issues of m-commerce. Though, no one has used the SLR methodology to find the challenges/barriers and critical success factors (CSFs) in Mobile Customer relationship management (MCCRM).

Mobile Commerce Customer Relationship Management Model (MCCRMM) is expected to bridge the identified gap.

VI. RESEARCH CARRIED UP TO-DATE

The following progress has been made till date.

- Background knowledge.
- Research questions identifications.
- Research methodology selection.
- Plan for building the MCCRMM.
- Method for evaluation.

SLR protocol developed and published [14]. SLR protocol implementation is in progress.

ACKNOWLEDGMENT

We are thankful to software engineering research group (SERG_UOM) in general and to Abdul Wahid Khan, Muhammad Ilyas Azeem and Muhammad Salam in particular for providing their valuable comments regarding this study.

REFERENCES

[1] M. Khalifa and K. N. C. Sammi, "Adoption of Mobile Commerce: Role of Exposure," Proceedings of the 35th Hawaii International Conference on System Sciences - 2002, 2002.

[2] T. Dahlberg, N. Mallat, J. Ondrus, and Zmijewska., "Past, present and future of mobile payments research: A literature review.," Electronic Commerce Research and Applications., vol. 7, 2008.

[3] C. Coursaris and K. Hassanein, "Understanding m-Commerce: A Consumer-Centric Model," Quarterly Journal of Electronic Commerce, vol. 3, pp. 247-271.

[4] U. M. Z. Usman and Z. B. Mohammed, "The Impact of Mobile Advertisement and Communication on Customer Relationship Management (CRM)," presented at 2012 International Conference on Economics, Business and Marketing Management, 2012.

[5] K. Siau and Z. Shen, "BUILDING CUSTOMER TRUST IN MOBILE COMMERCE," presented at Communications of the ACM., 2003.

[6] "webiste:http://www.corethree.net/mobile-commerce-crm.html."

[7] Wen J. and M. P., "M-commerce operation modes and applications," International Journal of Electronic Business., vol. 2, pp. 301-315, 2004.

[8] Zhi-shuang CHEN, Rui LI, Xi CHEN, and H. XU., "A Survey Study on Consumer Perception of Mobile-Commerce Applications," presented at 2011 2nd International Conference on Challenges in Environmental Science and Computer Engineering (CESCE 2011), 2011.

[9] R. Safeena, N. Hundewale, and A. Kamani, "Customer's Adoption of Mobile-Commerce A Study on Emerging Economy," International Journal of e-Education, e-Business, e-Management and e-Learning (IJEEEE), vol. 1, 2011.

[10] A. A. Abdelkarim and H. H. O. Nasereddin, "MOBILE COMMERCE," Journal of Mathematics and Technology, vol. 4, 2010.

[11] L. Jingfeng, ZHU Yuefeit, Z. Yajuan, and P. Heng, "Signcryption.Based Key Management for MANETs Applications in Mobile Commerce," WUJNS Wuhan University Journal of Natural Sciences, vol. 11, pp. 1621 - 1624, 2006.

[12] J. T. Cao, Y. X. Tong, Z. K. Che, and X. Li, "Analysis of M-Commerce in China," International Federation for Information Processing, Research and Practical Issues of Enterprise Information Systems, eds., vol. 205, pp. 691-693, 2006.

[13] S. S. Chan, X. Fang, J. Brzezinski, Y. Zhou, S. Xu, and J. Lam, "USABILITY FOR MOBILE COMMERCE ACROSS MULTIPLE FORM FACTORS," Journal of Electronic Commerce Research,, vol. 3, pp. 187-199, 2002.

[14] Muhib Ullah and S. U. Khan, "Challenges and Success Factors in Mobile Commerce Customer Relationships Management: A Systematic Literature Review Protocol," IOSR Journal of Computer Engineering (IOSRJCE), vol. Volume 2, pp. PP 25-34, 2012.

[15] B. Kitchenham and S. Charters., "Guidelines for performing Systematic Literature Reviews in Software Engineering.," Keele University, UK 9 july. 2007.

[16] W. Claes, H. Martin, and H. Kennet, "Empirical Research Methods in Web and Software Engineering," Empirical Methods and Studies in Software Engineering: Experiences from ESERNET, pp. 7-23, 2003.

[17] H. Zhang, M. A. Babar, and P. Tell, "Identifying relevant studies in software engineering," Information and Software Technology, vol. 53, pp. 625-637, 2011.

[18] S. U. Khan and M. Niazi, "A preliminary structure of software outsourcing vendors' readiness model," presented at Proceedings of the 11th International Conference on Product Focused Software, 2010.

[19] S. U. Khan, M. Niazi, and R. Ahmad, "Barriers in the selection of offshore software development outsourcing vendors: An exploratory study using a systematic literature review.," Information and Software Technology, vol. 53, pp. 693-706, 2011.

[20] M. I. Azeem and S. U. Khan, "Intercultural Challenges Mitigation Model for Software Development Outsourcing Vendors," International Journal of Advanced Science and Technology, vol. 47, 2012.

[21] Darja Smite, Claes Wohlin, Tony Gorschek, and R. Feldt., "Empirical evidence in global software engineering: A systematic review," Empirical Software Engineering., pp. 91-118, 2010.

[22] S. Darja, W. Claes, G. Tony, and F. Robert., "Empirical evidence in global software engineering: A systematic review," Empirical Software Engineering., pp. 91-118, 2010.

Software Re-engineering Risks Management Model
(SRE-RMM)

Nasir Rashid
Department of Computer Science and I.T
University of Malakand
KPK, Pakistan
nrswat1280@gmail.com

Dr. Siffat Ullah Khan
Department of Computer Science and I.T
University of Malakand
KPK, Pakistan
Siffat7@yahoo.com

***Abstract*— Software re-engineering is important for putting high software maintenance costs under control, recovering existing software assets and establishing a base for future software evolution. However software re-engineering process is not a risk free activity and has put many new challenges to software maintenance programming community. The objective of this research study is to empirically identify the critical risks and success factors while re-engineering legacy software systems and finally to develop a Software Re-engineering Risks Management Model (SRE-RMM) on the basis of these findings. The primary goal of this model is to help software maintenance engineers to effectively design and transform legacy systems using re-engineering strategies to generate successful target systems according to the changing demands of the customers with reduced risks. Systematic Literature Review (SLR) and empirical study will be used throughout this research project by the authors for customized search strings that will be derived on the basis of defined research questions..**

***Index Terms*— Software re-engineering, risks management model, legacy systems.**

I. INTRODUCTION

Software Re-engineering is the examination, reorganization, analysis and alteration of an existing software system. It helps to make software more maintainable and to reconstitute it in a new form and the subsequent implementation of the modified system. This process involves the restructuring or recoding of a component or all parts of legacy system without affecting its functionality [1].

Re-engineering is a combination of other processes such as reverse engineering, re-documentation, translation, and forward engineering. The main purpose is to understand the specification, design, and implementation of the legacy system and then to implement its modernized form to improve its overall functionality and performance [2].

Re-engineering a legacy software system has some major advantages over more fundamental approaches to system evolution.

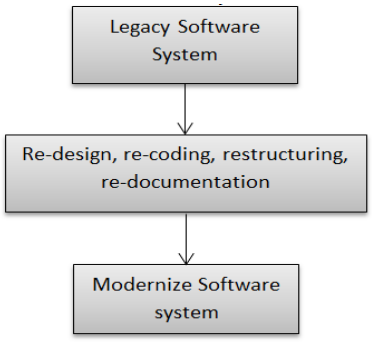

Fig 1. Basic process of Software Re-engineering

Software re-engineering helps in reducing the cost of writing a system from the scratch by using functionality of the legacy software. According to [3] risks/barriers should be considered as an important factor when redeveloping end re-engineering the existing system. Some of the success factors of software re-engineering are as follows:

- It helps to improve the business value of modernize application software by reiterating the critical data and business logic.
- It also helps in increasing system reliability and customer satisfaction.
- Reduces the operational costs.
- Efficient use of the existing staff.
- It improves system performance, accessibility and system documentation.
- It enhances software functionality [4].

The evolution of software systems need to be effective in software development and maintenance environment. The need for software re-engineering is increasing day by day, as legacy systems have become outdated in terms of their architecture, the platforms on which they run, and their feasibility for support evolution in order to support changing needs [5].

Software re-engineering, which is considered as one of the integral process of software maintenance and is the most costly phase of the software lifecycle. It has been estimated

that approximately 60 percent of the efforts in software life cycle has been related to software reengineering and maintenance [6].

II. GENERAL MODEL FOR SOFTWARE RE-ENGINEERING

The process of Re-engineering begins with the conceptual understanding and redesign of the legacy software systems and to re-implement it to improve the performance and to maintain the existing functionalities of the system for greater maintainability [7]. Figure 2 depicts the abstract model of software re-engineering that illustrates the basic activities involved.

In *forward engineering* informal requirements of existing software are converted into a semi-formal specification using notations regardless of the underlying precise semantics like entity-relationship diagrams, data-flow diagrams, natural language descriptions, or other problem oriented informal notations. The program is then designed and coded manually from the according to the specification by software programmer [8].

Reverse engineering deals with the study and analysis of existing software systems with an aim to better understand the structure and design of software as well as to support software maintenance and evolution activities through enhanced program of comprehension. In this process the system's components and their interrelationships are identified in order to develop more usable representations of existing system in higher level of abstraction [9].

Re-documentation is one of the crucial steps in software re-engineering that involves the revision of a semantically equivalent representation of existing software systems within the same level of abstraction. Data flow, control flow and entities relationship diagrams must be documented after restructuring [10].

Restructuring is the transformation of a software system from one representation to another that includes splitting, merging and reorganizing the existing software components without effecting the external behavior and functionalities [11].

Recode is concerned with changing the implementation characteristics of the source code. It involves translation of the source code to more object oriented programming. In addition to this, code standard and readability of existing system is also improved [12].

Redesign involves changing the design characteristics and interfaces of existing software systems to more interactive designs. Possible changes include design and architecture restructuring, redesign of the existing data model as incorporated in data structures and improving the efficiency of algorithms [13].

Re-specify involves changing the functional requirements of legacy software systems. This phase includes changing only the form of current requirements (i.e. to generate formal specifications from the informal requirements that have been collected in any contrived language such as Z or other uniform modeling language) [14].

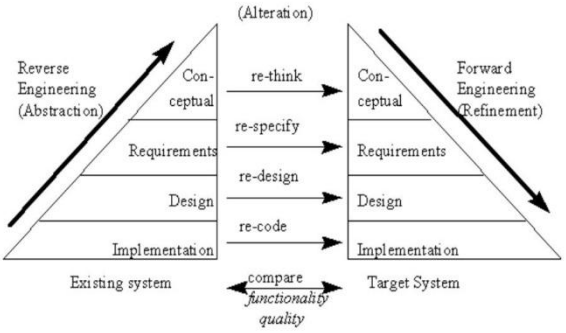

Fig. 2. General Model for Software re-engineering [13]

III. BACKGROUND AND RELATED STUDIES

Re-engineering process involves applying re-engineering principles to an existing system to meet new requirements. However, development of successful re-engineering efforts need take into consideration the re-engineering problems and risks involved from a number of different perspectives in various levels of software development organizations.

Harry M. Sneed in [15] estimates risk of re-engineering project in terms of project completion rates and cost overruns [16]. However re-engineering efforts also require to consider other factors like performance improvement, resource utilization, quality goals, user satisfaction etc.

As we know that all software do not have the same level of complexity. Different software has different complexities according to their functions. One of the factor for complexity estimations is the number of kilo source line of code (KSLOC). According to E.M Benettan [17] the engineering efforts required for the design and development can be calculated in terms of software engineering month (SEM) as shown in equation 1.

$$\textit{System Software: } SEM=3.6 \times (KSLOC)^{1.20}$$

$$\textit{Data Processing Software: } SEM=2.4 \times (KSLOC)^{1.05}$$

$$\textit{Algorithmic Software: } SEM=3.2 \times (KSLOC)^{1.15}$$

$$\textit{Service Software: } SEM=2.8 \times (KSLOC)^{1.10} \quad (1)[17]$$

So the SRE-RMM will be designed and developed having functionalities to deal with all categories of the software having different complexities. Reengineering efforts will be automated through the automated functions and procedures.

Eric K. Clemons in [18] suggests that two principal reasons for failure of re-engineering efforts are functionality risk and political risks. Though there are other serious risk involved such as technical risk, process risk, development environment risk, architecture risk, and risk related to stakeholders are also needed to take into consideration.

According to the current research, 88 percent of various organizations are involved in re-engineering the existing software systems. However failure rates of re-engineering activities for various types of software systems are at maximum. The reason being that most of software engineers failed to understand the conceptual needs and framework of software systems. Many firms either leave behind their huge efforts after making significant investment or fail to achieve desired outcomes from their re-engineering strategies due to unidentified risks.

Anand Rajavat [19] provides details of the major risks/barriers that need to be addressed in re-engineering process. Risks have been categorized into different domains according to its nature i.e. system, managerial and technical domains. Some of the major risks include resource, deployment, team, financial, performance, modularity and development process risks etc.

Varun Grover [20] argues that although the concept of re-engineering as a "radical change" is already evolving to a more moderated and possibly more realistic stance but re-engineering is not a risk free process to be conducted. Little research has been carried out to date. To evaluate on a broad basis the implementation efforts of software engineers that have pursued and successfully concluded re-engineering projects. This is of more concern according to the recent evidence that a large number of re-engineering efforts required for transforming legacy systems into more reliable and modern software systems.

According to Eric K [21], existing experience in re-engineering reveals that most of the transformation of software systems fail for the reason of unrelated technical abilities of software engineers to implement software systems. The two major risks addressed for more attention are functionality and political risks. These risks are the result of conflict among the organization's existing strategies. Risks information systems must have to be developed in order to cope with them accordingly for more robust re-engineering objectives.

IV. AIMS AND OBJECTIVES

The main objective of this research study is to develop Software Re-engineering Risks Management Model for software engineers and practitioners. This model will assist software engineers in evaluating the risks during re-engineering activities in terms of designing, code translation, structure migration, functionality enhancement for better maintainability and reliability of legacy systems so as to cope with new requirements. This model will also help them to successfully reengineer legacy systems by managing all the possible risks encountered during transformation so that we may not only in position to modernize the system, but also in a better position to provide quality to legacy system in a repeatable and productive manner in the form of a software tool.

V. RESEARCH QUESTIONS

There are six research questions that have motivated the work reported in this research project.

RQ1. What are the success factors, as identified in the literature, do software engineers need to address for reengineering legacy software systems ?

RQ 2: What are the risks/barriers faced by software engineers as identified in the literature, in transformation of legacy software systems in order to cope with new requirements?

RQ 3: What factors, as identified in real-world practice, do software engineers need to address for reliable transformation of legacy software systems?

RQ 4: What are the risks/barriers faced by software engineers as identified in real world practice, in transformation of legacy software systems in order to cope with new requirements ?

RQ 5: Are their differences between the factors and risks identified through literature and real world practices?

RQ 6: How can a practically robust Software Re-engineering Risks Management Model (SRE-RMM) be developed and adopted by software engineers?

Software re-engineering risks management model will be developed based on an empirical analysis of software reengineer's experiences and perceptions of factors that can have a positive or negative impact on the transformation of legacy software systems to a new reliable target systems for better performance and maintainability. We will automate the risks management model in the form of a software tool in order to facilitate software engineers in assessing the transformation process.

VI. RESEARCH METHODOLOGY

A. Data Collection and Analysis

In this research project, two types of data will be collected: firstly, the factors that can have a positive or negative impact on the transformation of existing software systems; and secondly, how one can implement these factors. For the identification of these factors, systematic literature review and empirical studies will be conducted. "A systematic literature review is a mean of identifying, evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest" [22]." SLR has become an important research methodology in software engineering, since the introduction of evidence-based software engineering (EBSE) in 2004. In order to conduct SLR for our desired results we have first designed a protocol that will be implemented in order to get the success factors and risks for answering RQ1 and RQ2. Empirical study will be conducted for the validation of SLR findings to answer RQ3, RQ4 using questionnaire survey and focus group session.

B. SRE-RMM Model Development

Various phases involved in designing and development of software re-engineering risks management model (SRE-RMM) are shown in Figure 3.

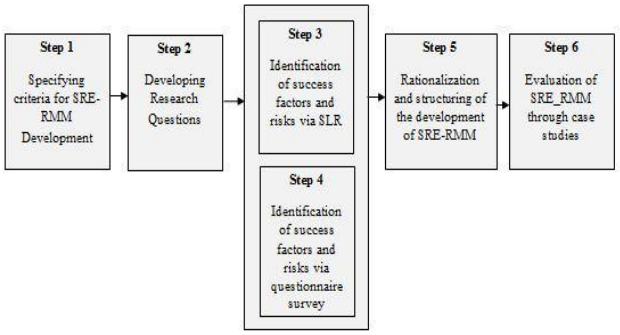

Fig. 3. Stages in development of SRE-RMM

C. SRE-RMM Evaluation

We aim to adopt the case study approach for the evaluation and review of the proposed model i.e. SRE-RMM. Case study has been proven to be a powerful tool for evaluating real world practices and information. For this research project, four case studies will be carried out at different software development organizations of various levels to assess the efficiency and effectiveness of SRE-RMM. For feedback, focus group sessions have been planned, that will be attended by the software engineers having at least five year practical experience in designing and development of software in various organizations in order to evaluate the model according to their level of expertise.

VII. NOVELITY AND TIMELINESS

Software engineers are facing many risks while transforming existing systems to new software systems according to the needs of organizations using various strategies. The reason being that there is a lack of reliable risks management model to cope with the possible risks that do come in transformation process. Research suggests that many companies have failed to realize the desired results. There are many reasons that transformation of legacy systems fail. Some of the risks include reliability, usability, security, training, technology and project complexity risks etc. Despite the importance of evolution of software systems, little empirical research has been carried out in this context. An empirical study in this research project has been suggested in order to provide useful insights about the factors that can have a positive or negative effect on re-engineering process.

The desired outcome of this research project will be software re-engineering risks management model to assist the software engineers in design, coding, testing and support for effective software re-engineering process of legacy systems. SRE-RMM is a risk mitigation and management model that attempts to address the software re-engineering risks during the transformation of existing legacy software systems and tries to modernize the systems and to provide quality to legacy software systems in a repeatable and productive manner through software.

VIII. CONCLUSION

Legacy software system re-engineering is a renowned system reforming technique which helps in effective design of cost control, quality improvements, time and risk reduction. Re-engineering is well-known technique to develop optimum version of software legacy systems to satisfy the end users with changing demands. However different projects can have different level of complexities and hence may face a lot of risks during the re-engineering. Over the past few years of software maintenance and evolution, software engineers are facing various risks that encounters in re-engineering process. It would be strategically beneficial for software engineers if they could predict the root causes to cope with re-engineering risk so that re-engineering may provide every attempt to grab and maintain true competitive advantages. Evolution of legacy systems through re-engineering requires attention to critical risks of re-engineering. Proposed SRE-RMM model addresses software re-engineering risks from different perspectives and analyzes strengths of those risks to mitigate. SRE-RMM will be adopted as a standard model as well as in the form of software by software engineers of CMMI level organizations for successful re-engineering of legacy systems.

ACKNOWLEDGMENT

We are very thankful to our PhD supervisor Dr. Siffat Ullah Khan for reviewing our PhD research proposal and providing valuable suggestions.

REFERENCES

[1] Cross.J.H, C.E.J., Reverse Engineering and Design Recovery : A Taxonomy. IEEE Software, 1990. 23: p. 13-17.

[2] Rosenberg, D.L.H., Software Re-engineering, Software Assurance Technology Center. p. 2-3.

[3] Nguyen, P.V., The Study and Approach of Software Re-Engineering. arXiv preprint arXiv:1112.4016, 2011.

[4] Rashid, N., et al., Analysis of Risks in Re-Engineering Software Systems. International Journal of Computer Applications, 2013. 73(11): p. 5-8.

[5] Joshphin Jasaline Anitha.X, M.K., K. Alagarsamy The Classification and Analysis of Risks in Reengineering System. international Journal of Computer Applications (0975 – 8887) 2012. Volume 39(8).

[6] Koschke, R., Software visualization in software maintenance, reverse engineering, and re‐engineering: a research survey. Journal of Software Maintenance and Evolution: Research and Practice, 2003. 15(2): p. 87-109.

[7] Manzella, J. and B. Mutafelija. Concept of re-engineering life-cycle. in Systems Integration, 1992. ICSI '92.,

Proceedings of the Second International Conference on System Integration. 1992.

[8] Waheed, M.A., A practical approach to re-engineering software, in Proceedings of the 1992 conference of the Centre for Advanced Studies on Collaborative research - Volume 11992, IBM Press: Toronto, Ontario, Canada. p. 295-298.

[9] Xiaomin, W., et al. A reverse engineering approach to support software maintenance: version control knowledge extraction. in Reverse Engineering, 2004. Proceedings. 11th Working Conference on Reverse Engineering (WCRE'04). 2004.

[10] Bennett, K.H. and V.T. Rajlich. Software maintenance and evolution: a roadmap. in Proceedings of the Conference on the Future of Software Engineering. 2000. ACM.

[11] Lung, C.-H., M. Zaman, and A. Nandi, Applications of clustering techniques to software partitioning, recovery and restructuring. Journal of Systems and Software, 2004. 73(2): p. 227-244.

[12] Yang, H., et al., Software Evolution for Evolving ChinaSoftware Evolution Software Evolution for Evolving China Evolving China, in The China Information Technology Handbook. 2009, Springer. p. 1-33.

[13] Abbas, A.S., W. Jeberson, and V. Klinsega, The Need of Re-engineering in Software Engineering. International Journal of Engineering and Technology, 2012. 2(2).

[14] Yang, H. and M. Ward, Successful evolution of software systems. 2003: Artech House.

[15] Sneed, H.M. Estimating the costs of a reengineering project. in Reverse Engineering, 12th Working Conference on Reengineering. 2005.

[16] Sneed, H.M. Risks involved in reengineering projects. in Reverse Engineering, 1999. Proceedings. Sixth Working Conference on Reverse Engineering. 1999.

[17] Bennetan, E.M., Software Project Management A Practitioner' Approach McGraw-Hill International, UK.

[18] Clemons, E.K., M.C. Row, and M.E. Thatcher. An integrative framework for identifying and managing risks associated with large scale reengineering efforts. in System Sciences, 1995. Proceedings of the Twenty-Eighth Hawaii International Conference on. 1995.

[19] Rajavat, A. and V. Tokekar, ReeRisk –A Decisional Risk Engineering Framework for Legacy System Rejuvenation through Reengineering, in Computer Networks and Information Technologies, V. Das, J. Stephen, and Y. Chaba, Editors. 2011, Springer Berlin Heidelberg. p. 152-158.

[20] Teng, J.T., S.R. Jeong, and V. Grover, Profiling successful reengineering projects. Communications of the ACM, 1998. 41(6): p. 96-102.

[21] Clemons, E.K., M.E. Thatcher, and M.C. Row, Identifying sources of reengineering failures: A study of the behavioral factors contributing to reengineering risks. Journal of Management Information Systems, 1995. 12: p. 9-36.

[22] Kitchenham, B.A. and S. Charters, Guidelines for performing systematic literature reviews in software engineering. 2007.

A Review on Web Design Processes: How They Incorporated in a Complex System

Siti Azreena Mubin, Azrul Hazri Jantan
Faculty of Computer Science & Information Technology
Universiti Putra Malaysia
Selangor, Malaysia
GS35679@mutiara.upm.edu.my, azrulhazri@upm.edu.my

***Abstract*—The process of designing Web Application (WA) is one of the big challenges nowadays. The designs should incorporate conceptual domain, complex process and interactions, hyperlinks structure and abstract presentation designs. The increasing complexity has raised several issues and raised many designing methods in WA. Thus, this paper will discuss on a number of common models in web design methods. We claimed that not much of web design models concentrate on the complexity of process flow design in Academic Management System (AMS). Our main objective is to describe how those complex process flows should be treated in AMS. A new design process will be proposed and briefly summarized at the end of the paper.**

***Index Terms*—Web application, web design, design processes, academic management system, complex processes.**

I. INTRODUCTION

These days, Web Application (WA) is one of the greatest approaches and most powerful tool being used to convey information and to handle extensive of data to a user. Thus, these conditions promise WA to be applied in most web management system such as in online banking, business, education institution and so on [1].

As WA getting complicated and demanding, practitioners believed a systematic design and comprehensive process of WA need to be developed well to support the complexity of the process [2]. Therefore, some design issues have been identified in web design such as in navigation process, complex business process and conceptual design process [3, 4].

In order to deal with those issues, several web design models have been introduced and proposed by researchers for the past few years such as Object-Oriented Hypermedia Design Method (OOHDM) [5], UML-based Web Engineering Design Method (UWE) [6, 7], Web Site Design Method (WSDM) [8, 9], Navigational Development Techniques [10], Web Modeling Language (WebML) [11] and the most recent web design model is known as Comprehensive Hypermedia Design Method for Complex Process Modeling (COM$^+$HDM) [12].

Each WA required a different design model for their structure and application [13]. They might need extra features and suitable design models according to their user requirements. Hence, in this paper, our objective is focusing on the implementation of web design models in Academic Management System (AMS). Besides, independent module issues in AMS will lead to poor navigation results [14]. Thus, a systematic process design needs to be proposed in AMS.

Academic Management System (AMS) can be defined as a web application that is capable to assist an education institution to manage student's records. The system mainly designed to handle academic matters [19] such as course registration, tuition fees payment, attendance and examination. Therefore, different WA will require different design method and the design method should be suited well with the application and web functionality [13]. However, it is very challenging to find a design method that fit for the development of any web engineering system due to rapid change in the technology and combination of different techniques and fields [10, 13].

AMS is created in huge databases which contain interrelated process. This interrelated process usually identified as collection of information from various module and the process are dependent with other process. As an example of attendance module, collection of students name registered for particular course exclusively depends on course registration. Nevertheless, in exam module, qualified students to sit for an exam will depends on attendance module. This kind of interrelated process could contribute to poor functionalities of AMS if it is not systematically designed during the web design process.

Existing WA design models still not adequate to support different way of dealing with complex process specifically in AMS. Generally, there are two main features to differentiate complex process in AMS from a general process. First, AMS complex process flow depends on user's input or user dependent processes. There are many type of users in AMS such as students, lecturers, programme leaders, finance personnel and etc. Each of them will give different type of inputs to the system. Hence, the input will be processed according to the input given by the respective users. Secondly, complex process can be defined as a process flow that has set of activities which need to be executed in a defined time frame. In example of course work assessment, the marks and grading need to be keyed in according to a specific time period defined by the respective departments. This action as to ensure students' results will be released based on the date scheduled.

The main objectives of this paper are to describe and review a few of recognized web design model for the past few years as well as to propose design process model in AMS. The structure of this paper is as follows. Section II will review on WA designs that have been introduced by earlier researchers. Section III will explain on complex process in AMS and it will discuss how our proposed design methods able to cope with the difficulties of the complex process in AMS and conclusion will be summarized in section IV as well as the future research works.

II. WEB APPLICATION DESIGN PROCESS

In the last few years, several web design models have been introduced and implemented among the practitioners such as OOHDM [5], UWE [6, 7], WSDM [8, 9], NDT [10], RMM [15], OO-H [16] and W2000 [17]. Most of the design models comprise of four phases which known as user requirement designs, conceptual designs, navigational design and presentation designs. Table 1 shows the summary of respective design process in the particular design models.

Table 1: A Summary of design phases in the design models

Design Model	*User Req. Design*	*Concept. Design*	*Navigational Design*	*Presentation Design*
OOHDM	✓	✓	✓	✓
UWE	✓	✓	✓	✓
WSDM	✓	✓	✓	✓
NDT	✓	✓	✓	
RMM	✓		✓	✓
OOH	✓	✓	✓	✓
W-2000	✓		✓	
COM$^+$HDM	✓	✓	✓	✓

In this paper, we will review four web design models, namely OOHDM [5], UWE [6, 7], NDT [10] and COM$^+$HDM [12] which we believe the design process are quite similar and certain processes might be practical to be implemented in our on-going proposed design process.

A. Object-Oriented Hypermedia Design Method (OOHDM)

Object-Oriented Hypermedia Design Method (OOHDM) [5] is an approach that is suitable in designing different kinds of applications such as multimedia systems and information systems. It is a model-based method which suitable to be designed in huge hypermedia applications. OOHDM involves four different phases which are known as conceptual design, navigational design, abstract interface design and implementation. One of the advantages of OOHDM is the separate phases of design enable reusable and flexible development of the design model. Figure 1 illustrates OOHDM design method.

Conceptual design is the first phase in OOHDM which captures the ideas or the static aspect of the particular system. During the design, a similar UML notation is being used as to construct the model in the application domain.

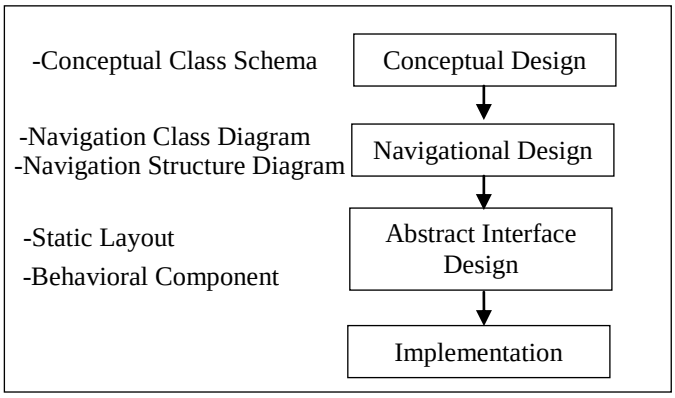

Fig. 1. Overview of the OOHDM design process [5].

In navigation design, it consist two navigational diagrams which are called navigation class diagram and navigation structure diagram. Navigation class diagram is design as to signify the static possibilities of navigation which exist in the system while navigation structure diagram represents contexts of navigation and access structures.

Third phase in OOHDM is known as abstract interface design. This particular design model identifies user interface design in abstract way which shows interface objects and how users observe it. For the last phase called implementation, web application is implemented and the designer implement the design based on the models build from previous phase.

B. UML-based Web Engineering Design Method (UWE)

UML-based Web Engineering Design Method (UWE) [6, 7] supports WA development with motivation on personalization, systematization and semi-automatic generation. UWE is based on object-oriented, iterative and incremental approach which centered on Unified Modeling Language (UML) [18]. It has four main designs process called Requirement Analysis, Conceptual, Navigation and Presentation Design and it shows in figure 2.

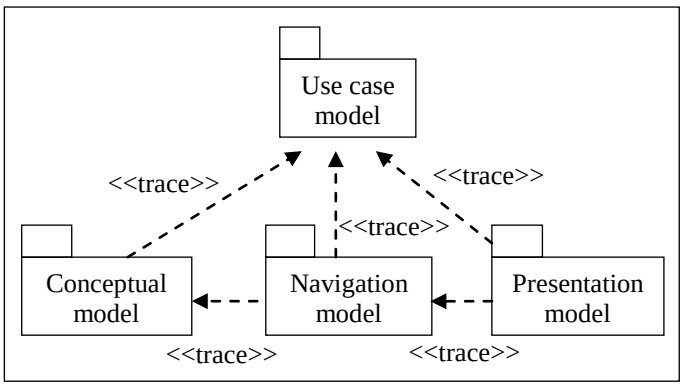

Fig. 2. Overview of the UWE design process [6].

In requirement analysis, use cases are being used for capturing the system's user requirements. User-centered technique applied in this design phase defined the actors involved in the application and represent activities as well as

the functionality of each actor. Conceptual design is derived from requirement analysis phase. The main modeling elements in this phase are class, association and package. All the modeling elements are graphically presented using the UML notation [18]. The class model consists of attributes, operations and variants while the association and package are implemented as in standard UML class diagrams.

Navigation design phase comprises two components which are called the navigation space model and the navigation structure model. For navigation space model, it identifies which objects can be visited and linked through the web application. As for navigation structure model, it defined how these objects can be reached in the navigation process. Presentation design phase is based on the navigation structure model constructed from previous phase. In this phase, UML composition notation for classes and stereotypes classes are being used in a class diagram. The idea of this presentation design is to present where and how navigation objects will be presented to the user.

C. Navigational Development Techniques (NDT)

Navigational Development Techniques (NDT) [10] is divided into three main stages called requirement treatment, analysis and prototyping. These three stages support methodological process for web application development which focused on the requirement and analysis phases. Figure 3 illustrates NDT concept.

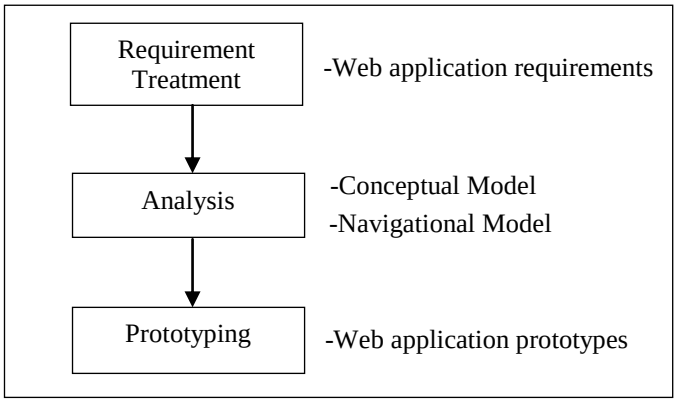

Fig. 3. Overview of the NDT design process [10].

In requirements treatment, web application requirements are collected and described extensively in textual templates. In analysis phase, analysis models will be systematically constructed from the requirements specification and it consist of conceptual model and navigational model. In prototyping, web application prototypes development is derived from analysis models and these prototypes are used for requirements validation.

D. Comprehensive Hypermedia Design Method for Complex Process Modeling (Com+HDM)

Comprehensive Hypermedia Design Method for Complex Process Modeling (Com+HDM) [12] is one of the latest design models in WEA which aim in complex process modeling issues. Com+HDM incorporates three design phases called conceptual design, navigation design and user interface design. Com+HDM design modeling is able to be executed in increment and iterative design manner. The object-oriented modeling used in Com+HDM is based on UML modeling paradigm [18]. Figure 4 illustrates Com+HDM design modeling.

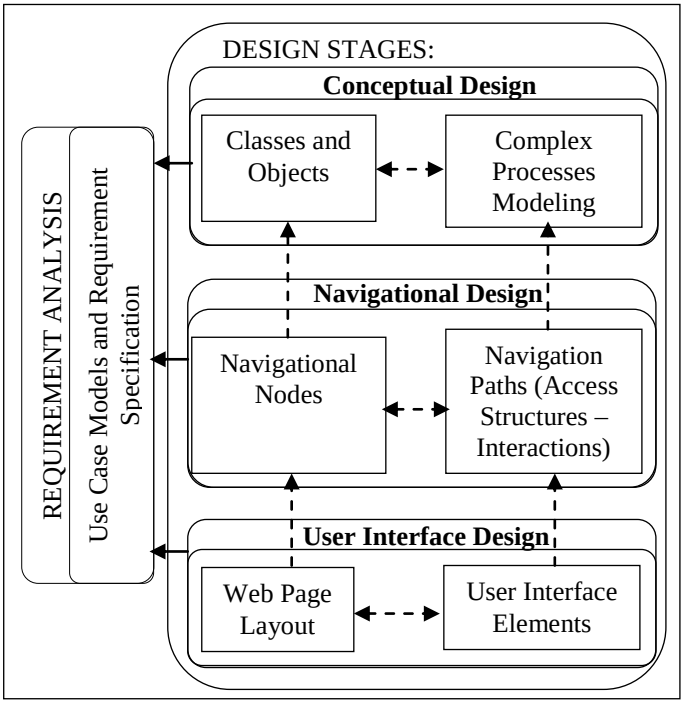

Fig. 4. Overview of the COM+HDM design process [12].

Conceptual design aims to present the application domain structures. It comprises of three domain called Conceptual Class Model (CCM), Complex Process Flow Model (CFM) and Conceptual Process Model (CPM). CCM is built according to UML Class Diagram standard. CFM is created as to design the flow of the complex process. The third stage is CPM which the designer can view details of classes in the complex process and how conceptual classes and the complex processes can be improved in order to construct the next phase named navigation models.

Navigation design is built in order to show how the navigation is conducted in WA. It offers the navigation design with four UML-based stereotypes which are known as navigation class, interaction class, hyperlinks and access structures. Thus, the implementation in navigation design phase is basically to present the details in navigation node including the interaction class, hyperlinks and access structures. In this design stage, Navigation Class Model and Navigation Access and Interaction Model are developed.

User interface design is focus on page layout presentations. In Com+HDM, the user interface elements are defined through graphical notations and the objective is to provide layout on how navigational model designed from previous phase able to be structured in WA. In this phase, user interface layout is being designed by the web developer according to his creativity and skills.

III. DISCUSSIONS: A SUPPORT FOR COMPLEX PROCESS IN AMS

Current research has shown that complex processes must be considered and relied on each design process. Our research idea is to build design method for AMS process flow in WA which starts from the earlier stage of the design process. We are not intend to have a new design method that develop from scratch, yet we compare and combine several design aspects from existing WA design models as to improve our propose design method. Hence, we propose the following design processes as concern to consider how complex process can be systematically modeled in a design method.

- User Requirement Design: Collect all information needed from potential users for AMS functional requirements and navigation preference requirements. In our AMS case study, potential users will be lecturers, students, finance personnel, exam unit personnel and student affair personnel. Their role and activities will be defined accordingly as to have a clear structured of user requirements models. The models will be constructed via Unified Modeling Language (UML) use case models.
- Conceptual Design: This design stage concentrates on information structures in AMS. It will present which information should be made available to AMS users, including classes, objects, attributes and association using UML standards. AMS complex process flow should be able to be identified at this stage. In AMS, the classes will be interrelated with each other. For example, course registration class need to correlate with attendance class which in this stage, all attributes, objects and association for course registration class and attendance class should properly captured and defined for next design stage model.
- Complex Process Design: In this new propose design stage, AMS complex process flow identified from conceptual design stage should be designed in more details and expanded into a much deeper view. The idea of this process is to capture in details complex process involved between interrelated classes. This particular design process should be modeled in a different way compare to other common processes in AMS. For example, course assessment class only will be opened in a defined period and lecturers need to key in students' marks according to the predefined period set up by the administrator. This activity involved need to be expanded in more specifics and structured in this stage. Using UML, the process can be detailed using sequence diagram and activity diagram. Additionally, appropriate UML stereotypes can be used to model the flow of complex process in this particular design stage.
- Navigation Design: The main objective of this design stage is to identify how information in complex process design stage is presented as navigation classes and defines the link directions for every class. For all classes determined in previous design stage, designers need to defined appropriate navigation path to the users to navigate in the real situation. Navigation access structures should be applied to each link to show how the users will navigate and reached those links. In AMS complex flow process, these navigation links will require a number of access elements to be presented such as index, guided tours and query. Besides, new UML's stereotype profiles will be introduced to ease the AMS complex process flow navigation structures designed.
- User Interface Design: This final design stage is built to present abstract user interface elements integrating the layout and visual characteristics information in AMS. Basically the idea is to present a screen layout including the navigation structures. User interface design stage not emphasize on the look and creativity of the screen layout. Nevertheless, it concern about how AMS navigation class and access structure are well configured especially among interrelated classes. A well-structured mechanism will be proposed in this particular design stage as to ease designers' implementation work later on.

We believe that by adding new stage called Complex Process Design will ease the design process of AMS and meet the functionalities requirement, well-structured navigation and proper user interface interaction of a campus management application in web design process. Figure 5 illustrates our propose design model.

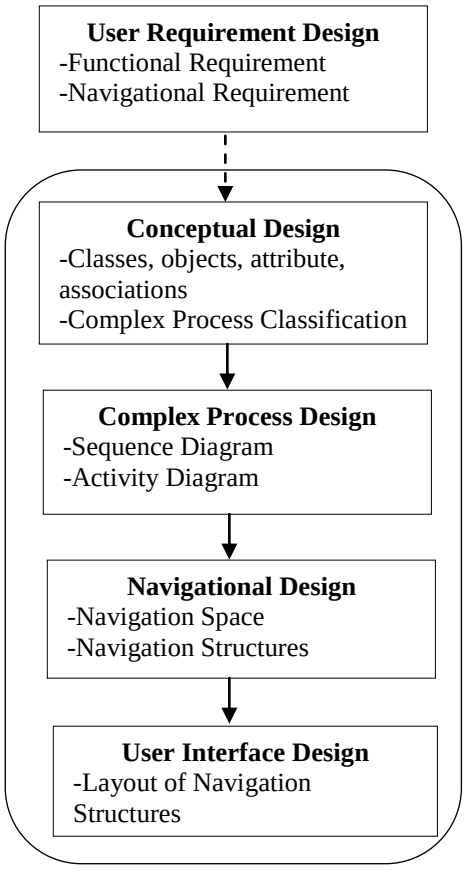

Fig. 5. Propose design model.

IV. CONCLUSION AND FUTURE WORKS

This paper has studied several web design models. Four design models have been chosen which called OOHDM [5], UWE [6, 7], NDT [10] and COM$^+$HDM [12]. The reasons behind these four web design models have been chosen are the similarity design aspects, graphical notation and the appropriateness of the models to be designed in AMS.

Some methods such as OOHDM and UWE focus more on systematic phases design model but less attention given to complex process. In NDT design model approach, it provides mechanism to properly describe the requirements of a web application. However, the design models not emphasize on complex process design and the use of textual templates could overload the requirements specification in complex web applications.

With regard to complex process, COM$^+$HDM has developed a design model which introduced UML-based modeling method and to extend the UML profiles into UML stereotypes as in order to provide more specific and expressive modeling facilities. In contrast, defining a large amount stereotypes could lead to massive confusion if they are not properly defined at earlier stage. Four design models studied use a well-known UML notation which offers advantages in terms of compatibility and easy to use features. Nevertheless, notation such as UML also has been greatly used by many case tools [7].

The issue arises when most of the WA not emphasize on complex process design model. According to our research and review done from previous section, we argue that involvement of complex process in WA could raise a number of functionality and navigational issues.

Our future work is to solve complex process design specifically in AMS. In AMS, complex process can be defined as interrelated process between modules such as course registration, attendance and exam. Thus, we propose five stages of design process and all these stages need to be systematically designed in the future. Involvement of complex process design stage will be able to solve complex process flow issues as the specific details of process flow will be designed and modeled in refined approach. Proposed design model will be relied on the object oriented approach called Unified Modeling Language (UML). Through this propose design model, we believe that our design approach could solve complex process design issues in WA specifically in AMS.

REFERENCES

[1] Murugesan, San, and Athula Ginige, Web Engineering: Introduction and Perspectives, Web Engineering: Principles and Techniques 1, 2005.

[2] M. Lang and B. Fitzgerald, "Hypermedia systems development practices: A Survey", IEEE Software, Volume 22, Issue 2, March-April, 2005, pp.68-75.

[3] Schmid, H.A. and Rossi, G., "Modeling and designing processes in E-Commerce applications", IEEE Internet Computing, Vol. 8, Issue 1, 2004, pp. 19-27.

[4] Selmi, S.S, Kraiem, N., and Ghezala, H.B., "Toward a comprehension view of web engineering", International Conference on Web Engineering (ICWE05), Australia, 2005, pp. 19-29.

[5] Schwabe D. and Rossi G., "Developing hypermedia application using OOHDM", Proceedings of Workshop on Hypermedia Development Process, Methods, and Model, 1998.

[6] Koch N., Kraus A., and Hennicker R. , "The authoring process of the UML-based web engineering approach", In Daniel Schwabe, 1st International Workshop on Web-oriented Software Technology (IWWOST01), online publication, June 2001.

[7] Koch N. and Kraus A. , "The expressive power of UML-based web engineering", 2nd International Workshop on Web Oriented Software Technology (IWWOST02), D. Schwabe, O. Pastor, G. Rossi, and L. Olsina, editors, CYTED, June 2002, pp. 105-119.

[8] De Troyer, "Audience driven web design in information modeling in the new millennium", Eds. Matt Rossi & Keng Siau, Publi., IDEA GroupPublishing, ISBN 1-878289-77-2, 2001.

[9] De Troyer and Leune C. J. , "WSDM: A user-centered design method for web sites", Computer Networks and ISDN Systems, Proceedings of the 7th International WWW Conference, Elsevier, 1998, pp. 85-94,

[10] Valderas, P., and Pelechano, V., "A survey requirements specification in model-driven development of web applications", ACM Transactions on the Web, Vol. 5, No. 2, May 2011.

[11] S. Ceri, P. Fraternali, A. Bongio, "Web modeling language (WebML): A modeling language for designing web sites", Computer Networks, Vol. 33, 2000, pp. 137-157.

[12] A. H. Jantan, P. Sumari, and S. Sulaiman, "Com+ HDM: Extending UML profiles for modeling complex web hypermedia applications", 2008 International Conference on Advanced Computer Theory and Engineering, 2008, pp. 290-294.

[13] M. Lang, "Issues and challenges in the development of hypermedia information systems", Proceedings of 11th Annual Business Information Technology Conference (BIT 2001), Manchester, England, October 30-31, 2001.

[14] Memon, Qurban and Khoja, Shakeel, "Academic program administration via semantic web – A case study", *International Conference on Computer, Electrical, and Systems Science, and Engineering, CESSE 2009*, Dubai, UAE, 28 - 30 Jan, 2009.

[15] T. Isakowitz, A.S Edward, P. Balasubramaniam, "RMM: A methodology for structured hypermedia design", Communications of the ACM, Vol. 38, Issue 8, 1995, pp. 34-44.

[16] J. Gomez and C. Cachero, "OOH-Method: Extending UML to model web interfaces", Information Modeling for Internet Applications, 2003, pp. 144-173.

[17] L. Baresi, F. Garzotto, and P. Paolini, "Extending UML for modeling web applications", Proceedings of the 34th Hawaii International Conference on System Sciences, 2001.

[18] G. Booch, I. Jacobson, J. Rumbaugh, "The unified modeling language user guide", The Addison Wesley Object Technology Series, 1998.

[19] Hassan, Wan Azlan Wan, et al., "An assessment study on usability of academic management system", User Science and Engineering (i-USEr), 2010 International Conference on. IEEE, 2010.

Predicting Hearing Loss Symptoms from Audiometry Data Using Machine Learning Algorithms

Nasir G. Noma
Biomedical Computing & Engineering Technologies (BIOCORE Malaysia) Applied Research Group,
Faculty of Information and Communication Technology,
Universiti Teknikal Malaysia Melaka
76100 Hang Tuah Jaya, Durian Tunggal, Melaka
Malaysia
nasnoma@hotmail.com

Mohd Khanapi Abd Ghani
Biomedical Computing & Engineering Technologies (BIOCORE Malaysia) Applied Research Group,
Faculty of Information and Communication Technology,
Universiti Teknikal Malaysia Melaka
76100 Hang Tuah Jaya, Durian Tunggal, Melaka
Malaysia
khanapi@utem.edu.my

Abstract-Hearing loss symptoms are detected by clinicians after the patient undergoes audiometry test conducted by an audiologist to determine the presence, severity and type of hearing impairment. This is often conducted via manual procedure of diagnosis that involves questions and other check-up procedures. This process increases the time spent on consultation thereby resulting to longer waiting time for other patients that needs to be attended to. A number of studies have justified connection between audiometry thresholds and some attributes in medical records of hearing loss patients, but none of these studies have analysed these audiometry thresholds to find any relationship between pure-tone audiometry thresholds and hearing loss symptoms or utilized these relationship to predict any of those attributes. This research focuses on finding relationship between pure-tone audiometry thresholds and hearing loss symptoms, and utilizing this relationship to predict hearing loss symptoms with high accuracy using combination of supervised and unsupervised machine learning algorithms.

Keywords **- Audiometry, thresholds, naïve Bayes, FP-Growth, sensorineural**

I. INTRODUCTION

There is potential knowledge inherent in vast amounts of untapped and possibly valuable data generated by healthcare providers. So often, clinicians rely in their skills and experience and that of other medical experts as their source of information. The healthcare sector is now capturing more data in digital and non-digital format that may potentially be mined to generate valuable insights. The overall aim of the research is to efficiently predict hearing loss symptoms from pure-tone air and bone conduction audiometry thresholds which can be used to assist otorhinolaryngology clinicians in quick detection of hearing loss symptoms. The process involves finding relationship that exist between pure-tone audiometry thresholds and symptoms and other attributes in patient's medical audiology datasets and utilizing these relationships in the prediction. The symptoms can be accurately predicted with the aid of a model that employs hybrid machine learning algorithms that can predict a class or label of a given input data.

II. MOTIVATION

Previous studies relating to analysis of pure-tone audiometry data or audiology medical records of hearing loss patients have concentrate more on analysing this data to find how certain environmental or genetic factors exacerbate hearing loss [1]. Some have found with success similarities in audiogram configuration of patients with common disease, gender or age group [2]; [3]. Others have employed statistical and computational techniques in their analysis with good result [4] [5]. None of these studies utilized the relationship between audiometry thresholds and hearing loss symptoms to predict any of those attributes. There is the need for efficient computational algorithms that can help otorhinolaryngology clinicians make quick and better decisions at the point care.

In order to address this, efficient computational algorithm with good accuracy is needed. An algorithm that can predict or classify with minimum error rate will go a long way in helping

clinicians make informed decision that can improve quality of care of the patient.

III. REVIEW OF THE LITERATURE

The relationship between hearing thresholds and other attribute in the patients' medical data is evidenced by several studies. [6] Have discovered a relationship between Meniere disease and low frequency sensorineural hearing loss. [3] Indicates that noise induced hearing loss or Meniere disease is associated with particular audiogram configurations. [7] have found a connection between high-frequency sensorineural hearing loss and noise exposure, and the association of cardiovascular risk generated by smoking and diabetes with both high and low-frequency hearing loss.

Pure-tone air and bone conduction audiometry thresholds were obtained from unscreened population of older adults to observe the rate of changes over the period of 10 years. The baseline measurement predictors for that period are reported for specific age groups of both genders. The threshold changes over the said period for those at the age of 50 to 60 years old were found to be more at the higher frequencies while changes in threshold were at lower frequencies for the aged that are 80 years and above [8]. Age-related changes in pure-tone audiometry were also observed in a longitudinal study on hearing thresholds of 813 adult males. The results shows steady rate of hearing loss at higher frequencies and increase in the rate of hearing loss at lower frequencies [9].

A study on the effect of age and noise related hearing loss on high frequency thresholds was carried out on 187 industrial noise-exposed and 52 non-industrial noise-exposed subjects. The test-retest in the study shows high frequency audiometry (HFA) as a technique that is as reliable as the conventional in indicating noise-induced hearing loss and can be used more reliably to monitor hearing loss of individual cases over time period. Results from this study show both exposed and non-exposed subjects to have hearing loss at high frequencies (10-18 kHz) [10].

A number of studies have employed different machine learning and statistical methods in exploring and analysing audiometric data. [4] Have used a combination of statistical and neural techniques on audiology medical records with the aim of looking for factors influencing which patient will benefit from hearing aid. Audiogram shapes were clustered into homogeneous and heterogeneous groups using K-means clustering with the aim of helping clinicians in the diagnosis of hearing loss in the future [11]. [12] Experiment with multi-layer perceptron neural network and support vector machine in classifying ear disorders from otoneurological data. Varpa et al. (2008) Uses K-nearest neighbour and naïve Bayes classification algorithms to determine the classification accuracy of combining machine learnt knowledge with expert's knowledge. The results showed accurate classification when machine learnt knowledge is combined with the experts' knowledge.

[1] Use multivariate logistic regression for hearing loss risk factor analysis of workers exposed to toxic substance and noise in working environment. Similarly, [14] used multivariate logistic regression to estimate the risk of hypertension between three groups of workers tested with noise-induced hearing loss. Another study used expert program to study individual risk factors for sensorineural hearing loss for individuals with hereditary hearing loss, exposed to noisy environment and using analgesic [15]. Slope of linear regression was used to measure the rate of change in pure-tone thresholds of men and women in Korean population [16]. [17] Proposed a speech processing algorithm that is based on machine learning technique to address the problem of distortion in hearing aids. Three supervised machine learning algorithms used to classify audiograms into different phenotypes to get insight about possible presbycusis (related hearing loss) etiologies [18]. [19] Developed a neural network-based diagnostic tool to detect sensorineural hearing loss for vestibular schwannomas audiometry data. [20] Used supervised machine learning techniques to determine auditory brainstem response. [21] Employs Bayesian machine learning approach for feature selection in order to implement an online personalization of hearing aid.

None of the above mentioned studies have made any effort to find the extent of connection between hearing thresholds and hearing loss symptoms or any other attribute in the patient's medical record or use these relationships to predict with high accuracy any attributes in the hearing loss patient medical data. . [22] made an attempt by developing audioGene, a software that employs machine learning techniques to use phenotypic information derived from audiometry data to predict the genetic cause of hearing loss in persons segregating Autosomal dominant nonsyndromic hearing loss. The result of the study shows the software has accuracy of 68% in predicting the causative gene within its top three predictions. However, this result cannot be considered as impressive given the level of accuracy and confidence that is often required in the medical field.

The hybrid of unsupervised learning algorithm like association rule or FP-Growth algorithm with classification algorithm has been referred to as associative classification which has shown to perform excellently and even better than the existing classification algorithms [23][24]. [25]

Have achieved high prediction accuracy with the combination of K-nearest neighbour (KNN) and FP-Growth algorithm. [26] Have also experiment with combination FP-Growth and classification algorithms in classification of brain tumour with high level of accuracy. Bayesian classifier was chosen for the classification phase because it is a popular classification algorithm that performs well despite its simplicity [27]. This has led to its resurgence and application in numerous fields [28]. It has been proven to have advantageous properties and effective at classification than other approaches like support vector machines, boosting, genetic algorithms, decision tree and k-means algorithms [29].

IV. PROPOSED MODEL

FP-Growth (Frequent Pattern) algorithm is used as unsupervised learning algorithm and naïve Bayes classifier (Multivariate Bernoulli Model) as supervised learning algorithm to implement the model. This is because FP-Growth algorithm store the datasets in the FP-tree structure which results to faster execution time than many of the item sets generation algorithm commonly used, with two orders of magnitude better in performance than algorithm like Apriori [30]. And naïve Bayes classifier is a simple approach that works well and can outperform more sophisticated classifiers on many datasets [31].

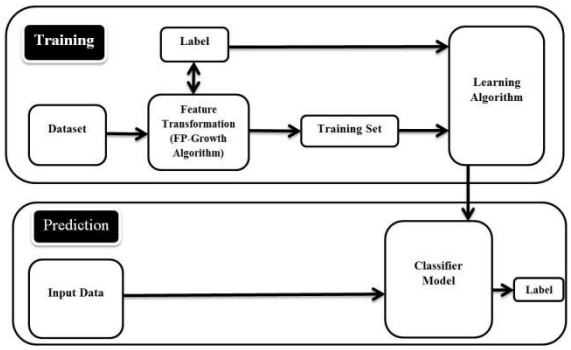

Fig. 1. Hearing loss Proposed Prediction Model

The prediction model in Figure 1 shows how FP-Growth algorithm is used as feature transformation technique to extract frequent itemsets from the dataset. These frequent itemsets represents relevant features that constitute the features in the vocabulary. Those training examples whose frequent itemsets (subset) passed the minimum threshold become part of the training set that is used to train the learning algorithm. In other words, out of the all training examples, only those particular ones with itemsets that passed minimum support threshold forms or becomes part of the training set for the Bayesian classifier. This approach reduces the dimensionality of the data, reduce storage requirement, removes redundancies and improve the algorithm performance. Hence, dimensionality of the overall training data for the classification is reduced when certain condition is required for a training example to be part of the classifiers training set. The classifiers vocabulary storage requirement shrinks when it is only composed of frequent features (itemsets) as opposed to the conventional way of including all the features from the training data. Noise and redundancies are elements that characterise many datasets. By selecting only frequent features in a data, a lot of redundancies will be eliminated. It is understandable that less data can boost processing speed thereby increasing algorithm performance. By using this technique, all the benefits of feature transformation such as feature construction, extraction and selection are all achieved.

Feature construction, feature selection and feature extraction are all form of feature transformation. It is a process in which new features are created [32]. Feature extraction is a process that extracts new features from the original features through some functional mapping [33]. Extracting frequent items or similar items in a dataset are two important techniques of feature extraction. While feature construction is process that uncovers missing information about relationships between features and replaces those missing information by inferring or creating additional features. However, FP-Growth algorithm utilized for this research serves as a linear and non-linear space embedding methods; a feature construction technique that reduce the dimensionality of the original data without losing the relevant information [32]. It is always easier to processes data in fewer dimensions. On the other hand, feature selection selects relevant features in a dataset in order to discard redundant and noisy features and thereby reducing storage requirements and improving algorithm performance [33].

Despite the fact that naïve Bayes classifier is considered to be stable classification algorithm with high bias and low variance [34], the proposed method of pre-processing with FP-Growth algorithm will not allow it to suffer from overfitting problem even as the training set grows because the FP-Growth algorithm provides the solution that is needed to overcome the issue of overfitting. That is, it serves as a model selection algorithm that automatically decides the sets of features to keep based on the minimum support threshold provided. This reduces the number of features were there are many available

V. RESEARCH METHODOLOGY

For this research, the combination of deductive and inductive approaches is adopted [35]. By

engaging with relevant body of literature and formulating a hypothesis that there is a relationship between hearing loss symptoms and pure-tone audiometry thresholds, deductive approach is adopted. Deductive approach is used when numerous literatures were found to show some level of connection between audiometry thresholds and hearing loss etiologies, patient's age and other attributes, and as a result of this, becoming interested in finding relationship between those hearing thresholds and hearing loss symptoms which also lead to uncovering the possibility these relationships might have in predicting hearing loss symptoms. While inductive approach is used when pure-tone audiometry data of 399 hearing loss patients were collected, analysed and the results used to construct generalizations, relations and literature supporting these results.

Although this research uses experimental research methodology which makes use of quantitative method of analysis, mixed method of research design which comprises of both qualitative and quantitative methods is the one adopted for this research [36]. This influences the choice of data collection methods that is adopted. It occurs in the data collection stage where qualitative data is obtained via the case study method by engaging the otolaryngology clinicians in discussion to give accounts of recent phenomenon within real-life context. Quantitative data is gathered in the form of secondary data. It is medical record for hearing loss patient including their audiometry data. The audiometric data sample of 399 patients across all ages (3-88 years) measured at 11 frequencies ranging from 0.125-8kHz from were collected over 10-year period from 2003 to 2012. The thresholds were the patients' first audiometry test results obtained from Hospital Pakar Sultanah Fatimah Muar, Johor, Malaysia file archive. Each patient's audiometric data includes the diagnosis information during the first visit. This is to avoid obtaining improved result after patient undergoes treatments. The data was collected using hand-held digital scanner that is used to scan the document. It is saved in portable document format (PDF) which is manually converted into digital format by entering every detail into a text document. The results are text data in form of diagnosis information and numeric data in the form of audiometry thresholds.

Hearing loss symptoms such as Tinnitus, Vertigo, Giddiness, Otalgia, Otorrhea, Rhinitis, Meniere and Presbycusis and other attributes in the medical records like age, gender of the patient were abbreviated. For example, the symptoms of tinnitus represented as (TNTS), vertigo as (VTG). Other symptoms like Meniere disease, Otalgia, Otorrhea, Giddiness, Prescubysis and Rhinitis were not abbreviated. The gender of male and female were represented as (F) for female and (M) for male. Ages of the patients were divided into three groups and represented as (E) for early, (M) for mid and (L) for late. For example, 55 years was represented as 5M (mid 50's), late 80's like 89 years was represented as 8L. Hearing thresholds were represented with colon (:) in between the frequencies and the sound dB. For example, 500:45R signifies threshold at frequency of 500Hz at 45dB for the right ear (R) and 8000:80L signifies high frequency of 8kHz at 80dB in the left ear (L). In their study, [4] have abbreviated hearing loss symptoms, structured data (gender, date of birth, hearing aid device type) and other free text data to use neural and statistical methods to cluster patients in order to find those that most benefit from being fitted with hearing aid. Before data is fed into the algorithm, there is the need to change the format to a format the algorithm can accept [30].

In mixed research methods, quantitative and qualitative research can be used in different sequence [35]. They can be used interdependently or concurrently. In other words, qualitative followed by quantitative, quantitative followed by qualitative or the two used independently and not in any particular order. There are benefits in mixing the two research design methods. One method can compensate for the weakness of the other like in a situation where the quantitative data deemed not enough; it can be supplemented with qualitative data.

VI. RESULTS AND CONTRIBUTIONS

The importance of feature extraction techniques cannot be over emphasized. It is important in the success of many machine learning algorithms [33]. The performances of the classifier using both multivariate Bernoulli and multinomial naïve Bayes models with and without feature transformation are compared. This section presents the validation results of the proposed technique performances on data set. The performance of the classifier is measured by calculating the error rate. This is done through cross-validation technique; repeated random sub-sampling validation method that partitioned the data set into training set and validation set. The partition is done multiple times and the training examples in each partition are randomly selected. There are 5 partitions and each containing 10, 20, 30, 40 and 50 training set respectively. Ten iterations are performed using each partition and the errors averaged [28]. It is often a good idea to use the algorithm on the data before pre-processing and compare the performance obtained with either representation [33]. [33] Further argued that it is better to make mistake on the side of being too inclusive rather than risk discarding relevant information.

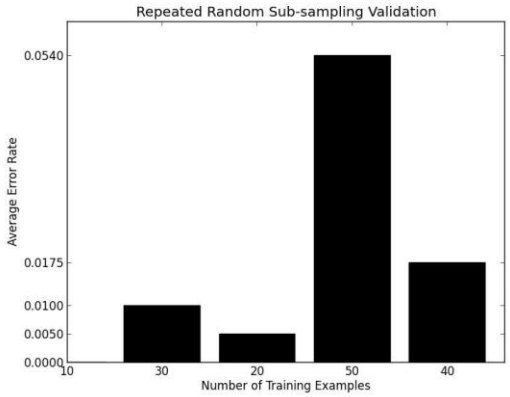

Fig. 2. Validation results using multivariate Bernoulli model with feature transformation

Figure 2 shows the average error rate from 10 iterations using 5 different partitions. There is 100% prediction accuracy using partition with 10 training examples, 99.5% accuracy with 20 training examples, 99% with 30 training examples, 98.25% with 40 training examples and 94.60% accuracy with 50 training examples. In other words, the average error rates for the 5 different partitions are, 0, 0.5, 1, 1.75 and 5.4% respectively. The classifier performs excellently with multivariate Bernoulli model with feature transformation.

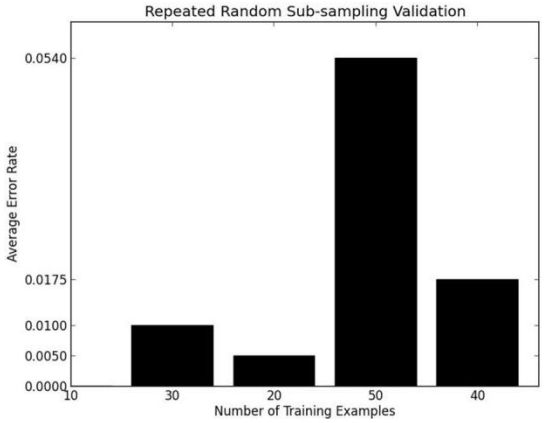

Fig. 3. Validation results using Multinomial model with feature transformation

Figure 3 visualizes the average error rate based on ten iterations using different partitions of validation set. This is the result from multinomial naïve Bayes model with feature transformation using FP-Growth algorithm. The partition with 50 training examples is the highest with 10% error rate averaged over 10 iterations. This means the prediction rate is 90% accurate. The lowest average error rate is 2% which is for partition with 10 training examples. This is followed by average error rate of 3%, 3.9% and 8.5% for 20, 30 and 40 partitions respectively.

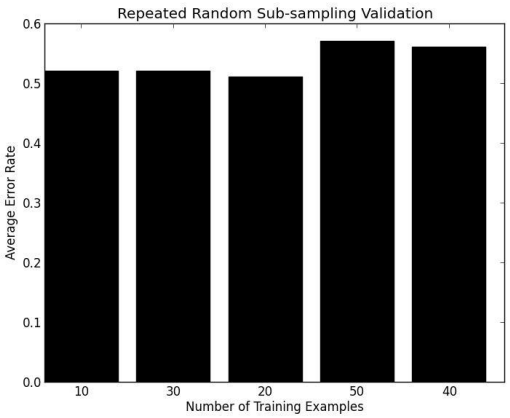

Fig. 4. Validation results using Multivariate Bernoulli model without feature transformation

Figure 4 depicts the validation results over 10 iterations using different partitions. The performance obtained from the classifier without feature transformation is contrary to the results obtained with feature transformation using both multivariate Bernoulli and multinomial models. The bar chart shows the average error rates using multivariate Bernoulli model without feature transformation. The average error rate for each partition is high. The partitions with 50 and 40 training examples have the worst average prediction inaccuracy with 57% and 56% average error rates respectively. The partitions with 10, 20 and 30 training examples have up to 50% error rate.

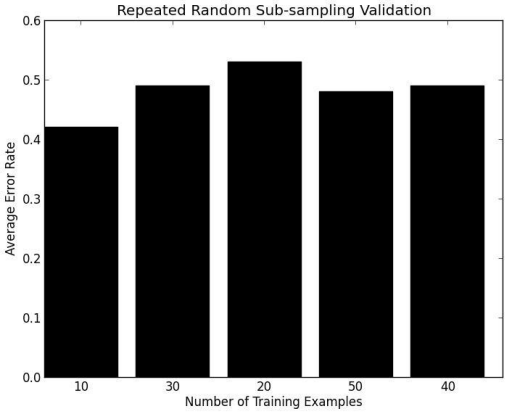

Fig. 5. Validation results using Multinomial model without feature transformation

Figure 5 shows the validation results using multinomial model without feature transformation. The partition with 10 training examples has the least average error rate of 42%. The highest is partition with 20 training examples having 53% average error rate. The partitions with 30, 40 and 50 training examples all have error rates of 48%. All the error rates were averaged over 10 iterations.

The approach used show that when the vocabulary in the learning is composed of this frequent itemsets, the accuracy of the prediction will be increased. Comparing the two Naïve Bayes classification models showed that the multivariate Bernoulli model prediction result is more accurate than the multinomial model with and without using FP-Growth algorithm as the feature transformation technique. Multivariate Bernoulli model with feature transformation was found to have minimal average prediction error rate of just 5.4% using 50 training examples in 10 iterations of repeated random sub-sampling validation as against multinomial model with feature transformation which also have an acceptable average error rate of 10% using 50 training examples in 10 iterations. The partition with 50 training examples is more preferable because testing or validating with more training examples in a random seems to show reliability and accuracy of the approach. Both multivariate Bernoulli and multinomial model without feature transformation gave the highest average error rates. In partition with 50 training examples, the multivariate Bernoulli model error rate after is 57% and that of multinomial model is 48%. This shows that the two models perform poorly on this size of dataset without feature transformation technique that pre-process the dataset. It is interesting to note that in all the 5 partitions in both models, the average error rates are without FP-Growth algorithm feature transformation.

Surprisingly, it was found that the average error rates in all the 5 partitions after 10 iterations are a bit higher in multivariate Bernoulli model without feature transformation compared to multinomial model without feature transformation. This has corroborated the findings of a previous work that indicates multinomial model outperforms multivariate Bernoulli model and supported his argument by using the two models on four different datasets (McCallum and Nigam, 1998). Eyheramendy et al. (2003) also found multinomial model to outperform multivariate Bernoulli model and three other probabilistic models on three text classification tasks. But this is not the same when the FP-Growth algorithm feature transformation is used. All the 5 partitions have lower average error rates in multivariate Bernoulli model with feature transformation compared to multinomial model with the same feature transformation.

However, in contrast to earlier findings on the performance of the two models, the result of this study do not support the argument by [38] that multinomial model of naïve Bayes have better classification accuracy than the multivariate Bernoulli model because of word frequency count as the reason why the former outperform the later. [38] Further argued that based on their experiment results, multivariate Bernoulli model performs better on smaller vocabularies and multinomial model performs better on larger vocabularies. But [27] disagree with this and argues that word frequency information is not the reason multinomial Naive Bayes classifier outperforms multivariate Bernoulli model. He argues that the removal of the word frequency information only increase performance of the algorithm rather than decrease it. Although, these studies argue in support of multinomial model Naïve Bayes, the result in this study has shown forming the vocabulary based on frequent itemsets that are subsets of each training example in the dataset can make multivariate Bernoulli model outperforms multinomial model.

There is large amount of data that pervade the healthcare industry. This data needs to be utilized using the proper techniques in order to realize the value and the knowledge that may be inherent in it. With the advancement in information technology and the wide adoption of health information system (HIS) healthcare providers can no longer be complacent as regard to embracing techniques that enables quality health services.

The healthcare worker goes through multi-spectral data and various information sources when diagnosing a disease in order decide on the appropriate treatment strategy. This research can help in the discovery of new and useful patterns in audiometry datasets of patients. The computational algorithm can be used to implement audiology decision support system (ADSS) that learns from past experiences and predict likely symptoms with high accuracy and minimal error rate. The clinician can use both his knowledge and the system to make better analysis of patient hearing test result and make more informed and better decision than either him or the ADSS could make.

The introduction of FP-Growth unsupervised machine learning algorithm as feature transformation techniques has open a new dimension in data pre-processing using one algorithm that presents the benefit of both feature construction, selection and extraction. This has also help enhance the classification accuracy of the multivariate Bernoulli naïve Bayes classification model which was otherwise believed to be less accurate than the multinomial model in previous studies [38] [39].

AKNOWLEDGEMENT

The authors wish to thank the Department of Otolaryngology, Hospital Sultanah Fatimah, Muar, Johor for their support and cooperation and for making the sample medical audiology records available to us.

REFERENCES

[1] S.-J. Chang, C.-J. Chen, C.-H. Lien, and F.-C. Sung, 'Hearing Loss in Workers Exposed to Toluene and Noise', Environ. Health Perspect., vol. 114, no. 8, pp. 1283–1286, Aug. 2006.

[2] E. Toppila, I. Pyykkö, and J. Starck, 'Age and noise-induced hearing loss', 12-Oct-2009. [Online]. Available: http://informahealthcare.com/doi/abs/10.1080/01050390152704751. [Accessed: 10-Aug-2013].

[3] S. Dalebout, The Praeger Guide to Hearing and Hearing Loss: Assessment, Treatment, and Prevention. ABC-CLIO, 2009.

[4] M. N. Anwar and M. P. Oakes, 'Data mining of audiology patient records: factors influencing the choice of hearing aid type', in Proceedings of the ACM fifth international workshop on Data and text mining in biomedical informatics, New York, NY, USA, 2011, pp. 11–18.

[5] C.-Y. Lee, J.-H. Hwang, S.-J. Hou, and T.-C. Liu, 'Using cluster analysis to classify audiogram shapes', Int. J. Audiol., vol. 49, no. 9, pp. 628–633, Sep. 2010.

[6] J. H. Hwang, H. C. Ho, C. J. Hsu, W. S. Yang, and T. C. Liu, 'Diagnostic value of combining bilateral electrocochleography results for unilateral Ménière's disease', Audiol. Neurootol., vol. 13, no. 6, pp. 365–369, 2008.

[7] Y. Agrawal, E. A. Platz, and J. K. Niparko, 'Risk factors for hearing loss in US adults: data from the National Health and Nutrition Examination Survey, 1999 to 2002', Otol. Neurotol. Off. Publ. Am. Otol. Soc. Am. Neurotol. Soc. Eur. Acad. Otol. Neurotol., vol. 30, no. 2, pp. 139–145, Feb. 2009.

[8] T. L. Wiley, R. Chappell, L. Carmichael, D. M. Nondahl, and K. J. Cruickshanks, 'Changes in hearing thresholds over 10 years in older adults', J. Am. Acad. Audiol., vol. 19, no. 4, pp. 281–292; quiz 371, Apr. 2008.

[9] L. J. Brant and J. L. Fozard, 'Age changes in pure-tone hearing thresholds in a longitudinal study of normal human aging', J. Acoust. Soc. Am., vol. 88, no. 2, pp. 813–820, Aug. 1990.

[10] H. O. Ahmed, J. H. Dennis, O. Badran, M. Ismail, S. G. Ballal, A. Ashoor, and D. Jerwood, 'High-frequency (10-18 kHz) hearing thresholds: reliability, and effects of age and occupational noise exposure', Occup. Med. Oxf. Engl., vol. 51, no. 4, pp. 245–258, Jun. 2001.

[11] C.-Y. Lee, J.-H. Hwang, S.-J. Hou, and T.-C. Liu, 'Using cluster analysis to classify audiogram shapes', Int. J. Audiol., vol. 49, no. 9, pp. 628–633, Sep. 2010.

[12] M. Moein, M. Davarpanah, M. A. Montazeri, and M. Ataei, 'Classifying ear disorders using support vector machines', in 2010 Second International Conference on Computational Intelligence and Natural Computing Proceedings (CINC), Sept., vol. 1, pp. 321–324.

[13] K. Varpa, K. Iltanen, and M. Juhola, 'Machine learning method for knowledge discovery experimented with otoneurological data', Comput. Methods Programs Biomed., vol. 91, no. 2, pp. 154–164, Aug. 2008.

[14] T.-Y. Chang, C.-S. Liu, K.-H. Huang, R.-Y. Chen, J.-S. Lai, and B.-Y. Bao, 'High-frequency hearing loss, occupational noise exposure and hypertension: a cross-sectional study in male workers', Environ. Heal. Glob. Access Sci. Source, vol. 10, p. 35, 2011.

[15] R. Kaksonen, I. Pyykkö, J. Kere, J. Starck, and E. Toppila, 'Hereditary hearing loss--the role of environmental factors', Acta Oto-Laryngol. Suppl., vol. 543, pp. 70–72, 2000.

[16] S. Kim, E. J. Lim, H. S. Kim, J. H. Park, S. S. Jarng, and S. H. Lee, 'Sex Differences in a Cross Sectional Study of Age-related Hearing Loss in Korean', Clin. Exp. Otorhinolaryngol., vol. 3, no. 1, pp. 27–31, Mar. 2010.

[17] J. Bondy, 'Applying Machine Learning to Speech Processing in Hearing Aids', Open Access Diss. Theses, Aug. 2005.

[18] J. R. Dubno, M. A. Eckert, F.-S. Lee, L. J. Matthews, and R. A. Schmiedt, 'Classifying Human Audiometric Phenotypes of Age-Related Hearing Loss from Animal Models', J. Assoc. Res. Otolaryngol. JARO, Jun. 2013.

[19] S. A. R. Nouraei, Q. J. M. Huys, P. Chatrath, J. Powles, and J. P. Harcourt, 'Screening patients with sensorineural hearing loss for vestibular schwannoma using a Bayesian classifier', Clin. Otolaryngol. Off. J. ENT-UK Off. J. Neth. Soc. Oto-Rhino-Laryngol. Cervico-Facial Surg., vol. 32, no. 4, pp. 248–254, Aug. 2007.

[20] P. McCullagh, H. Wang, H. Zheng, G. Lightbody, and G. McAllister, 'A comparison of supervised classification methods for auditory brainstem response determination', Stud. Health Technol. Inform., vol. 129, no. Pt 2, pp. 1289–1293, 2007.

[21] A. Ypma, J. Geurts, S. Özer, E. van der Werf, and B. de Vries, 'Online Personalization of Hearing Instruments', EURASIP J. Audio Speech Music Process., vol. 2008, no. 1, p. 183456, Jun. 2008.

[22] K. R. Taylor, A. P. Deluca, A. E. Shearer, M. S. Hildebrand, E. A. Black-Ziegelbein, V. N. Anand, C. M. Sloan, R. W. Eppsteiner, T. E. Scheetz, P. L. M. Huygen, R. J. H. Smith, T. A. Braun, and T. L. Casavant, 'AudioGene: predicting hearing loss genotypes from phenotypes to guide genetic screening', Hum. Mutat., vol. 34, no. 4, pp. 539–545, Apr. 2013.

[23] C. Rules, W. Li, J. Han, and J. Pei, CMAR: Accurate and Efficient Classification Based on Multiple Class-Association Rules. 2001.

[24] B. Liu, W. Hsu, and Y. Ma, 'Integrating Classification and Association Rule Mining', 1998, pp. 80–86.

[25] J. L. Mao, M. D. Li, and M. Liu, 'A Multi-Label Classification Using KNN and FP-Growth Techniques', Adv. Mater. Res., vol. 791–793, pp. 1554–1557, Sep. 2013.

[26] M. M. P Rajendran, 'Hybrid Medical Image Classification Using Association Rule Mining with Decision Tree Algorithm', 2010.

[27] K.-M. Schneider, 'On Word Frequency Information and Negative Evidence in Naive Bayes Text Classification', in ESPAÑA FOR NATURAL LANGUAGE PROCESSING, ESTAL, 2004.

[28] 'IBM Information Center masthead', 09-May-2011. [Online]. Available:

http://pic.dhe.ibm.com/infocenter/spssstat/v20r0m0/topic/com.ibm.ibmcom.doc/banner.htm. [Accessed: 10-Aug-2013].
[29] A. Ghosh, R. K. De, and S. K. Pal, Eds., Pattern Recognition and Machine Intelligence: Second International Conference, PReMI 2007, Kolkata, India, December 18-22, 2007, Proceedings, 2007th ed. Springer, 2008.
[30] P. Harrington, Machine Learning in Action. Manning Publications, 2012.
[31] I. H. Witten, E. Frank, and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, Third Edition, 3rd ed. Morgan Kaufmann, 2011.
[32] H. Liu, Feature selection for knowledge discovery and data mining. Boston: Kluwer Academic Publishers, 1998.
[33] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, Eds., Feature Extraction: Foundations and Applications, 2006th ed. Springer, 2006.
[34] Mathematical Reviews. American Mathematical Society, 2004.
[35] D. E. Gray, Doing Research in the Real World, Second Edition. SAGE Publications Ltd, 2009.
[36] J. W. Creswell and V. L. Plano Clark, Designing and conducting mixed methods research. Los Angeles: SAGE Publications, 2011.
[37] O. Z. Maimon and L. Rokach, The Data Mining and Knowledge Discovery Handbook. Springer, 2005.
[38] A. McCallum and K. Nigam, A comparison of event models for Naive Bayes text classification. 1998.
[39] S. Eyheramendy, D. D. Lewis, and D. Madigan, On the Naive Bayes Model for Text Categorization. 2003.

The Impact of EMR User Interface Design on Doctor Satisfaction

Shatilah Zakaria
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
me870409@gmail.com

Mohd Khanapi Abd Ghani
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
khanapi@utem.edu.my

***Abstract*— Many complex systems such as electronic medical records (EMR) system fail to provide user with a good user interface. As a result, patient medical records are mostly on paper and are scattered and often inaccessible by doctor. Aim to identify and design the most usable EMR user interface for capturing medical records and to explore user satisfaction of quality with EMR. To achieve this, review and comparison of the existing EMR user interface application will be conducted. Then, follow by a primary data collection on EMR interface and doctor satisfaction. Finally, to design, develop and validate the interface design.**

***Index Terms*— Electronic Medical Record; user interface design; principle of user interface design; User Satisfaction, healthcare**

I. INTRODUCTION

User interface design or user interface engineering is the design of computers, mobile communication devices, software applications, websites with the focus on the user's experience and interaction. The goal of user interface design is to improve users experience with simplicity and efficiency while keeping the design ergonomically sound. The term user friendly is often used as a synonym for usability which de- notes the ease with which people can employ a particular tool or other.[1]

User interface design is also a subset of a field study called human-computer interaction (HCI). Human –computer interaction is the study, planning, and design of how people and computer work together so that a person's needs are satisfied in the most effective way. The user interface is part of a computer and its software that people can see, hear, touch, talk or otherwise understand or direct. The user interface has essentially two components: input and output. The best interface is one that is not noticed, one that permits the user to focus on the information and task at hand, not mechanism used to present the information and perform the task.[2], [3]

Electronic record-keeping can dramatically change the work of a physician. With a few taps on a keyboard, a doctor can call up all of the office notes on a patient, as well as reports from specialists, test results, family history, and lists of past treatment and medication. With another keystroke or two, a prescription can be sent to a pharmacy.

The shift of media from paper to screen has some consequences that have to be considered. We need to understand these in order to avoid making things that were easy with the paper-tools more difficult in the computerized media but also to be able to exploit the possibilities offered by the new media for enhancement of human cognitive skills such as using dynamical pictorial representations, color, symbols and other visual codes[2]

Medical concept representation is not only a matter of words and terminology. There is also a need for development of representations that provide overviews in some specific frequently encountered decision-making situations. Such representations should be rich in information content, elaborated and optimized for fast information extraction.

EMRs offer several key advantages over paper medical records. First, the quality of care is enhanced due to the fact that information is always on hand. Second, diagnosis and treatment speed increases and becomes more efficient therefore increasing revenue. Third, an EMR data trial is much easier to retrieve than a paper trial and facilitates medical auditing. However, there has been resistance to converting to EMRs mainly due to economic reasons [4].

Poor user interface design has been pointed out as a factor contributing directly to the problems of computerized medical records [5][6]and became the main reason the doctors reluctant to use computer in recording clinical finding. Examples of design problems found were inconsistency, poor screen design, insufficient navigation control and difficulties to access important procedures.[6], [7]

The term user friendly is often used as a synonym for usability which de- notes the ease with which people can employ a particular tool or other human- made object in order to achieve a particular goal[1]. Hence, it is important to develop user interface design especially for EMR system and a good system only would not help the doctor.

This research suggests that a compilation of user requirement and existing system's user interface design to produce a validate user interface design for EMR system. As such, this research aims to follow research question:

A. *What is the crucial design factor that could make user interface usable to be used by the doctor?*

This question is related to the user interface design factors, user interface principles and what are the correct design that could be used to satisfied doctor because the medical business flow are different from common business or user interface, it could affect the doctor-patient relationship. Many principles of design an interface should be count. From research in usability engineering we know that there are five essential usability characteristics that are vital for any Human–Computer Interaction (HCI) and these characteristics are not easy to achieve, especially when one has to deal with as heterogeneous a user group as in healthcare. [8]

II. Aim of Investigation

The aim of the research is to produce a validate user interface design for EMR system. The new design interface of EMR should provide the best user interface that could help the doctors feel ease in using the EMR system without feeling of guilt and threaten.

III. Problem Statement

Poor user interface design has been pointed out as a factor contributing directly to the problems of computerized medical records[9] and became the main reason the doctors reluctant to use computer in recording clinical finding. Examples of design problems found were inconsistency, poor screen design, insufficient navigation control and difficulties to access important procedures.

IV. Research Objectives

1. To review and compare and contrast the existing EMR user interface model of application.
2. Conduct primary data collection on EMR user interface and doctor satisfaction using interview and questionnaire and case study.
3. Design and develop EMR user interface usability model EMRuism.
4. Validate and test EMRusim through system prototype.

V. Related Work

There are works has been done to enhance the user interface for the existing medical system. For example, Else Nygren [2] mention that the interface should be easy to learn; short learning time, efficiency in daily work and accommodate excellence. In a study of how physician worked with the paper-based information tools, physicians were not reading the text pages from top to bottom but shifted rapidly between reading, skimming and skipping parts of the contents because reading of continuous text is generally slower from screen than from paper. It also mentions that type of font does not have significant effect on reading speed. The shift of media from paper to computer screen means that limitation in presenting data no longer exists. However, some user interface issues will require special attention such as, disposition of the screen area, naming screen-objects, design of presentation format for initial display of medical record, design of dedicated overview formats, design of an efficient way of presenting multiple pages of textual descriptions and design of data entry methods.

Furthermore, there are work by Larkin [10] that a common reason for being unaware of certain EMR features was lack of software training and some doctor mention that the idea of EMR is something that interferes with personal nature of doctor-patient exam. In order to improve the design, we must enhance the personal connections such as using free text fields; non medical term patient notes and memory cues can assist in building and sustain personal connection. Other than that, using seamless input of contextual data such as, the ability to quickly note supplemental information and reminders is critical to physician workflow.

Another work from A.F.Rose et al. [11], one of the greatest design in identified challenges of EMR user interface design was balancing the clinician's information needs with the limited of available screen, too many information on the screen and screen contrast also an important on consideration for user interface designer. When designing a system for user group, it is important to know about diversity of expectation, experience and prior knowledge. In either care, careful attention to usable design in medical applications can impact the user's perception of the application's utility and potentially patient safety and quality care.

Other work from Y.Batu Salman [12], many end user of medical input systems encounter usability problems such as learnability, flexibility and robustness with the interfaces of computer applications. By employing icons,it provides an easy and friendly interaction method with systems. Iconic interfaces serve to reduce system complexity and so the mental workload of end users when the icons are designed properly.

Using the theories of human intuition in designing the user interface proposed by IIlie [13] also can help the design of the interface for not being too theoretical and more to human need. As it is difficult to design a system that fit everyone's needs. Many complex electronic medical records (EMR) fail to provide users with an intuitive interface because the systems were not initially designed for environment where there is an opportunity-cost time pressure such as healthcare.

VI. Our Research

A. *Research Methodology*

This section outlines the activities conducted in the research. The aim of this research is to produce a validate user interface design for EMR system. The new design interface of EMR should provide the best user interface that could help the doctors feel at ease in using the EMR system without feeling of guilt and threaten.

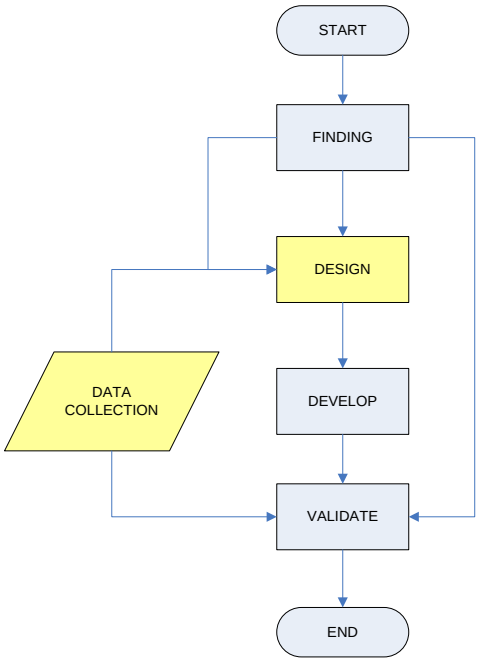

Figure 1. Research Approach

1. *Provide literature review*

The first objective in this research is to know about the current knowledge and existing EMR user interface design in order to achieve the primary aim of the research. Due to this, the literature dealing with all design that related EMR should be in count such as reviewed, compared and contrasted with other user interface design of an EMR model. By doing this, the best design will be as an output.

2. *Design and develop*

The research represents the gap between the existing user interface and the research aim. This third objective is about identifying the means to close this gap. This objective is to design a suitable user interface that achieves research aims.

3. *Primary Data Collection*

The second objective is to conduct primary data collection on EMR user interface and doctor satisfaction. The data collection could help to determine the best user interface that existing for doctor's use. The data will be collected through questionnaire and target user to answer the data collection will be divided into three part; Senior Doctor, New Doctor and Medical Students. During the first preliminaries survey, users were divided into three groups of users whilst for Validate phase, questionnaire will be given to the medical doctor that involve in validating the interface.

4. *Validating*

After the proposed user interface is established, it is important to validate and identify the main gaps. The identified gaps would help in order to solve the refinement and improvement of the user interface. During the validation of the interface based on user requirement, another survey will be conducted in order to verify whether users were satisfied or not. After the validating, user interface will be refine based on the satisfaction survey.

VII. Research Results And Progress

As shown in Figure 1, Data Collection phase will be as external input in order to start the Design phase other than literature Finding.

After obtaining the best theoretical guidelines for the user interface design in Finding phase, the outcomes will be merged with the existing data collection in order to meet user requirement and satisfactions.

As for the ongoing data collection, currently we have 32 collected data on user requirement for user interface design from Medical Students and Doctor in Medical.

VIII. Research Implication

The research is expected to guide EMR designer and system implementers in their efforts to design usable interfaces. Second, results from this research can resolve important issues in healthcare today: what are the user requirements from medical field toward electronic medical record system (EMR)?

IX. Acknowledgement

We would like to acknowledge the support of Universiti Teknikal Malaysia Melaka (UTeM) and to thank all friends and colleagues for their helpful comments and courage.

X. References

[1] B. Qureshi, "Applied Computing and Informatics," in in *Applied Computing and Informatics*, vol. 9, no. 1, King Saud University, 2011, pp. 79–82.

[2] E. Nygren, "From Paper to Computer Screen . Human Information Processing and Interfaces to Patient Data . The user-interface challenge Experiences of

computerized medical," vol. 222, no. 1, pp. 317–326, Sep. 1997.

[3] A. Sutcliffe, S. Fickas, and L. A. Ehlhardt, McKay Moore Sohlberg, "Investigating the usability of assistive user interfaces," *Interact. Comput.*, vol. 15, no. 4, pp. 577–602, Aug. 2003.

[4] W. Chen and M. Akay, "Developing EMRs in developing countries.," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 1, pp. 62–5, Jan. 2011.

[5] B. Shneiderman and S. Ben, *Designing The User Interface: Strategies for Effective Human-Computer Interaction, 4/e (New Edition)*. Pearson Education India, 2003.

[6] K. Zheng, R. Padman, and M. P. Johnson, "User interface optimization for an electronic medical record system.," *Stud. Health Technol. Inform.*, vol. 129, no. Pt 2, pp. 1058–62, Jan. 2007.

[7] K. Zheng, R. Padman, M. P. Johnson, and H. S. Diamond, "An interface-driven analysis of user interactions with an electronic health records system.," *J. Am. Med. Inform. Assoc.*, vol. 16, no. 2, pp. 228–37, Jan. 2007.

[8] D. Herzberg, N. Marsden, P. Kübler, C. Leonhardt, S. Thomanek, H. Jung, and A. Becker, "Specifying computer-based counseling systems in health care: a new approach to user-interface and interaction design.," *J. Biomed. Inform.*, vol. 42, no. 2, pp. 347–55, Apr. 2009.

[9] S. Tsumoto, "Web based medical decision support system for neurological diseases," in *Proceedings IEEE/WIC International Conference on Web Intelligence (WI 2003)*, 2003, pp. 629–632.

[10] K. Larkin and A. Kelliher, "Designing flexible EMR systems for recording and summarizing doctor-patient interactions," *Proc. 2011 Annu. Conf. Ext. Abstr. Hum. factors Comput. Syst. - CHI EA '11*, p. 1609, 2011.

[11] A. F. Rose, J. L. Schnipper, E. R. Park, E. G. Poon, Q. Li, and B. Middleton, "Using qualitative studies to improve the usability of an EMR.," *J. Biomed. Inform.*, vol. 38, no. 1, pp. 51–60, Feb. 2005.

[12] Y. B. Salman, H.-I. Cheng, and P. E. Patterson, "Icon and user interface design for emergency medical information systems: a case study.," *Int. J. Med. Inform.*, vol. 81, no. 1, pp. 29–35, Jan. 2012.

[13] V. Ilie, O. Turel, and P. D. Witman, "Towards a New Design Paradigm for Complex Electronic Medical Record Systems: Intuitive User Interfaces," in *2013 46th Hawaii International Conference on System Sciences*, 2013, pp. 2585–2594.

Migration Traditional Projects Framework to Enterprise Resource Planning System: An Empirical Study

Basem Zughoul
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka (UTeM)
Melaka, Malaysia
bzughoul@hotmail.com

Burairah Hussin
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka (UTeM)
Melaka, Malaysia,
burairah@utem.edu.my

***Abstract*—Traditional software projects are designed and developed in house to fit a specific or small set of functions. Typically the software developments required time, high cost to develop, not integrated, unclear customer requirements, and it's limited to specific business functions. However, from business perspective information systems must need versatile and comply with various business needs, like implementation the on-shelf system. Many companies rely on Enterprise Resource Planning (ERP) systems to manage and process their business needs; in spite of its expensive cost, high risk, large, and complex. However, ERP systems suffer fails due to the minimum customization, or not fit with business Standard Operating Procedures (SOP) for the organization. The aim of this framework is to seek a solution for software companies, to develop traditional projects according to ERP characteristics to be more adaptive, with environment change based on business process concepts. The proposed framework is developed based on integration among three structures for software development models: Waterfall, Iterative and XP model, and adaptive with unique features of ERP. The framework was compared and evaluated as an empirical study applied in Jordanian software companies. The results highlight on problems faced these companies, which lead to failed projects. In addition, phases and activities that did not find in companies model, also the phases and activities exist in the company's model and not found in the proposed framework.**

***Index Terms*—Enterprise Resource Planning, Information system, Iterative implementation, System development methodology, Traditional software project**

I. INTRODUCTION

Traditional software project is known as custom built software, in house developed software, or mainframe based on organization requirement. Generally with this setting, the company retains full ownership rights of the source code and security of data. Thus, the organization customized the software according to the required and needs. Traditional project's emphasis on the specification of requirement analysis and then everything evolve around it, until delivery of business functions. The execution of this type of software according to business requirements needs a high cost, time and efforts for the organization to achieve competitive goals [1]. Finally there are important problems that face these systems: the compatibility with developments of software and hardware, integration and redundancy of data between modules.

Enterprise Resource Planning (ERP) system solutions are currently in high demand by both manufacturing and service organizations. Therefore they provide a tightly integrated solution to an organization's information system needs [2]. ERP systems have been known as systems that bring integration to several business activities within complex organizations. Also ERP systems are becoming a standard information system, irrespective of the size and nature of the organization. Thus, they classify ERP implementation into technological and operational factors [3]:

- Technological factor: replacement of disparate systems (hardware/software) because of difficulty of maintenance, integration and the systems do not meet business needs.
- Operational factors: customer satisfaction, improvement of process, simplification of complex process and standardization of process.

Many companies regard and rely on ERP systems to manage and process their business data, although the high cost associated with enterprise software [4]. [5] Presented four reasons that urge organizations to implement ERP system: (1) simplify and standardize systems, (2) replace legacy systems, (3) gain strategic advantage, (4) improve interaction between suppliers and customers. ERP market was $16.67 billion in 2005 and forecasted to be over $21 billion in 2010 [6]. Over 35 hundred million U.S. dollar has been invested in the deployment of ERP, and over 60,000 companies have already implemented ERP system in the whole world [7].

The ERP system has the ability to automate and integrate the business processes in the organization, to share common data and practices across the organization, and to produce and access information in real time [8]. Generally, ERP systems are becoming one of today's most widespread Information Technology (IT) solutions. Thus, organizations move from functional to process-based IT infrastructure [9].

However, ERP suffers a major defect in its implementation, because of the complexity of the system or the cost involve during installation [10].

The development of information systems is still high risk and complex, because it contained several applications, large size and development teams, high cost. Therefore, it usually required substantial investments of time, cost and effort. Where, it combines the implementation of new information system with a redesign of work process [11], [12].

The implementation of these systems is difficult and high cost that places tremendous demands on corporate time and resources, and associated massive change in organizational structure, business process, and people work. So many ERP implementations have been classified as failures, because they did not achieve organizational goals [11], [13]. Therefore, ERP systems are more than just new information systems. In fact, they are software packages, which is acquired as an application software package or developed in house or by independent software vendors, to meet the general requirements and needs to organize [14].

For these reasons, organizations must upgrade and develop their traditional projects, to replace old standalone systems with a unified system, which divided into software modules linked together with integration similar to ERP systems. The business environment is significantly changing. Companies today face the challenge of expanding markets, increasing competition, business requirements and rising customer expectations. This increases the pressure to reduce total costs with less time. There are some reasons to migrate to ERP systems as follows [15]:

- Most traditional projects rely on developing a software system due to user needs. But ERP projects consist of tightly linked interdependencies of software systems, business processes, and process reengineering.
- Traditional software packages implementation emphasis on technical analysis and programming, while ERP projects emphasis on business process design, software configuration and legacy data maintenance.

Many companies are changing their IT strategies to maintain a competitive advantage. Therefore, become more responsive to changing markets, and deliver better service at lower cost by purchasing off-the-shelf integrated ERP software, instead of developing systems in-house [16]

In addition, there are a lot of reasons for buying ERP system: ERP system uses a common database shared by all departments, moving from client-server architecture to internet architecture, and allow the organization to move agilely and adapt quickly to changes in the business environment [17], [18].

II. Software Development Methodologies

Software development methodology or model is a set of standard steps using by most organizations to develop and support their information systems, such as many processes, so the development of these systems often follows a life cycle [19]. System life cycle is one of the key concepts in information system development in many organizations.

Many different models and techniques used to direct the software project based on the life cycle concept, such as waterfall model, iterative model and XP model [19].

Waterfall Model

The waterfall is the traditional and classical model used for software development, sometimes called the classic life cycle, linear sequential model. It is the first software process model since 1970, the main phases of this model are initial investigation, requirements definition, system and software design, development and testing, implementation and operation & support) [20], [21].

Iterative Development model

This model is solving problems of waterfall model and which could provide faster results. Where, it requires less in advance information and offer greater flexibility. The main idea of this model is based on divided project into small parts, like a mini-waterfall to allow the development team to show results earlier, and get valuable feedback from system users [20].

Extreme Programming (XP)

Extreme Programming (XP) is an agile software development model. It handles rapid changes in requirements and environments in short increments and release cycles. Also XP can be described as a development method for projects, where not all requirements and features can be defined and planned before the development start [12], [19].

III. Main Characteristics of ERP System

Many studies show different characteristics of ERP. Thus, for simplifying the characteristics of ERP we consider the most commonly referred from the literature review. Therefore, in this work the main features of ERP project were identified and help to understand what they are, what they can do, how they differ from other IT packages [22]. There are the main characteristics that affect the ERP life cycle as follows:

- Complexity: ERP system is the most complex software in the most information systems. It usually includes one application, which can be operated separately of other organizational applications with friendly user interface [10], [16], [13], [22].
- Flexibility and Adaptability: organizations perform customization and configuration ERP system in order to achieve objectives and reach to competitive advantages [16], [15], [3].
- Software Packages: it is one of the basic characteristics that distinguishes ERP system from custom built software, because it is developed to meet general needs of a group of organizations [19], [22].
- Integration: this property represents an extremely important part and provides a tightly integrated solution to the ERP system. Also it's one of the core functions that differ from previous IT project [10], [22].
- Best Practices: ERP system implementation is the best business practices to serve a large variety of organizations. It adopted many options to fit and support various business processes used in different type of organization [10].
- Completeness: ERP is much broader scope than that early software packages. Also it designed to affect the entire enterprise structure rather than set of function for traditional software packages [16], [22].
- Openness: ERP system is evolving over several years. Most organizations need to implement an ERP system to replace or upgrade the existing infrastructure. It was

gradual and continuous developments without considerable changes [16], [15].

- Change Process: ERP project can be viewed as an organizational change due to the large number of changes. Therefore, it brings to an organization structure like workflow processes, business process and the way people do their job [8], [15], [17].

Finally we evaluate these characteristics according to software development models (waterfall, iterative and XP model) as shown below in Table I.

TABLE I
EVALUATION OF SOFTWARE DEVELOPMENT MODELS AGAINST ERP CHARACTERSTICS

ERP System Characteristics	System Development Models		
	WATERFALL	ITERATIVE	XP
Flexibility		√	√
Integration	√		
Complexity	√		
Software Packages			√
Completeness	√		
Change Process			
Best practices			√
Openness			

IV. PROPOSED DEVELOPMENT FRAMEWORK

This section presents and describes a proposed framework for ERP system development life cycle. Consequently, the definition of this framework will take into account main ERP characteristics, requirements and the aspects of ERP systems. This framework will apply for migrating traditional software projects to the ERP system, based on the combination of the main characteristics and structure of the three models: Waterfall, Iterative and XP model according to ERP needs, figure 1 shows the structure of proposed framework.

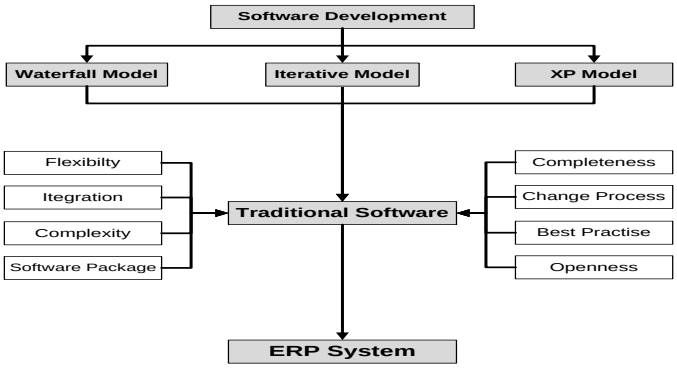

Fig. 1. Structure of proposed framework

Proposed framework has many phases that proceeded by a critical look at an organization as enterprise architecture, involving several business process components surrounded by change management. The enterprise architecture analyses the business, information technology, and motivating for implementing ERP systems. Change management seeks to address problems early of implementation, and integrate the human resource dimension with the business process design prospectively. Figure 2 shows the proposed framework phases.

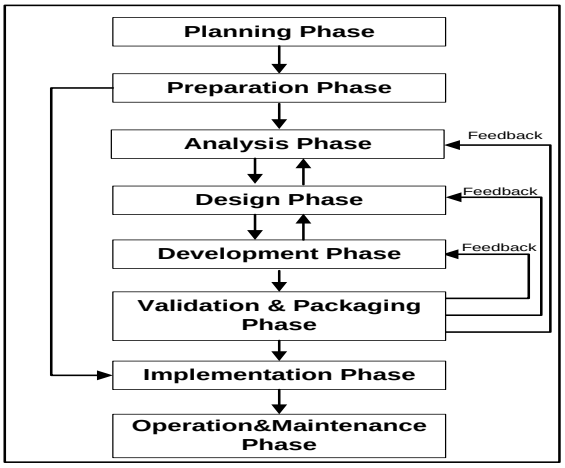

Fig. 2. Proposed framework phases

The following is a breakdown of the propose framework of each phase:

1. Planning phase: The major activities of this phase are:
   - Define the project objective
   - Define project scope
   - Define the project package
   - Create a project schedule
   - Establish implementation team and trained
   - Collect information about the technological infrastructure
   - Prepare risk plan
   - Prepare test cases
2. Preparation phase: The steps of this phase in this model are:
   - Organization structure
   - People structure
   - Technology infrastructure
3. Analysis phase: The major process of this phase emphasizes on:
   - Analyze organization requirements(RFP)
   - Requirements specification
   - Partition requirements
   - Analyze process sequence
   - Validation activity(business process and priority)
4. Design phase: The activities that involved in this phase are:
   - Design system architecture
   - Partition system into subsystem
   - Define subsystem
   - Build prototype
   - Verify system design
5. Development phase: The major steps of this phase:
   - Build system database
   - Code modules
   - Build interfaces
   - Testing (code, integration, and all systems)

6. Validation and Packaging phase: The main objective of this is to perform a validation for the whole system and readiness for packaging.
7. Implementation phase: Major processes to avoid and minimize ERP implementation risk and complexities:
   - Training end users and management
   - Iteratively implementation
   - User acceptance tests
   - Data migration/conversion
   - Run pilot system
   - System go-live
8. Operation and Maintenance phase: Major activities that perform during this phase:
   - Establish support centers
   - System upgrading
   - System maintenance to continuity

V. EMPIRICAL STUDY

Two case studies were conducted on software companies in Jordan. In order to verify and compare the proposed framework with models that used by these companies. A first case study was conducted on the large business company, second case study were conducted on medium business company. These companies operate in software that deals with ERP systems and traditional projects. The evaluation and comparison of this study did through interviews and questionnaires. The target respondent in each company was project manager (PM), director of MIS and IT manager, or any senior person responsible and involved in developing information systems. The case study taken up to collect data due to; the company process model for each company that used to build their software projects. Also the problems that faced company that leads to failure the project [11], [12].

VI. RESULTS AND DISCUSSION

Response data were collected and analyzed. Hence data were analyzed and compared with proposed framework. A descriptive difference carried out in order to, understand the effectiveness of phases and activities that obtained from the study. Some of phases and activities that concerns in proposed framework and not exist in the company's model, or in one of them, also some activities found in another phase.

Difficulties and Problems Faced ERP Development

According to project manager, there are some difficulties and problems that faced implementation projects. So sometimes lead into project failure such as: project management, change management, end user awareness, user qualifications, communication and the organization's readiness.

Comparison Between Proposed and Company Model

Consequently, the proposed framework compared with company's software model according to major phases. There are main observations that represent the difference between two models according to the analysis and comparison:

1. Preparation phase is not part of the both companies model, whereas the system in both companies designed to fit specific function or small set of functions. Also the implementation affects only part of the organization. But in ERP project, there are massive changes in organizational structure, business process, and people. Therefore preparation phase takes place at the beginning of the project, to address problems of implementing an ERP system early enough, to complete all other pre request items.

2. The validation and packaging phase which is very important for ERP package before the implementation phase. So software packages are one of ERP characteristics that means the system will be tested and validated to produce software package, as Commercial Off-The-Shelf (COTS) before starting ERP system implementation.

3. The implementation phase is the basic principle of any ERP system. Therefore, implementation is the correlation between the system and the organization. Establish implementation team and training them in planning phase to keep up on what progress during system development. Also Iterative implementation reduces the implementation risk and the complexity of the ERP system.

4. Business Process Reengineering (BPR) is an integral part of the framework. BPR is the redesign of business processes for substantial improvement. In other words, it redesigns the way of work, so the definition of process is any series or combination of tasks or activities, which produce a result.

5. Some problems that impact system implementation are the readiness of environment. Therefore framework combines the activities that lead to change in the technological infrastructure by the beginning of project, and give organization enough time.

Table II shows the comparison between proposed framework and company A and B.

TABLE II
COMPARISON BETWEEN PROPOSED FRAMEWORK AND COMPANY A&B

Proposed Framework	Company A	Company B
Planning Phase		
1. Definition project objective	√	√
2. Define project scope	√	X
3. Defining a project package	√	√
4. Create a project schedule	√	√
5. Establish implementation team and	x	X
6. Define technological infrastructure	√	X
7. Preparing a risk plan	√	X
8. Preparing test cases	x	X
Preparation Phase		
1. Organization structure	x	X
2. People structure	x	X
3. Technology infrastructure	x	X
Analysis Phase		
1. Analysis organization requirements	√	√
2. Requirements specification	√	√
3. Partition requirements	x	X
4. Analyze process sequence	√	X
5. Validate business process and priority	x	X
Design Phase		
1. Design system architecture	√	√
2. Partition system into subsystem	x	x
3. Define each subsystem	x	x

4. Build prototype	√	√
5. Verifying system design	√	x
Development Phase		
1. Build system database	√	√
2. Coding system modules	√	√
3. Build interfaces	√	√
4. System testing	*	x
Validation & packaging phase		
1. Validate system and readiness for	x	x
Implementation Phase		
1. Training end users and management	√	x
2. Iterative implementation	x	x
3. User acceptance tests	√	x
4. Data migration	√	x
5. Run pilot system	x	x
6. System go-live	√	√
Operation & Maintenance Phase		
1. Establish support centers	x	x
2. System upgrading	x	x
3. System maintenance	x	x

√ This activity is performed in this phase, x This activity is not performed in this phase, * This activity is performed in another phase

VII. Conclusion

Most of project managers and senior teams work strongly agreed to the proposed framework, as a software development life cycle. Therefore, they are looking to apply to their projects. This framework is theoretical based, rather than practical works in real life project. To practice and apply this framework to build ERP or migrate from traditional project to the ERP system in some Jordanians software companies. So the results and lessons of applying this framework will be evaluated.

The proposed framework is characterized by looking for organization, as enterprise architecture involving several business processes and surrounded by change management. Furthermore, addresses problems of ERP implementation (people, change organizational structure and technology problems), in early enough by introducing preparation phase in the ERP life cycle. The proposed framework combines the activities that attend to change in the technological infrastructure of the organization, also using prototyping and business process reengineering of the ERP implementation scope and system complexity of ERP, the implementation phase is performed iteratively to simplify the process.

Acknowledgment

I would like to acknowledge and thank the Universiti Teknikal Malaysia Melaka (UTeM) for this support and help toward my research and studies.

References

[1] Rizvi, S. A., & Ahmad, S. "Legacy Code Components Based Estimation Modeling."International Review on Computers and Software(IRECOS).(2008) 3. n. 2, 145-147.

[2] Shehab, E. M., M. W. Sharp, and T. A. Spedding. "Enterprise resource planning An integrative review." Business Process Management Journal 10, no. 4 (2004): pp. 359-386.

[3] Al-Mashari, M., Al-Mudimigh, A. and Zairi, Z. "Enterprise resource planning: A taxonomy of critical factors" European Journal of Operational Research 146, no. 2 (2003): 352-364.

[4] Borovskiy, Vadym, and Alexender Zeier. "Enabling Enterprise Composite Application on Top of ERP Systems." IEEE Services Computing Conference. 2009. 492-497.

[5] Sumner, Mary. Enterprise Resource Planning. Prentice Hall, 2005.

[6] Leon, Alexis. "ERP Demystified second edition-study guide." In ERP Demystified(Second Edition), by Alexis Leon. Singapore: McGraw-Hill, 2008.

[7] Chen, C.-C., Shih, D.-H., Chiang, H.-S., & Lai, C.-Y. The "Affecting on the Performance of Introducing Different ERP Strategies." International Review on Computers and Software(IRECOS). (2011). 6. n.5, 765-773.

[8] Nah, Fiona Fui-hoon, Janet Lee-shang Lau, and Jinghua Kuang. "Critical factors for successful implementation of enterprise systems." Business Process Management Journal 7, no. 3 (2001): 285-296.

[9] Motwani, J., R. Subramanian, and P. Gopalakrishna. "Critical factors for successful ERP implementation: Exploratory findings from four case studies." Computers in Industry 56, no. 6 (2005): 529-544.

[10] Negi, T., and Veena Bansal. "A Metric for ERP Complexity." Business Information Systems, 2008.

[11] Umble, Elisabet J., Ronald R. Haft, and M. Michael Umble. "Enterprise resource planning: Implementation procedures and critical success factors." European Journal of Operational Research 146, no. 2 (2003): 241-257.

[12] Cordeiro, Lucas, and Carlos Mar. "An Agile Development Methodology Applied to Embedded Control Software under Stringent Hardware Constraints." Edited by Lucas Cordeiro, Carlos Mar, E. Valentin and F. Cruz. ACM SIGSOFT Software Engineering Notes 33, no. 1 (2008): 1.

[13] Appuswamy, R. "Implementation iss in ERP." International Conference on System Thinking in Managment. 2000. 1-8.

[14] O'Leary, Daniel E. "Discussion of Information System Assurance for Enterprise Resource Planning Systems: Unique Risk Consideration." Journal of Information Systems 16 (2002): 115-126.

[15] Iskanius, Päivi. "Risk Management in ERP Project in the Context of SMEs." Iskanius, Päivi 17, no. 4 (2009).

[16] Ahituv, Niv, Seev Neumann, and Moshe Sviran. "A System Development Methodology for ERP Systems." Computer Information Systems, 2002: 56-67.

[17] Yusuf, Yahaya. "Enterprise information systems project implementation: A case study of ERP in Rolls-Royce." International Journal of Production Economics 87, no. 3 (2004): 251-266.

[18] Kapp, K., F. William, and N. Hester. Integrated learning for ERP success: a learning requirements planning approach. New York: Amazon, 2001.

[19] Hoffer, A. Jeffery. Modern System Analysis and Design (Sixth Edition ed). Upper Saddle River, New Jersey 07458: Pearson Education, Inc, 2011.

[20] Qureshi, M, and S Hussain. "A reusable software component-based development process model." Advances in Engineering Software 39, no. 2 (2008): 88-94.

[21] Sommerville, I. Software Engineering (8th Edition ed.). Addison Wesely , 2006.

[22] Brehm, L. "Tailoring ERP Systems : A Spectrum of Choices and their Implications." Proceedings of the 34th Hawai'i International Conference on System Sciences. 2001.

Software Quality

Software Quality Assessment Model for Global Software Development (SQAM-GSD) **Sadia Rehman and Siffat Ullah Khan**
Web Usability Management Model (WEB-UMM) From Customers' Perspective **Muhammad Salam and Siffat Ullah Khan**
Classification Techniques in Blood Donors Sector - A Survey **Nor Syuhada Che Khalid, Burhanuddin M.A., Asmala Ahmad and Ghani M.K.A**
Improving Data Accessibility Using QR Code in Healthcare Domain **Fatin N. M. Leza, Nurul A. Emran and Mohd. Khanapi A. Ghani**
Improving Large Database Performance Using SQL Query Rewriting in MES **Nor'azah Md. Khusairi, Nurul A. Emran and Mohd Mokhtar Mohd Yusof**
Test Case Generator for Embedded Real Time Software **Nabilah Kamarul Baharin and Radziah Mohamad**
A Study on Code Peer Review Process Monitoring using Statistical Process Control **Alhassan Muhammad Abubakar and Dayang N. A. Jawawi**

Software Quality Assessment Model for Global Software Development

(SQAM-GSD)

Sadia Rehman
Software Engineering Research Group (SERG_UOM),
Department of Computer Science and IT,
University of Malakand, Khyber Pakhtunkhwa, Pakistan.
srehman71@gmail.com

Siffat Ullah Khan
Software Engineering Research Group (SERG_UOM),
Department of Computer Science and IT,
Department of Software Engineering,
University of Malakand, Khyber Pakhtunkhwa, Pakistan
siffatullah@uom.edu.pk

***Abstract*--Global Software Development (GSD) is a modern software engineering paradigm adopted by many client organisations in developed countries to get high quality products at low cost in low wage countries. Fabrication of high quality software is considered as one of the key factor in GSD. However GSD projects face many challenges related to software quality. In order to address these challenges software quality models are frequently used in organisations to fabricate high quality products. The objective of research project is to develop a Software Quality Assessment Model for Global Software Development (SQAM-GSD) to assist vendor organisations to assess the quality of their software products so that they can improve their quality for GSD. Software quality attributes and software quality metrics will be identified through Systematic Literature Review (SLR) and validated by questionnaire survey. Correlation analysis and SWOT analysis of software metrics will be performed on quality attributes and quality metrics respectively. The results of Correlation analysis and SWOT analysis will be combined to form Software Quality Assessment Model. Model will also have real world practices to address identified quality attributes. The anticipated outcomes of this review will be Software Quality Assessment Model for Global Software Development to evaluate the quality of software products.**

***Keywords* - Software Quality Metrics, Global Software Development, Software Quality Assessment Model**

I. INTRODUCTION

Global Software Development (GSD) is a modern software engineering paradigm adopted by many client organisations in developed countries to get high quality product at low cost in low wage countries. Production of high quality software is considered as one of the key factor in the rapid growth of GSD. However GSD projects face many challenges related to software quality. In order to address these challenges software quality models are frequently used in organisations to fabricate high quality products [1].

Global Software Development (GSD) has become the new approach for software development from the last two decades. Vendor organisations develop their software products at low cost in low-wage countries [2]. Organisations compete in global market and maintain their products according to the needs of software industry [3]. Providing high quality products is challenging for vendor organisations in GSD. GSD put more challenges to vendors rather than standalone software development such as timely delivery of products, skilled human resources, sharing of knowledge, contract management, hidden costs and Quality of product and service [2].

Change in the Requirements and desires of customers are natural phenomenon [4]. Organisations working in GSD have to provide high quality products to compete in Global industry to cope with the demands of customers. Software quality became more critical for GSD, that is why organisations spend much of their time and capital on improving quality of software products [3]. In order to cope with change and maintain quality of the products, software quality is monitored through-out the software developmental cycle [5]. According to IEEE "Software quality is the degree to which a system, component, or process meets specified requirements and meets customer's needs or expectations" [6]. Quality of software often expresses in its attributes which can be further subdivided into other attributes [7]. Quality attributes help us to measure the entity and evaluate the characteristics of a particular entity [8].We measure things to understand it and make our lives comfortable [9]. "Measurement is the process by which numbers or symbols are assigned to attributes of entities in the real world so as to describe such entities according to clearly defined rules" [10].

GSD induces the organisations to invest more cost and efforts to get high software quality by selecting appropriate Software Quality Metrics according to their requirements at the right time. The use of software metrics becomes ineffective

and extra burden on organisations when metrics do not observe the goals of organisations [11].

The use of software quality metrics can aid to minimize the software crises in software industry [12]. Identification of quality attributes plays a vital role to improve software quality. Vendor organisations use metrics to improve its quality by measuring its capabilities and efficiencies. Software metrics in GSD plays a vital role to avert on hand and future risks and to improve its software quality [6].

The purpose of software measurement is to quantify all attributes of quality and predict the future quality of the software [13]. The use of appropriate software metrics at right time helps the organisations to achieve their required and expected outcomes. The use of Software Quality Metrics facilitates the organisations to get both short and long term advantages by introducing high quality products to the global market [12].

Software measurement makes it easy to maintain the quality of software products. Humans are familiar with the concept of measuring in their everyday life. Measuring quality of software makes the product reliable to users [10]. Identification of quality attributes (QAs) plays a vital role to improve software quality. It is also necessary to show the impact of one QA on another. In other words the relationships among QAs are important to identify critical QAs that the organisations have to focus on and measure. This helps the organisations to improve, control and predicate the quality of their products [14]. In GSD the identification of QAs and Software Quality Metrics to measure these QAs is quite challenging. The main idea is to choose appropriate metrics at right time that help the organisations to develop high quality products [6].

This research will help the organisations to choose appropriate Software Quality Metrics and QAs for GSD at right time to develop high quality products [15]. The outcomes of this research will be SQAM-GSD that will give different practices for QAs, correlation Analysis of different QAs and SWOT analysis of currently used Software Quality Metrics in software industry to measure these QAs. In order to make work more reliable we are using systematic literature review approach and an empirical study. After performing SLR an empirical study will be conducted to validate the findings of SLR.

II. RELATED WORK

Production of GSD products are more challenging as vendor organisations have to compete in the global market. In order to fabricate quality software products software quality metrics are used to measure the quality of products. Organisations have to identify quality attributes that are important to measure and control.

Software metrics in GSD plays a very important role to minimize risks and control quality of software products [16]. ISO standards of quality are being adapted by organisations in GSD to excel their performance. ISO/ICE 9126 quality model have various internal and external quality factors [17].

Quality attributes define criteria for the selection of appropriate metrics. Organisations should establish quality manual that contains metrics, measurement tool, and measurement guidelines and data analysis method. The organisations should select Software Quality Metrics according to their requirements at the right time [6].

"Measurement, metrics, and statistical analysis of data are the basic tools of science and engineering. Unfortunately, the software industry has existed for more than 50 years with metrics that have never been formally validated, and with statistical techniques that are at best questionable" [18].

Software Quality Metrics for GSD aid to control quality as according to Tom DeMacrio "You cannot control what you cannot measure. Measurement is the prerequisite to management control" [14]. Barbara Kitchenham has conducted a survey to describe advancement in software metrics research. The study assessed 103 papers published between 2000 and 2005. She has suggested that researchers in software metrics domain need to refine their empirical methodology to answer empirical questions [19].

Dr. Deepshikha Jamwal discussed different quality models (McCall's Quality Model, Boehm's Quality Model, Dromey's Quality Model, FURPS Quality Model, ISO 9126 Quality Model**)** and concluded that only "reliability" is the common attribute in all models. A criteria has been defined based on some questions in order to choose quality model for any organisation that will save organisation's time [20].

Software quality models aid vendors to select which quality attributes are essential for their organisations. Software quality model should be made reusable for other projects. Organisations must define their product quality [21]. Software quality is measured and validated by different model; ISO/IEC 9126 is international standard for software quality widely used in industry [22]. Luijten at el, conducted an empirical study that showed the correlation between measured software quality by Software Improvement quality model and the speed of solving these issues by teams[23].

Capability Maturity Model Integration (CMMI) also encourages the use of software metrics in software development. Organisations involve in GSD try their best to achieve CMMI high level to survive in global market [24].

Linda L. Westfall has presented seven steps process to design an affective software metrics.

According to the author well designed software metrics facilitate the organisations to improve their software products, processes and services[25].The following figure represents the concept of software metrics in an organisation.

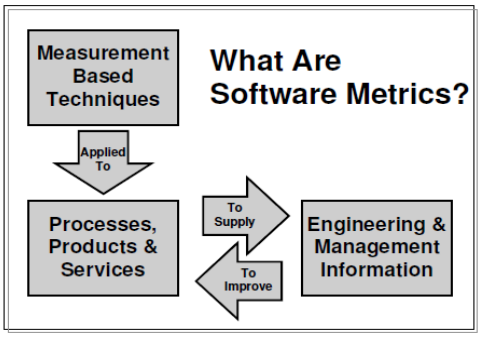

Fig. 1. Software Metrics

A survey is conducted by Rawat et.al, to show the impact of software Metrics on software quality, according to which software metrics become the foundation of software development, the use of software metrics will increase software quality and productivity in future [4].

In literature many researchers worked on software quality and software measurement but SLR has not been conducted yet. There is a need of model that assesses the quality of software products and identifies critical Quality Attributes and Software Quality Metrics for GSD. The model should also suggest real world practices to address critical quality Attributes in GSD. Still there is a need to show the relationships among different Quality Attributes so that vendors keep an eye on critical attributes.

This research will identify critical Quality Attributes and Software Quality Metrics for GSD. SWOT Analysis of quality metrics will be performed to help the organisations working in GSD to select appropriate metrics for their organisations. This research will help the organisations to assess the quality of their software products.

III. AIMS AND OBJECTIVES

The main objective of this research project is to develop Software Quality Assessment Model for Global Software Development. This model will help GSD vendor organisations to evaluate the quality of their software products. Correlation analysis and SWOT analysis will be performed on quality attributes and quality metrics respectively. These findings will be combined to form Software Quality Assessment Model with real world practices to address quality attributes**.** Correlation Analysis of Quality Attributes will show the relationships among different Quality Attributes. Similarly SQAM will propose Software Quality Metrics based on the SWOT analysis for the efficient evaluation of the identified critical Quality Attributes. This research will help researchers and practitioners working in Global Software Development.

IV. RESEARCH QUESTIONS

The following six research questions (RQs) were formulated to develop SQM-GSD.

RQ1. What are the different quality attributes, as identified in the literature, which affects software quality in GSD?
RQ2. What are the existing software quality metrics in GSD, as identified in the literature?
RQ3. What is SWOT analysis of existing Software Quality Metrics for GSD?
RQ4: What is correlation analysis of different Quality Attributes for GSD?
RQ5. What are the different quality attributes, as identified in the real world practice, which affect software quality in GSD?
RQ6. What are different software quality metrics, as identified in the real world practice, which affect software quality in GSD?

V. METHODOLOGY

Empirical research methodology will be used for this research project in three phases. In phase one data will be collected and analyzed, in second phase collected data will be validated. In the final phase empirical method will be used to evaluate SQAM-GSD. The detail of each phase is given below.

A. Data collection and Analysis

In first phase data will be collected via Systematic Literature Review (SLR). A systematic literature review is a way of discovering, assessing and inferring all available research relevant to a particular research question, or topic area [26]. Systematic Literature Review (SLR) will be conducted for the identification of software quality attributes and existing software quality metrics. The rationale behind the use of SLR is that it is unbiased, thorough and fair. Formally SLR has three phases, planning, conducting and reporting the review. SLR has been developed validated and published for this research project [15].

Second phase consists of an empirical study in order to validate the findings of SLR by conducting an empirical study i.e. Questionnaire survey and interviews. Similarly we will identify real world practices for the identified QAs and Software Quality Metrics. These practices will help us in the implementation of these QAs and Software Quality Metrics. After identification and validation of findings we will perform correlation analysis for QAs to identify the relationships among different QAs, and SWOT analysis of Software Quality Metrics .The result of correlation analysis and SWOT Analysis will be used with practices to

develop our model. SQAM will assist vendor organisations to address QAs at the right time by choosing right metrics to produce high quality software products.

B. SQAM-GSD Development

The following phases illustrate the fabrication of Software Quality Assessment Model for GSD. These phases have been adapted from stages of SOVRM [2].

Phase 1: Planning
a) 1.1: Set criteria for project
b) 1.2: Formulate Research Questions

Phase 2: Data Collection and Analysis
a) 2.1: Identification of QAs and Software Quality Metrics
b) 2.2: Correlation analysis of QAs
c) 2.3: SWOT Analysis of Software Quality Metrics

Phase 3: Development and Evaluation of SQAM-GSD
a) 3.1: Systemization of Result
b) 3.2: SQAM-GSD Development
c) 3.3: SQAM-GSD Evolution

The following figure shows the different phases of SQAM-GSD.

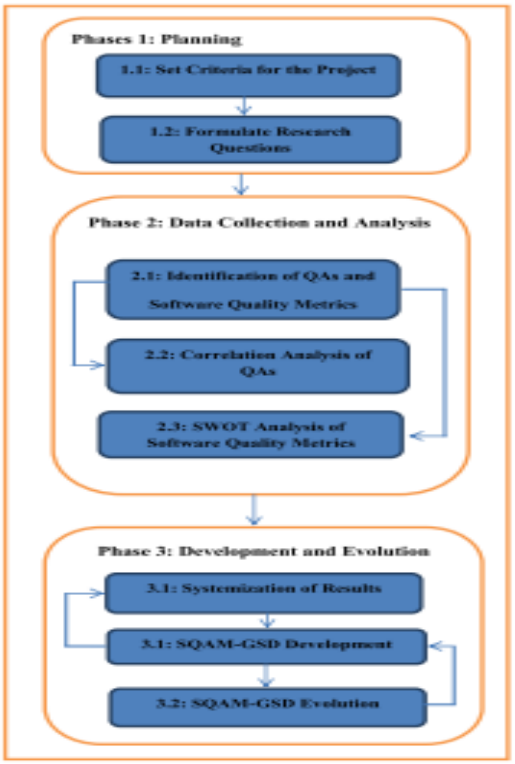

Fig. 2. Phases of SQAM-GS

C. Evaluation of SQAM-GSD

The SQAM-GSD will be evaluated by conducting five case studies at different GSD organisations. The logic of using case studies is that research revealed that case study is most widely used empirical method [22] .Based on the results of the case studies any changes proposed by the practitioners to the identified data will be added in order to make the model robust. After each case study conducted a focus group session will be arranged to in order to obtain feedback from practitioners. The following figure shows the structure of SQAM.

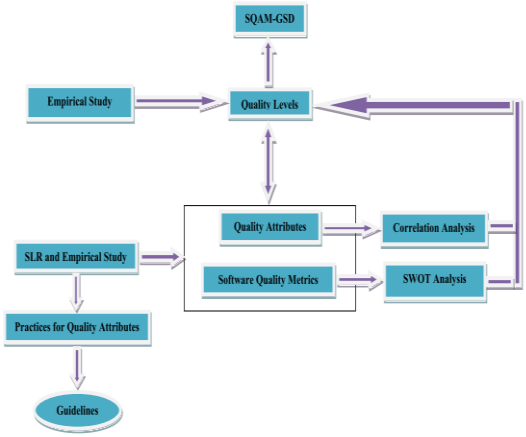

Fig. 3. Structure of SQAM-GSD

VI. TIMELINESS AND NOVELTY

This research work will help the organisations to assess the quality of their work at right time to fabricate high quality product in GSD. Quality levels of this model will assess the quality status of an organisation's software products in GSD. In the literature many researchers have worked out on different quality models and quality attributes but still there is a need for a model that will give critical QAs for GSD and will show how to measure these QAs and how to assess quality level of software products. There is also a need of a model that presents the real world practices to address these QAs.

SQAM-GSD will be the first model that will point out QAs along with practices and Software Quality Metrics to address these QAs. Correlation analysis of QAs will show the relationships among different QAs to make the picture clear; to focus on what actions should be perform at right time. This model will not only help the organisations to evaluate quality of their products and improve their work but also help the researchers and end-users to better understand the quality level of a software products and organisations. SQAM-GSD will assess what is the quality level of software products of an organization and how to improve it.

VII. RESEARCH PROGRESS TO DATE

Software Quality and its measurement is the most important for vendor organisations. In order to cope with the quality demands of customers, organisations use different metrics program. Providing high quality products in GSD develop trust among customers and vendors. Our model will assess the quality of software products and provide guidelines to achieve high quality products. So far we have done the following steps in our proposed model.

- Problem identification
- Research Questions identification
- Research Methodology selection
- Evaluation Method selection
- SLR Protocol developed and published
- SLR Protocol implementation in progress

ACKNOWLEDGEMENTS

We are obliged to Software Engineering Research Group at University of Malakand (SERG_UOM) in general and to Muhammad Ilyas Azeem in particular for the review and their valuable comments in validation process of the protocol**.**

REFERENCES

[1] IEEE, "IEEE Std 610.12-1990-IEEE Standards Glossary of Software Engineering Terminology, IEEE Computer Society," IEEE Std 610.12-1990, 1991.

[2] S. U. Khan, "Software Outsourcing Vendors' Readiness Model (SOVRM) PhD Thesis," Keele University UK, 2010.

[3] B. N. S. Sahay, and S. Krishna, "Global IT outsourcing," University of Cambridge, 2003.

[4] I. Sommerville, Software Engineering, 9th ed: Pearson Education, Inc, 2009.

[5] T. Andersson, "A Survey on Software Quality Metrics," 1999.

[6] M. Xenos, "Software Metrics and Measuremnts," Encylopedia of E-Commerce,E-Government and Mobile Commerce, pp. 8, 2006.

[7] S. R. Pressman, Software Engineering: A practitioner's Approach, Sixth ed, 2005.

[8] S. E. S. o. t. I. C. Society, "IEEE Standard for a Software Quality Metrics Methodology," presented at IEEE Std 1061™-1998 (R2009), 1999.

[9] E. F. Norman and S. L. Pfleeger, Software Metrices :A Rigorous and Pratical approach: PWS publishing Company, 1997.

[10] N. E. Fenton and S. L. Pfleeger, Software Metrics: a Rigorous and Practical Approach, 2nd Edition ed, 1997.

[11] H. Douglas, "The Darker Side of Metrics," 2000.

[12] F.Norman and Schneidewind, "Knowledge Requirements for Software Quality Measurement," Empricial Software Engineering, 2001.

[13] J. Magne, "Software quality measurement," University of Oslo, Department of Informatics, Norway.

[14] T. DeMacri, "Controlling Software Projects," 1982.

[15] S. Rehman and S. U. Khan, "SWOT Analysis Of Software Quality Metrics For Global Software Development: A Systematic Literature Review Protocol," IOSR Journal of Computer Engineering (IOSRJCE), vol. 2, pp. PP 01-07, 2012.

[16] Sanjay Misra and L. Fernández-Sanz, "Quality Issues in Global Software Development," presented at The Sixth International Conference on Software Engineering Advances, 2011.

[17] ISO, "ISO 8402," 1994.

[18] C. Jones, "Strenght and Weakness of Software Metrics," Software Productivity Research LLC, 2006.

[19] B. Kitchenham, "What's up with software metrics? – A preliminary mapping study," The Journal of Systems and Software, pp. 37–51, 2009.

[20] S. Malathi and S. Dr S, "Analysis of Size Metrics and Effort Performance Criterion in Software Cost Estimation," Indian Journal of Computer Science and Engineering (IJCSE), vol. 3, 2012.

[21] V. R. Basili and G. Caldiera, "Improve software quality by reusing knowledge and experience," Sloan Management Review, vol. 37, 2012.

[22] J. P. Correia, Y. Kanellopoulos, and J. Visser, "A survey-based study of the mapping of system properties to ISO/IEC 9126 maintainability characteristics," presented at Software Maintenance, 2009. ICSM 2009. IEEE International Conference on, 2009.

[23] D. Bijlsma, M. A. Ferreira, B. Luijten, and J. Visser, "Faster issue resolution with higher technical quality of software," Software quality journal, vol. 20, pp. 265-285, 2012.

[24] Z. Dave, "Measuring Software Product Quality: the ISO 25000 Series and CMMI," European SEPG, 2004.

[25] L. L. Westfall, "Twelve Step to Useful Software Metrics," presented at Proceedings of the Seventeenth Annual Pacific Northwest Software Quality Conference, 2005.

[26] B. Kitchenham, "Guidelines for performing Systematic Literature Reviews in Software Engineering," Software Engineering Group, School of Computer Science and Mathematics and Department of Computer Science University of Durham, ST5 5BG,Durham,, EBSE Technical Report 9 July 2007.

Web Usability Management Model

(WEB-UMM) From Customers' Perspective

Muhammad Salam
Department of Computer Science and I.T
University of Malakand, Pakistan
msmdir@gamil.com

Dr.Siffat Ullah Khan
Department of Computer Science and I.T
University of Malakand, Pakistan
Siffat7@yahoo.com

Abstract*-- The effective and efficient completion of a task by web users is called web usability. The web site is said to be usable or user friendly if the end user uses the website in an easy manner without any special training. Website having good usability has positive impact on web user. Web usability plays an important role in the business promotion and success. The first objective of this study is to identify the CSFs (Critical Success Factors) and CBs (Critical Barriers) that have a positive and negative impact on the web users in the context of web usability. The second objective is to develop a Web Usability Management model (WEB-UMM) based on basis of identified CSFs and CBS. The proposed model will help web development organizations measuring the level of the website in terms of web usability. Systematic literature review (SLR) and empirical software engineering methods will be used for the identification of CSFs and CBs. We will use empirical software engineering methods to validate the findings of SLR protocol. In order to represent the proposed model mathematically; the concept of fuzzy logic will be used. The anticipated outcome of this research will be a Web Usability Management model (WEB-UMM).

Index Terms—Web Usability, website usability, website usability model, WEB-UMM.

I. INTRODUCTION

Usability is a measure of how quickly or easily users of a system are able to achieve their desired goal [1].The definition of usability by Nielsen [1993]. "Usability has multiple components and is traditionally associated with these five usability attributes: learnability, efficiency, memorability, errors, and satisfaction". *Learnability*: how easy is it for users to accomplish basic tasks the first time they visit the webpage. *Efficiency:* how quickly users can perform tasks after getting used to the design and organization structure of the website. *Memorability:* how easily users can re-establish their proficiency after a period of time not using the website. *Errors:* How many errors do users make, how severe are these errors, and how easily they can recover from the errors. *Satisfaction:* How pleasant it is to use the design [2].Web usability is the efficient, effective and satisfying completion of a specified task by any given Web user [3]. In general, Web usability was defined as the ability of people who use the product to accomplish their own tasks in a quick and easy way. However, previous studies show that there was a lack of concern in Web usability when designing website, which caused website failures [4]. The concept of usability can be defined as how well and how easily a user, without formal training, can interact with an information system of a website [5]. Since the web is used for almost every aspect of our daily life so it is important to concentrate on the website usability in order to make useful for user's perspective. Web is the most interactive medium in the context of information retrieval and searching. A large number of people use web for different purposes like information retrieval, online games, online chatting, on-line shopping, on-line education, etc. keeping in view these facts it is important to focus on the website usability. A user who was interested to buy online product deviates from the website because of error in the user interface, this lead to financial loss. Another situation in today's cyber world is the growing diversity of devices used to access the websites, in early days traditional PCs was used as a standard device for web browsing, so the web developers design the website keeping in mind the main properties of the PCs like screen resolution, etc. Nowadays web browsing facilities are also available in other small devices like mobile phones. Keeping in view these facts it is extremely important for web developers to give proper attention to the usability of the website [6].The web site is said to be usable or user friendly if the end user use the website in easy manner without any special training. Website having good usability has positive impact on web user. Web usability plays an important role in the business promotion and success [5].As we know that internet and web is the easiest way of getting information and services, as a result more companies and organizations go for online solutions. As we know that online stores and online consumers are increasing day by day. Users will not stick with the company's website if they do not have a pleasant usability. Web usability play an important role in attracting customers to the website. People pretends to dislike the website if it is difficult to use (have low usability).people may give up visiting the website if the contents are hard to read or the interface of the key functions is difficult to figure out.

Web user normally deviates from the website whose contents is poorly managed, and visits other competitive user-friendly website on the cyber space. The website can achieve its goals very easy if it is user friendly (having good usability) [5].Web usability is one of the key characteristic that decides the success of any type website. Visitors of the website will deviate from the website before finding the information they were seeking if the usability of the website is poor. Usability is mostly related to the web interface design, if the web interface is user friendly then it is said to be good one. Web usability is not only permitting net users to

navigate easily and expediently but also help them to find what they want. It is revealing from the current literature that many companies have lost lot of money and potential surfers because of web usability problems. According Rowland's study, a clothing merchant in UK suffers huge losses due to low number of site visitors, despite spending millions on advertisement [7].Another example is the study by zona Research: A survey of 239 longtime internet users while shopping online were conducted. In this survey 28 percent user finds online shopping very difficult. 62% of the users has given up looking for a specific product online [1]. If the usability of a website is low then web surfers will leave the site before finding the information they were looking for [8]. Usability is a primary motivating factor in any development [9]. When designing the website it is extremely important to consider usability of the website. Gradually the need to measure the usability of software becomes more significant. Poor web usability will obviously reduce user productivity [10].One of the key success factors of business website is good usability. The website should be design in such a manner that that enables users to accomplish their tasks more easily. Many commercial website have failed during the last couple of years, because of missing customer retention and repeat sales. Poor usability is one of the major causes of their failure [11].

Keeping in view all of the above facts related to web usability the website should be designed to make easy and support well-organized and defective human-computer interactions. Every possible attempt should be made to reduce the consumer's work load in the context of web usability. Consumers will make the best use of the website when the website is usable.

Web usability management model WEB-UMM will help web development organizations to improve the usability of a web site which will lead to the success of the website in term of usability. If the proposed model is successfully adopted in the website development industry, website users may accomplish their own tasks in a quick and easy way.

II. BACKGROUND AND RELATED STUDIES

Some of the problems and issues of web usability have already been addressed by scholars and researchers for example following are some of these studies related to web usability. The main purpose of this study is to develop a consistent, small and suitable professed usability measurement tool. The objective is to build up an economical tool that can be used to measure website usability. Thus the measurement tool could be used for assessment purposes within the organizations. For example, an organization could gauge customer's observation of its website usability and of their competitor's website in order to scale their website with the antagonism. The development of a usability dimension tool that shows proofs of consistency and build soundness would also be helpful to researchers in order to examine the relationship between apparent usability and other applicable constructs such as thoughts toward the website and intention revisit to the website [5].According to this study for software quality models usability in an imperative feature. In the development of quality and successful software applications it an important factor. In the software engineering field usability is the most broadly used idea which is responsible for software demand and use. Keeping in view such an extensive significance of this quality factor various usability assessment methods are projected by usability experts and researchers. This article presents a detail study of various usability assessment methods. The article will be advantageous for the both the students and the researchers who are working in the field of software engineering and usability [12]. According to this study/ thesis web applications success and failure depends upon the web usability and user experience. There are many challenges measuring usability during SDLC (Software development life cycle). On the basis of systematic literature review, this thesis explains the current usability and user experience assessment and dimension methods and the definite procedures as well as their applicability during the SDLC. The challenges of those methods also identified. For more elaboration of the challenges an informal interviews with Software Company were conducted. On the basis of these findings a usability and user experience dimension and assessment road map for the web applications development companies were define. The roadmap contains a set of usability assessment and dimension methods as well as measures that is suitable to be used during the web development lifecycle. In order to authenticate the applicability of the definite roadmap, a case study was conducted on a real time market oriented real domain web application. During this study a number of aspects of web usability are identified that require attention. Studies on the topic of customer perspective on usability and user knowledge also seem to be vital from the usability assessment perspective. As this would help software developer humanizing their perceptive regarding what the customer expects and accomplish their requirements. The study mainly focuses on previous web measurement studies, especially web quality, web metric, models and methods. Several research gaps have been identified in the review. Even though our conclusion may be indicative of the field, additional reviews are required to verify the outcome obtained. Future effort includes the extension of this study by including other sources [13].According to [14] it is essential for the companies to carry out usability test in order to conclude whether their products are user friendly or not. For the end user companies it equally important to perform usability studies as well. In order to run usability testing more efficiently and effectively this study presents the development of Usability Management System (USEMATE), and present system as an alternative solution to help usability tester or practitioners. The core objective of USEMATE is to improve the present systems which are paper based, concerning manual score computation using excel and manual response time recording into a computer based management system (web-based management system) [14].

Presently there is an increasing demand for quality software. However the majority of the software systems cannot meet the users' requirements because of its low usability. As usability has been considered a significant quality parameter of the software system, in this regard to build an excellent software system it is important to have an applied definition of usability. Since software engineers and

researchers have not build yet any model that accurately explain usability definition and all its parameters that take into account the varying aspects of usability. The objective of this study is to high light the various parameters defined that are important regarding usability. The paper is useful for software systems' designers, developers and users [15].In this study systematic literature review (SLR) methodology is used. The study mainly focuses on previous web measurement studies, especially web quality, web metric, models and methods. Several research gaps have been identified in the review. Even though our conclusion may be indicative of the field, additional reviews are required to verify the outcome obtained. Future effort includes the extension of this study by including other sources [9].According to this research paper when a customer shopping online or searching for some information on the web, the customer may find what they are searching in several attempts. These unsuccessful attempts (dead-ends) of searching information and services has negative impact on the customers regarding the usability of the website. In this research 204 consumers participated over two websites, regarding measuring the relationship between the numbers of dead-ends and website usability. It is clear that many websites are not fulfilling the customer needs; consumers find it hard to shop or find information online. In best case scenario users will get the information and services easily. In worst case the leave the website [16].Today most of the software are difficult to operate, hard to learn and complicated in understanding. These issues wastes the user's time, causes worry and discourage further use of the software. While usable software increases employee satisfaction. If the software in not usable it may cause then it is time consuming and the user may be discouraged from using advanced features. Difficult to learn software also increases the cost of training. While usable software increases user's satisfaction. Non user friendly software reduces motivation of using software [17].

Keeping in view all of the above facts and the identified gap between web developers and web user, there was a strong motivation for developing web usability management model to bridge the gap. For this purpose Web Usability Management Model (WEB-UMM) is proposed.

III. AIMS AND OBJECTIVE

The main objective of the current research study is to develop Web Usability Management Model (WEB-UMM).The model will help web development organizations to improve the usability of a web site which will lead to the success of the website in term of usability. If the proposed model is successfully adopted in the website development industry, website users may accomplish their own tasks in a quick and easy way. Figure 1 shows the basic idea of web usability. Clients check out the web page having bad usability and varies from the web page to another web page having good usability.

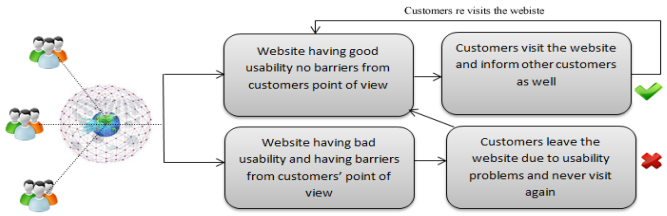

Fig. 1. The Basic idea of web usability

IV. RESEARCH QUESTIONS

There are six research questions that have motivated the work reported in this study.

RQ1. What are the challenges/barriers, as identified in the literature, faced by the internet users in the context of web usability that have negative impact on the users?

RQ2. What are the critical success factors, as identified in the literature, to be adopted by the web developer/vendors that have positive impact on the web users in the context of web usability?

RQ3. What are the challenges/barriers, as identified in the website development industry, faced by the internet users in the context of web usability that have negative impact on the users ?

RQ4. What are the critical success factors, as identified in the website development industry to be adopted by the web developer/vendors that have positive impact on the web users in the context of web usability?

RQ5. Are there differences between the factors and barriers identified through the literature and real-world practice?

RQ6. Is the WEB-UMM robust in terms of measuring the website usability level from customers' perspective?

For RQ1 and RQ2, the identification of CBs (Critical Barriers) and CSFs (Critical Success Factors), we will use systematic literature review (SLR). We will conduct empirical study (questionnaire survey, interviews) for the validation of SLR findings and to answer research questions RQ3, RQ4, RQ5 and RQ6.

Web Usability Management Model WEB-UMM will be developed using SLR guidelines[18, 19].The model will be built on web users experiences i.e. what are the CSFs that have positive impact on the web users in the context of web usability. Also CBs from users prospective will be identified in the litterateur via SLR (Systematic Literature Review). Web usability Management model WEB-UMM will help web development organizations to improve the usability of a web site which will lead to the success of the website in term of usability. If the proposed model is successfully adopted in the website development industry, web users may accomplish

their own tasks in a quick and easy way, which will enhance business credibility.

V. RESEARCH METHODOLOGY

For the development of WEB-UMM (Web Usability Management Model) we will use the same software engineering research methods as used by [20-22].

A. Data Collection And Analysis

For RQ1 and RQ2, the identification of CBs (Critical Barriers) and CSFs (Critical Success Factors), we will use systematic literature review (SLR). A systematic literature review is a means of determining, examining and understanding all available research appropriate to a particular research query, or specialized niche, or trend of interest [23].Systemic literature review SLR) has become an important research strategy in software engineering, since the release of evidence-based software engineering(EBSE) in 2004 [24]. In order to conduct SLR for our desired results first we will design a protocol based on SLR, and then we will implant the protocol to get the CBs (Critical Barriers) and CSFs (Critical Success Factors) for answering RQ1 and RQ2. We will conduct empirical study (case studies, interviews and questionnaire) for the validation of SLR findings and to answer research questions RQ3, RQ4, RQ5 and RQ6.

B. Steps in WEB-UMM Development

The model will be built in eight steps as shown in Figure 3. The first step in the development of WEB-UMM is to set criteria for its success. Second step is the development of research questions. Step 3 and 4 are the steps where data will be collected and analysed. Step 5 is the stage where rationalisation and structuring of results will be performed. Based on the empirical findings, the WEB-UMM model will be developed. In the final stage an evaluation of the WEB-UMM will be performed through case studies. The model will be mathematically represented using fuzzy logic concepts.

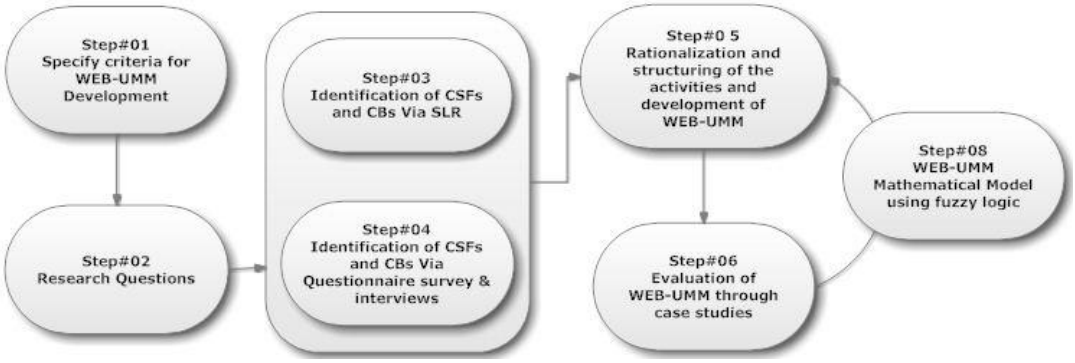

Fig. 3. Stages involved in building WEB-UMM

C. WEB-UMM Assessment

The case study approach will be used in order to assess and review the proposed model i.e. WEB-UMM. For this purpose four case studies will be conducted at web development organizations to assess its efficiency of WEB-UMM. In order to get feedback about WEB-UMM from the participants, focus group sessions will be conducted at the end of these case studies. The focus group sessions is chosen for feedback the reason being data from focus groups is mainly from the interaction between the members in the group without any suggestion or help from the members in the group without any implication or help from the researchers.

VI. CONCLUSION

We have made the following progress so far: knowledge about the background of web usability, identification of research questions, selection of research methodology and identification of activities involved in building WEB-UMM. We have designed the systematic literature review (SLR) protocol. We have now been entered in the implementation phase of the SLR protocol for WEB-UMM.

REFERENCES

[1] Farmer, J., Content and Useability of Australian Life Insurance Companies' Web Sites, ed. R.P.N. 1999/02. September 1999: DEPARTMENT OF ACTUARIAL STUDIES RESEARCH PAPER SERIES.

[2] Backlund, J., Web interfaces and usability. 2001.

[3] Lee, A.T., Web usability: a review of the research. Access Date: 11 July 2012. http://www.acm.org/sigchi/bulletin/1999.1/lee.pdf, 1999. 31(1).

[4] Nga, N.Q., Evaluating the Web Usability – An Interpretive Study of the Bazar portal in Oslo Public Library, Norway. June 2010.

[5] Wang, J. and S. Senecal, MEASURING PERCEIVED WEBSITE USABILITY. Journal of internet commerce. - Binghamton, NY: Haworth Press, ISSN 1533-2861, ZDB-ID 21369008. - Vol. 6.2007, 4, p. 97-112, 2007.

[6] Atterer, R., Usability Tool Support for Model-Based Web Development, PhD Thesis. 2008, Ludwig-

Maximilian University München (http://edoc.ub.uni-muenchen.de/9296/2/Atterer_Richard.pdf).

[7] Hassan, S. and F. Li, Identifying Web usability criteria: The 'Scanmic' Model. 2001.

[8] Farmer, J., Content and Useability of Australian Life Insurance Companies' Web Sites. September 1999: Division of Economic and Financial Studies Macquarie University Sydney NSW 2109 Australia.

[9] Kaur, D.A. and D. Dani, The Systematic Review of Web Measurement. International Journal of Computer Science & Engineering Survey (IJCSES), August 2011. Vol.2(No.3).

[10] Alva, M., et al., Comparison of Methods and Existing Tools for the Measurement of Usability in the Web, in Web Engineering. 2003, Springer Berlin / Heidelberg. p. 386-389.

[11] Tiedtke, T., C. Märtin, and N. Gerth, eds. AWUSA – A Tool for Automated Website Usability Analysis. 2002, Augsburg University of Applied Sciences Department of Computer Science, Baumgartnerstrasse 16 D-86161 Augsburg.

[12] MADAN, A. and S.K. DUBEY, USABILITY EVALUATION METHODS:A LITERATURE REVIEW. International Journal of Engineering Science and Technology (IJEST), February 2012. 04(02).

[13] Muhammad Arslan and M.A. Riaz, A Roadmap for Usability and User Experience Measurement during early phases of Web Applications Development. April 2010, School of Computing,Blekinge Institute of Technology Box 520 SE – 372 25 Ronneby Sweden.

[14] Ahmad, W.F.W., S. Sulaiman, and F.S. Johari. Usability Management System (USEMATE):A Web-Based Automated System for Managing Usability Testing Systematically. in 2010 International Conference on User Science and Engineering. 2010.

[15] DUBEY, S.K. and A. RANA, Analytical Roadmap to Usability Definitions and Decompositions. International Journal of Engineering Science and Technology (IJEST), 2010. 2(9): p. 4723-4729.

[16] Nantel, J., S. Sénécal, and A. Mekki-Berrada, The Influence of "dead-ends" on Perceived Website Usability. Working paper, RBC Chair of e-Commerce. Journal of E-Business, 2004. 5.

[17] Bevan, N. and M. Macleod, Usability measurement in context. Behaviour and Information Technology, 1994. 13: p. 132-145.

[18] Kitchenham and Barbara, Procedures for Performing Systematic Reviews. 2004, Keele University.

[19] Kitchenham, B. and C. Charters, Guidelines for performing Systematic Literature Reviews in Software Engineering, Keele University and Durham University Joint Report. 2007.

[20] Khan, S.U. and M.K. Niazi, Systematic Literature Review Protocol for SOVRM Software Outsourcing Vendors Readiness Model. Aug-2008, School of Computing and Mathematics Keele University ,Keele, Staffordshire, ST5 5BG, U.K.

[21] Nidhra, S. and M. Yanamadala, Knowledge Transfer Challenges and Mitigation Strategies in Global Software Development. May 2012, School of Computing Blekinge Institute of Technology SE-371 79 Karlskrona Sweden.

[22] Ahmad, A. and S.J. Kolla, Effective Distribution of Roles and Responsibilities in Global Software Development Teams. December 2011, School of Computing Blekinge Institute of Technology SE-371 79 Karlskrona Sweden.

[23] Kitchenham, B., Procedures for Performing Systematic Reviews. 2004, Keele University.

[24] Zhang, H., M.A. Babar, and P. Tell, Identifying relevant studies in software engineering. Information and Software Technology, 2011. 53(6): p. 625-637.

Classification Techniques in Blood Donors Sector – A Survey

Nor Syuhada Che Khalid[1], Burhanuddin M.A.[2], Asmala Ahmad[3], Ghani M.K.A[4]
Faculty of Information Technology and Communications
Universiti Teknikal Malaysia Melaka, Malaysia
norsyuhadachekhalid@gmail.com, {burhanuddin[2], asmala[3], khanapi[4]}@utem.edu.my

***Abstract*—This paper focuses on the classification and the recent trends associated with it. It presents a survey of the classification system and clarifies how classification and data mining are related both to each other. Classification is arranging the blood donor dataset into the predefined group and helpful to predict group membership for data instances. This enables users to search target donors become easier because the blood stocks always required replacing expired stocks after a certain period and useful in emergency demands such as surgery and blood transfusion. This paper has also sought to identify the research area in classification to fulfill gaps where further work can be carried on.**

Keywords: Classification, Data Mining, Blood Donors Prediction

I. INTRODUCTION

Safe blood donor recruitment is a challenging task. It is necessary for the community to realize that blood donation is their responsibility. Blood bank, hospital or government cannot sustain health care without sufficient blood from such donor and blood donor organizations play a more crucial role in this endeavor[8]. The problem concerns the identification of blood donors patterns on the basis of blood donor attribute measurements[7]. The blood donors data collection always in large and unexpected amount. Classification is important to predict a certain outcome based on a given input, that is blood donors dataset. The classification algorithm will discover the relationships between the blood donors attributes that would make it possible to predict the outcome[13]. Data mining is a step in knowledge discovery in databases (KDD) which is used to extract interesting patterns from the blood donor data that are easy to perceive, interpret, and manipulate. Several major kinds of classification methods such as decision tree and artificial neural network (ANN) are surveyed in this paper.

II. CLASSIFICATION AND DATA MINING

A. Classification

Classification is important to predict a certain outcome based on a given input, which is the blood donors dataset. Classification is the process of finding a model or function that identifies and distinguishes blood donor data classes or concepts, for the purpose of being able to use the model to predict the class of objects whose class label is unknown. The derived model is based on the analysis of a set of training data when data object whose class label is known[7].

The derived model which is suitable for blood donors dataset could be demonstrated in several forms, such as classification (IF-THEN) rules, decision trees, mathematical formulas, or neural networks. A decision tree is a flow-chart-like tree structure, where each node denotes a test on an attribute value, each branch represents an outcome of the test, and tree leaves represent classes or class distributions. Decision trees could only be converted to classification rules. A neural network, when used for classification, is typically a collection of neuron-like processing units with weighted connections between the units. There are many other methods for building classification models for the blood donors dataset, such as naïve Bayesian classification, support vector machines, and *k*-nearest neighbor classification. Prediction models are continuous-valued functions, while classification predicts categorical (discrete, unordered) labels. It is specifically used to predict missing or unavailable numerical data values rather than class labels. Regression analysis is a statistical methodology that is most often used for numeric prediction, although other methods exist as well. Prediction also comprises the identification of distribution trends based on the available data.

Classification and prediction may need to be preceded by relevance analysis of blood donor data, which attempts to identify attributes that do not contribute to the classification or prediction process. These attributes can then be omitted[4].

B. Data Mining

Data mining is defined as the process of discovering patterns in the blood donor data. The process must be automatic or semiautomatic. The patterns discovered must be significant in that they lead to some advantage, usually an economic or strategic advantage for blood donor analysis. The data includes blood donor data regularly present in significant quantities. Useful patterns allow nontrivial predictions on new data. There are two extremes for the expression of a pattern: as a black box whose innards are effectively incomprehensible and as a transparent box whose construction reveals the structure of the pattern. Both should make good predictions. The difference is whether or not the patterns that are mined are represented in terms of a structure that can be examined, reasoned about, and used to inform future decisions. Such patterns are called structural because they capture the decision structure in an explicit

manner. In other words, they assist to explain something about the data analyzed[4].

The data mining models and tasks are shown in Fig. 1 given below:

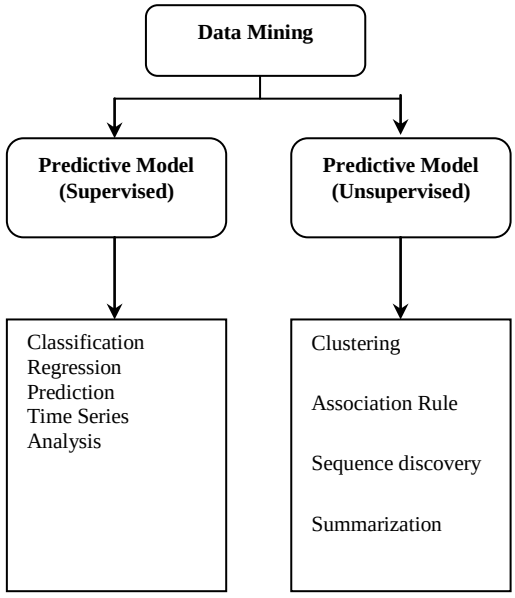

Figure 1: Data Mining Models

The predictive model makes prediction about the unknown data value using the known values. The descriptive model identifies the patterns or relationships in data and investigates the properties of the data analyzes.

There are different types of data mining tasks depending on the use of data mining results. These data mining tasks are categorized as follows:

1) *Exploratory Data Analysis:* It is merely exploring the data without any clear ideas of any searched target. These techniques are interactive and visual.

2) Descriptive Modeling: It describes all the data, it includes models for overall probability distribution of the data, partitioning of the p-dimensional space into groups and models describing the relationships between the variables.

3) Predictive Modeling: This model permits the value of one variable to be predicted from the known values of other variables.

4) Discovering Patterns and Rules: It is concerns with pattern detection, the aim is spotting fraudulent behavior by detecting regions of the space defining the different types of transactions where the data point significantly different from the rest.

5) Retrieval by Content: It is finding pattern similar to the pattern of interest in the dataset. This task is most commonly used for text and image datasets[1].

III. CLASSIFICATION TECHNIQUES IMPLEMENTATION SURVEY

In this section, a lot of research work has been recorded from past few years. They are presented here in chronological order.

First, the surveyed blood donor data are utilized for machine learning techniques of Artificial Intelligence to classify the blood donor group into donors and non-donors. Moreover, the accuracy testing of the surveyed information is performed using the artificial neural network (ANN) and decision tree techniques in order to predict from a series of individual blood behavior data whether or not each individual is a donor. The results indicate that the accuracy, precision, and recall values of ANN technique are higher than those of the decision tree technique. The problems noted in this research are inefficient to collect blood based on patients' usage and difficulty to stabilize blood demand and supply consequently[2][9]. This research has done comparative of neural network and decision tree evaluation based on model accuracies. However, these comparisons are not indicating how model structures will affect the model accuracies results such as how input nodes and the number of hidden layers in an MLP neural network is better than Tree-Building and Tree-Pruning in CART and C4.5 decision tree process making. As a conclusion, this research did not show how advantageous of a neural network classification model than a decision tree by mentioned their model structures and their own advantage and disadvantage factors.

Then, an attempt has been made to classify and predict the number of blood donors according to their age and blood group. J48 algorithm and WEKA tool have been used for the complete research work. The problems found in this attempt are to evaluate J48 decision tree classification through the blood donors training dataset and verify the result using a test set of data[3]. However, this effort has not analyzed through process of WEKA's usage and visual result by how J48 structure is produces the dataset training and evaluations.

The proposed intelligent system that has combined cluster and classification analysis provided the blood center management a more understanding in blood donors' intentions and behaviors. The classification analysis is combined decision tree, NaiveBayes and NBTree. The results are to be used as strategic references for the purpose of enhancing service quality to blood donors and blood supply quality. This system is intended to improve the performance of blood donation, persuade the blood donors to contribute to the high frequency of blood donation and maximize blood collection. In the classification analysis step, the dataset has adopted a single decision tree, naïve bayes, NBTree and bagging essembles to compute accuracy, sensitivity, receiver operating characteristics (ROC curve), specificity and area under the ROC curve[5]. However, this proposed intelligent system has not focused on the classification structures that have integrated with each other and produced the accuracy, sensitivity, receiver operating

characteristics (ROC curve), specificity and area under the ROC curve.

A specific case of real world Blood Transfusion dataset to demonstrate how each KDD step can be completed interactively to produce useful and required knowledge in less time/efforts with the synergy of human and computer system. The classification technique used is the decision tree[10]. However, this case has not explained how decision tree classification in TANAGRA software is making the prediction result.

Subsequently, the scoring algorithm implemented which is developed using CART classification for the dashboard also helps with the optimized deployment of budget resources and budget allocation determination for blood donation campaigns. The goal of this classification model is to provide the capability to determine voluntary blood donor based on blood donation patterns[11]. However, this implementation has not explained how RVD structure is implemented together with the CART classification structure.

A comparison of donation models using the classification algorithms for data mining that are the extended RVD model and DB2K7 which enable representation as decision trees. The analysis provides insight into the development of donor classification which enables blood banks to determine the kinds of donor profiles and manage blood donation related activities like recruitment and campaigns for blood donations. The goal of this classification model is to provide the capability to determine voluntary blood donor based on blood donation patterns[12]. However, the comparison of classification accuracies has not shown the difference in model structure of RVD, DB2K7 and extended RVD which have different capability to each classification.

An analysis had been carried out using a standard blood group donor's dataset and using the J48 decision tree algorithm implemented in WEKA. The research work is used to classify the blood donors based on the sex, blood group, weight and age[7]. However, the tree diagram that has produced J48 decision tree model through WEKA should be explained clearly to define the result produced.

A fuzzy sequential pattern mining algorithm is used to mine fuzzy sequential patterns from the Blood Transfusion Service Center dataset. It helps to predict future patterns of blood donating behavior. The problem of this mining algorithm before is the sequential pattern mining algorithms do not allow the processing of numerical data and require converting this data into binary representation[15]. In the end, the proposed model is shown the better improvement by algorithm structure explanation.

An identification of the blood donation behavior is done by applying the classification algorithms for data mining. The analysis had been implemented using a standard blood transfusion dataset and using the CART decision tree algorithm implemented in WEKA[8]. However, the comparison of CART and RVD is incomplete because no detailed analysis of how the classification accuracy that has improved by a leaf node number of tree complexity CART algorithm rather than RVD algorithm.

A particle swarm based hybrid system for remedying the class imbalance problem in medical and biological data mining. This hybrid system combines the particle swarm optimization (PSO) algorithm with multiple classifiers and evaluation metrics for evaluation fusion. Samples from the majority class are ranked using multiple objectives according to their merit in compensating the class imbalance, and then merged with the minority class to form a balanced dataset[14]. However, this system analysis has not specified about the conditions or criteria of imbalanced dataset happened such as random undersampling, random oversampling and clustering based sampling and their examples.

Finally, two artificial neural network models (multi-layer perceptron neural network [MLP] and probabilistic neural network [PNN]) are compared to a standard statistical method (linear discriminant analysis [LDA]). The variable sets considered are sex, age, educational level, altruistic values, the perceived risks of blood donation, blood donation knowledge, attitudes toward blood donation, and intention to donate blood. This comparison demonstrates how it is possible to identify various dimensions of blood donation behavior by uncovering patterns in the dataset, and also indicates the classification abilities of two neural network techniques[6]. However, the comparison of MLP and PNN with LDA only give results on classification accuracy without classification abilities, such as how their strength or weakness makes differences in accuracies that have produced.

As a conclusion, the occurrences of classification techniques in blood donor data research papers from 2009 to 2012 are in Table I as follow.

TABLE I. THE OCCURRENCES OF CLASSIFICATION TECHNIQUES IN BLOOD DONORS DATA RESEARCH PAPERS 2009 – 2012

Techniques		**The occurrences in research papers based on year**			
		2009	***2010***	***2011***	***2012***
1.	Artificial Neural Network (ANN) – Backpropagation (BP)				
2.	Artificial Neural Network (ANN) – Multi-layer Perceptron (MLP)	1			1
3.	BPSO hybrid module with multiple classifiers – Decision tree (J48), k-Nearest Neighbor (kNN), Naïve Bayes (NB), Random Forest (RF) and Logistic Regression (LOG).	1			
4.	Decision tree – C4.5				1
5.	Decision tree – CART algorithm		1	1	
6.	Decision tree – DB2K7			1	
7.	Decision tree – extended RVD based model			1	
8.	Decision tree – ID3			1	
9.	Decision tree – J48			1	3

10.	Decision tree – RVD			1	
11.	Decision tree – single decision tree			1	
12.	Fuzzy sequential pattern mining			1	
13.	K-Nearest Neighbor (kNN)				
14.	Linear discriminant analysis (LDA)	1			
15.	NaiveBayes			1	
16.	NBTree			1	
17.	Probabilistic neural network (PNN)	1			
18.	Sequential pattern mining			1	
19.	Bagging essembles			1	

IV. PROPOSED METHODOLOGY

In the proposed methodology, based on surveys and some analysis, a blood donors classification implementation supposedly has the following objectives:

- To propose better blood donor classification module for data mining purpose.
- To show clear analysis of strength and weakness classification technique in blood donors that will produce the classification accuracy.
- To increase and attract blood donor's rate and their attitude.
- To assist blood bank professionals in making policy on the acquisition of blood donors and new blood banks.

There are several major classification had been acquired and used in data mining research projects that recently including neural network, decision tree, statistics, distance and rule based. These methods in classification need brief mention to have a better understanding and potentially used for future research. Some of these methods have already occurred in research work with a few of classification model limitations that could suggest as further research gaps.

The estimation and prediction may be viewed as types of classification. The problems usually are evaluating the training dataset and then to apply them into the future model. The following Table II shows different classification algorithms [1].

TABLE II. CLASSIFICATION ALGORITHMS

Type	**Name of Algorithm**
Statistical	Regression Bayesian
Distance	Simple distance K-nearest neighbors
Decision Tree	ID3 C4.5 CART SPRINT
Neural Network	Propagation NN Supervised learning Radial base function network
Rule based	Genetic rules from DT Genetic rules from NN Genetic rules without DT and NN

Details of classification algorithm are given as follow:

A. Statistical

Statistical approaches are usually characterized by having an explicit underlying probability model, which provides a probability of being in each class rather than just a classification. Additionally, it is commonly taken for granted that the techniques will use by statisticians, and hence some human intervention is assumed with respect to variable selection and transformation, and overall structuring of the problem[16].

B. Distance

Traditionally, distance-based classifiers are instance based: they classify a test instance by the computation of a similarity measure between that instance and the instances in the training-set and assigning it the same class of the most similar k instances. This method is simple but has some disadvantages, among which there is the greater sensitivity to the local presence of noise in the training-set, the need to compute many similarity measures and the difficulty to decide the value of k[17].

C. Decision Tree

A decision tree is a flowchart-like tree structure, where each internal node (non leaf node) denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (or terminal node) possesses a class label or class distribution. The highest node in a tree is the root node. Decision trees could easily change to classification rules. The structure of decision tree classifiers does not require any domain knowledge or parameter setting, and therefore is appropriate for exploratory knowledge discovery. Decision trees can handle high dimensional data. Their representation of acquired knowledge in tree form is intuitive and generally comfortable. The learning and classification steps of decision tree induction are simple and quick. Generally, decision tree classifiers have good accuracy. However, successful use may depend on the data at hand. Decision tree induction algorithms have been used for classification in many application domains. Decision trees are the basis of some commercial rule induction systems. However, there are concepts that are hard to learn because decision trees do not express them clearly such as XOR, parity or multiplexer problems. Then, the decision tree becomes too large. Another issue is data that include categorical variables with different number of levels make an information gain in decision trees is biased in favor of those attributes with more levels [9].

D. Neural Network

The neural network is applied in data mining for prediction or classification information in the domain of ideas or feelings and behaviors of consumers effectively. An artificial neural network, often simply called a neural

network, is a mathematical model inspired by biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. The neural network is the model learning patterns of the data to solve the problem of classification and clustering, which are effective to analyze the marketing databases. However, issues occurred in a neural network are overtraining and local optimum in backpropagation[9].

E. Rule based

Rules are a good way of representing information or bits of knowledge. A rule-based classifier uses a set of IF-THEN rules for classification. The rule ordering scheme prioritizes the rules beforehand. The ordering may be class based or rule-based. The classes are sorted in order of decreasing "importance," such as by decreasing order of prevalence, with class-based ordering. All of the rules for the most prevalent or most frequent class come first; the rules for the next prevalent class come next, and so on. Alternatively, they may be sorted based on the misclassification cost per class. The rules are not ordered because they all predict the same class within each class. Therefore no class conflict occurs. The rules are organized into one long priority list, according to some measure of rule quality such as accuracy, coverage, or size (number of attribute tests in the rule antecedent), or based on advice from domain experts with rule-based ordering. When rule ordering is used, the rule set is known as a decision list. With rule ordering, the triggering rule that appears earliest in the list has highest priority, and then it gets to fire its class prediction. Any other rule that satisfies the similar class is ignored. Most rule-based classification systems use a class-based rule-ordering strategy. Remember that in the first strategy, overall the rules are unordered. They can be used in any order when classifying a tuple. A disjunction (logical OR) is implied between each of the rules. Each rule represents a stand-alone nugget or piece of knowledge[9].

V. CONCLUSION

This paper presents a survey of classification and its techniques which have been used to extract interesting patterns and to develop significant relationships among variables stored in a huge dataset, such as blood donors data. Classification is needed in many fields to extract the useful information from the large quantity of data. Large quantity of data is maintained in every field to keep different records. Consequently, different ways have been discovered to analyze the data automatically, to summarize it, to discover and characterize trends in it and to flag anomalies automatically. The various classification techniques are introduced by the different researchers. These techniques are used to do (IF-THEN) rules, decision trees, mathematical formulas, or neural networks, to find interesting patterns. In future work, the classification techniques will be implemented on the blood donors dataset for predicting the blood donor's behavior and attitude, which have been collected from the blood bank center and the limitations of classification methods implemented could be improved as research gaps.

ACKNOWLEDGMENT

The authors are grateful to Faculty of Information and Communications Technology and Universiti Teknikal Malaysia Melaka for providing facilities and financial support for this research project.

REFERENCES

[1] A. Bhardwaj, A. Sharma, and V. K. Shrivastava, "Data Mining Techniques and Their Implementation in Blood Bank Sector – A Review," *International Journal of Engineering Research and Applications (IJERA)*, vol. 2, no. August, pp. 1303–1309, 2012.

[2] W. Boonyanusith and P. Jittamai, "Blood Donor Classification Using Neural Network and Decision Tree Techniques," in *Proceedings of the World Congress on Engineering and Computer Science*, 2012.

[3] T. Franks, "Predicting the Number of Blood Donors through their Age and Blood Group by using Data Mining Tool," *American Journal of Medicine and Health*, vol. 2, no. 2, pp. 8–14, 2012.

[4] J. Han and M. Kamber, "Data Mining: Concepts and Techniques", pp. 1–743, 2006.

[5] W. Lee, "An Intelligent System for Improving Performance of Blood Donation," *Journal of Quality*, vol. 18, no. 2, pp. 173–185, 2011.

[6] M. M. Mostafa, "Profiling blood donors in Egypt: A neural network analysis," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5031–5038, Apr. 2009.

[7] P. Ramachandran, N. Girija, and T. Bhuvaneswari, "Classifying Blood Donors Using Data Mining Techniques," *IJCSET*, vol. 1, no. 1, pp. 10–13, 2011.

[8] T. Santhanam and S. Sundaram, "Application of CART Algorithm in Blood Donors Classification," *Journal of Computer Science 6*, vol. 6, no. 5, pp. 548–552, 2010.

[9] A. Sharma and P. C. Gupta, "Predicting the Number of Blood Donors through their Age and Blood Group by using Data Mining Tool," *International Journal of Communication and Computer Technologies*, vol. 01, no. 6, pp. 6–10, 2012.

[10] V. Singh and S. Nagpal, "Interactive Knowledge Discovery in Blood Transfusion Data Set," *VSRD International Journal of Computer Science & Information Technology*, vol. 1, no. 8, pp. 541–547, 2011.

[11] S. Sundaram and T. Santhanam, "Real-Time Blood Donor Management Using Dashboards Based on Data Mining Models," *International Journal of Computer Science Issues*, vol. 8, no. 5, pp. 159–163, 2011.

[12] S. Sundaram and T. Santhanam, "A Comparison of Blood Donor Classification," *Journal of Theoretical and Applied Information Technology*, vol. 30, no. 2, pp. 2–5, 2011.

[13] F. Voznika and L. Viana, "Data Mining Classification." pp. 1–6.

[14] P. Yang, L. Xu, B. B. Zhou, Z. Zhang, and A. Y. Zomaya, "A particle swarm based hybrid system for imbalanced medical data sampling.," in *BMC genomics*, pp. 1–14, 2009.

[15] F. Zabihi, M. Ramezan, M. M. Pedram, and A. Memariani, "Rule Extraction for Blood Donators with Fuzzy Sequential Pattern Mining," *The Journal of Mathematics and Computer Science Ava*, vol. 1, no. 1, pp. 37–43, 2011.

[16] E. D. Michie, D. J. Spiegelhalter, and C. C. Taylor, "Machine Learning , Neural and Statistical Classification," 1994.

[17] D. Bachar and R. Meo, "A distance-based classifier built on class model.

Improving Data Accessibility Using QR Code in Healthcare Domain

Fathin N. M. Leza, Nurul A. Emran, Mohd. Khanapi A. Ghani
Department of Software Engineering
Faculty of Information Technology and Communication Universiti Teknikal Malaysia Melaka (UTeM)
Melaka, Malaysia
M031220043@student.utem.edu.my, {nurulakmar,khanapi}@utem.edu.my

***Abstract*—In electronic healthcare environment, one factor that determines the quality of healthcare services is the accessibility of patient data. Fast and seamless access to patient data is necessary in order to support sharing of complete lifetime patient data among healthcare providers while promoting the greater use of it. Nevertheless, studies on how complete lifetime patient data can be acquired by improving the way these data are accessed are limited. In this paper, the conceptual model of data accessibility in healthcare domain will be presented. QR code technology will be adopted in the model as a tool that enables multi-platforms data access points. Incremental data updates will be used as a mean to share complete, up-to-date lifetime patient data. Software prototype architecture and the interface design of the model will be also presented in this paper. The contribution of the work is on improving lifetime patient data accessibility that consequently promotes sharing of complete patient data among healthcare providers.**

***Index Terms*— Data Accessibility, Mobile Application, QR Code, Healthcare Records**

I. INTRODUCTION

In electronic healthcare environment the prominent concern is not only on the accessibility of specific data of the patient under treatment, but also on the accessibility of his/her lifetime patient data. Acquiring complete lifetime patient data is crucial as information about patients' past treatments (and past healthcare providers), allergies, basic information (e.g. blood type and inherited diseases) will determine how accurate doctors/physicians can diagnose certain disease before suitable type of treatments and prescriptions can be recommended. Ideally, complete lifetime patient data hold information about a person's medical history from born until now. As one usually gets medical treatments from more than one healthcare provider (i.e., clinics, hospitals, rehabilitation centers and etc.) sharing of complete lifetime patient data among healthcare providers is a requirement that demands for fast and seamless access to these data. The consequence of incomplete patient data unfortunately must be borne by the patients. For example, as some data about a patient under treatment is missing he/she will be required experiencing similar tedious, time-consuming and painful medical procedures every time they are admitted for treatment (in one or more healthcare provider) until patient data that are acceptably 'complete' can be acquired.

The example just given portrays a lack of data sharing problem among healthcare providers which is caused by limited data access. In reality, one obstacle of data sharing that can be observed is the complexity of performing full data integration among healthcare providers' data sources. Even though full data integration can provide unified view and access of data from different data sources, one requirement of this approach is to resolve differences (called heterogeneity) among the participating healthcare systems (e.g., in term of hardware and data management software). This requirement makes the task of accessing and gathering complete lifetime patient data expensive and impractical. Therefore in this paper, we propose a data accessibility model which is a departure from the full data integration approach that eases the task of gathering complete lifetime patient data.

Before we can design (and evaluate) the model prototype, we must understand the state-of-the art of data accessibility methods available. Therefore in the next section, works related to data accessibility will be presented. This is followed by description about the model we proposed.

II. RELATED WORK

Data accessibility is the ability to access the data regardless of natural or man-made disasters, hardware failures or others [1]. There are many issues regarding to data accessibility has been pointed out by researchers in several domains. For example, Smith pointed the problem for controlling illegal access of data [2] . The author suggests to utilize multi-level security database to authorized personnel only. This is for restricting the access for confidential information in the database problem [3]. However, solution proposed by Smith do not result in direct revelation of data and unreliable. This is because it violates the integrity constraints of the database.

Hara highlights data accessibility in ad-hoc network [4]. The problem raised is regarding restricted data accessibility for individual mobile host against data held by mobile hosts in other network. To deal with the problem just described, Hara suggested data replication. Through data replication has multiple replicas of data items are created in a network. As the result data accessibility can be improved by increasing the probability of finding one copy of data in the network. Nevertheless, data replication is only possible when they are

plenty of storage space, bandwidth and power in mobiles nodes [5].

In healthcare domain Gorea and Felea proposed a medical system called as Telemon in order to ease access for critical patient data during emergency situation by medical staff on-duty [6]. These data were hard to be accessed in real-time by the medical officers on-duty when needed. A similar system called as TELEASIS has been proposed to improve accessibility of patient data such as patients' data [7]. Purnama and Hartati emphasize a web-based system of Electronic Medical Record (EMR) to improve the availability and completeness of data retrieved [8]. However, even though accessing complete patient data is important in healthcare domain, little attention has been given on how these complete patient data can be accessed by from multiple healthcare providers.

The work that we described so far unfortunately lack of discussion on how can we improve data accessibility in order to acquire complete patient data sets. In addition, these works are limited in terms of addressing how individual lifetime patient data can be accessed (and updated) by multiple healthcare providers in real-time. Therefore, in depth investigation is needed to address the gap in data accessibility studies just mentioned. We choose the healthcare domain to validate the usefulness of the model that we proposed in this paper. In the next section we will present the flow of the work that will be implemented to develop the model proposed.

III. RESEARCH METHODOLOGY

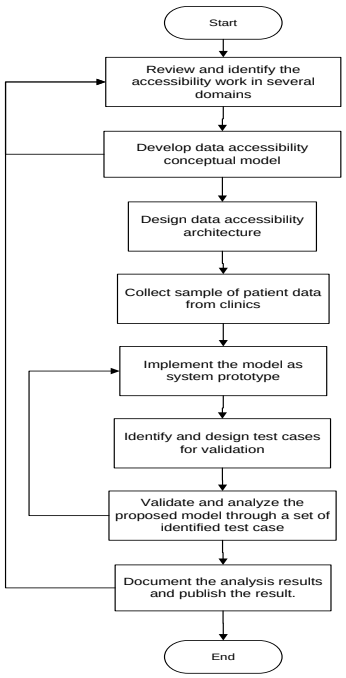

Fig 1. Flowchart of Methodology

This research will be pursued by following the stages as shown in Fig 1.

To identify gaps in data accessibility studies through literature review, every work in data accessibility will be critically evaluated. Then, data accessibility conceptual model will be developed. This is followed by designing data accessibility architecture. Samples of patients data will be collected before we can implement the model as a system prototype. A set of software test cases will be identified and designed for validation purposes. A standard software testing method [9] will be used during the validation stage. Then, the testing results will be analyzed. Finally, the results of analysis will be documented and published.

The implementation of the research stages just described has been scheduled for 18 months. The milestone of the work is shown in TABLE I.

TABLE I. WORK SCHEDULE AND MILESTONE

Milestone	Duration (Month)	Date Started
a) Review and identify accessibility work in several domains	18	April 2013
b) Develop data accessibility conceptual model	1	September 2013
c) Design data accessibility architecture	2	September 2013
d) Collect samples of patient data from clinics	2	November 2013
e) Implement the model as a system prototype	6	December 2013
f) Identify and design test cases for validation	1	March 2014
g) Validate and analyze the proposed model through a set of identified test case	2	April 2014
h) Document the analysis results and publish the result.	5	May 2014

IV. DATA ACCESSIBILITY CONCEPTUAL MODEL

Using UML uses case diagram, the conceptual data accessibility model for healthcare domain is shown in Fig 2. The model, when implemented as a system prototype is called as Patient Lifetime Health System or shortly, PLHS.

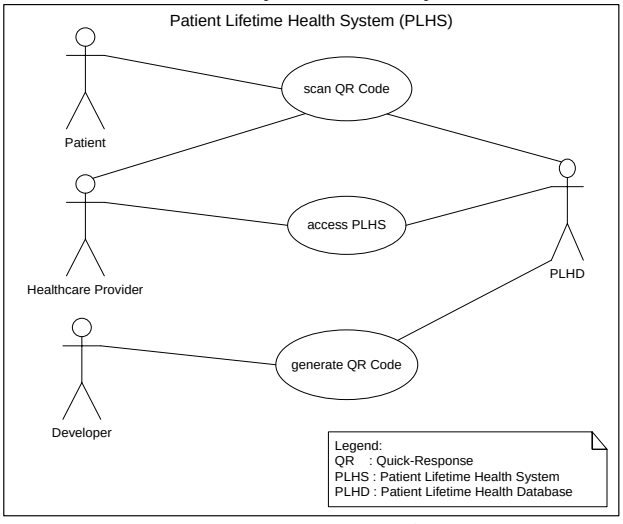

Fig 2. A Conceptual Model of PLHS

This model has four actors who interact with the PLHS prototype. The details are as follow:

- Patient: a person who seeks medical service or treatment from healthcare provider and use the system to scan the QR Code. (Description of QR Code will be provided in sub-section A).
- Healthcare Provider: an institution where the medical services are provided (e.g., clinics and hospital). Medical officers such as doctors/physicians who work at the healthcare provider have the privilege to view and update patient's data.
- Developer: a person who is responsible to generate QR Code that contains patient data.
- PLHD : database that stores patients data

To describe the activities in 'scan QR Code' and 'access PLHS' use cases, we use a flow chart as shown Fig 3. .

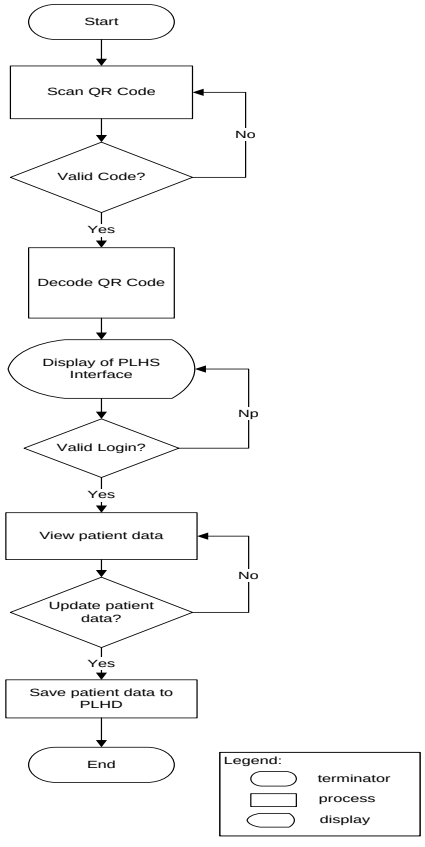

Fig 3. Flow of PLHS Usage

The system begins when a patient provides the QR Code to the medical officer for scanning where he/she seeks for treatment. The code can be scanned either through a desktop application QR Code scanner or a Smartphone scanner (where the QR Code application is installed). Once the QR code is successfully decoded, an interface displaying patient's demographic data will be displayed. Then, the authorized medical officer (i.e. doctor) can login into PLHS. Every valid login will allow the medical officer to view the patient's data and update them. The patient's data will be saved into PLHD for every successful updates, In the next sub-section, the usability of QR Code in several domains will be presented.

A. QR Code Applications

QR stands for Quick Response codes created in 1994 by Toyota subsidiary Denso-Wave where this special code is created as two-dimensional symbology [10]. This technology is the code advanced from the current barcode. It is initially used for tracking parts by vehicle manufacturers [11].

This QR code is quite useful for small businesses that wish to market their products and services. Using QR code are one of the marketing strategies that can be used [12][13]. For example, information displayed on the business card can be shortened using the QR Code. This code is commonly used for labeling purposes in inventory management. The codes that store inventory information are printed on stickers or on papers. With the inventory system, the audit staff is able to scan the coded stickers on company properties to conduct regular audit [14].

QR Code is also used for sending short messages in SMS (short message service) which is generated for "ready –to-send" purpose [15]. By inputting recipient phone number and the content of messages, QR Code is encoded and ready-to-send the message to the recipients. By this way, error caused by user inputs can be avoided. On the other side, QR Code is used as an authentication tool for online-banking [16]. It is used to confirm user's identity which is strengthened by one-time password (OTP). By this way, the system could verify the user's identity who attempt to access their account.

In the healthcare domain, QR codes are mainly used in the hospitals in developed countries. For example, QR codes are used to make online appointments to schedule mammograms in Athens [17]. It is also used to promote hospital services in Nebraska and in labor and delivery room in Forte Collin [17].

The applications of QR Codes in various domains show its usefulness as data accessibility tool. It's popular in marketing, shipping, trading, inventory management and more has been driven by several factors. These factors are its simple features, ease of use, cost-effective and more importantly instant access to the information [18]. In the next sub-section we will describe the architecture of PLHS.

B. System Prototype Architecture

In practice, PLHS will be implemented on the two platforms namely desktop and mobile. Fig. 4 depicts the architecture of PLHS for mobile platform. As depicted in the figure, suppose that a patient (Patient X) seeks healthcare services from several healthcare providers (A, B and C), at a different time interval. The QR Code is scanned in the PLHS interface that developed in the application which later, is decoded into readable patient data. Every time Patient X gets the treatment, the corresponding healthcare provider's doctors can view and update the patient's data which is stored in PLHD. Through this way, the patient's data will be incrementally updated throughout the patient's lifetime.

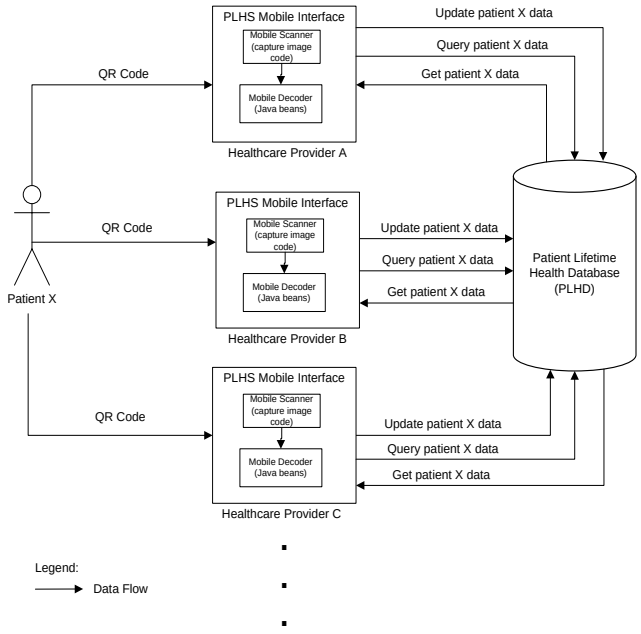

Fig 4. PLHS architecture for mobile platform

Fig. 5 shows the architecture of PLHS for desktop platform. Unlike mobile platform, QR Code scanner is not embedded in the interface as it is a separate device connected to a desktop. A decoder application which is installed on the desktop will be used to decode the QR Code. PLHS Desktop architecture will be developed by using web-based open-sources. Both architectures provide the same interface functionality (for decoding and retrieving data) but they will be implemented in a different way. In the next section we will present the design of PLHS interface.

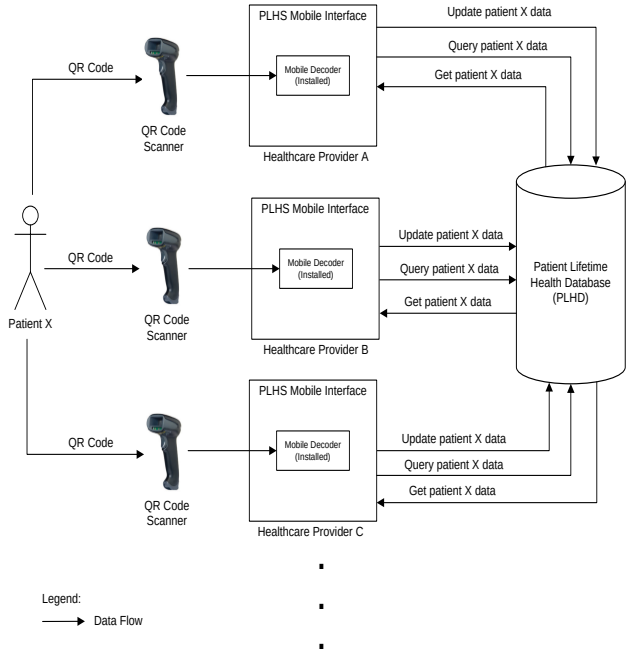

Fig 5. PLHS architecture for desktop platform

C. *PLHS User Interface*

In this paper, user interface of PLHS for mobile platform will be shown. User interface for the mobile platform is developed using Java language with the latest language of HTML (HTML5) which is linked with a Zxing library for barcode application. HTML5 is chosen for its compatibility with many devices and platforms including Android, IOS or Blackberry. PLHS prototype will be developed in Android version 4.2.2 (Jelly Bean) environment. Figure 6 to 9. show the design of interface for mobile devices.

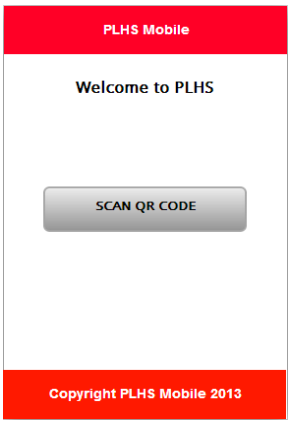

Fig 6. Design for Main User Interface

Fig 7. Design of QR Code Scanner Interface

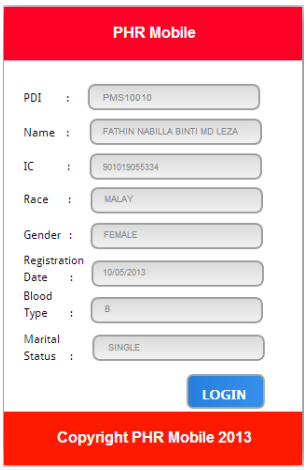

Fig 8. Design for Patient Demographic Display Interface

Fig 9. Design for Updating Patient Data Interface

V. EXPECTED RESULT

The expected result of this research is a new data accessibility model that improves completeness of lifetime patient data through incremental updates by multiple healthcare providers. A system prototype called as PLHS is the result implementing the conceptual model proposed in this research that will be available for desktop platform, as well as mobile platform.

VI. Research Contribution

This research will contribute to the usage of complete lifetime patient data by improving data accessibility among healthcare providers. In addition to its real-time characteristic, the model that we propose support unlimited (and multi-platforms) access points that will be motivate patients' data sharing among healthcare providers.

VII. Conclusion And Future Work

In conclusion, we presented the motivation of the work and the gaps in data accessibility studies. The result of this research will contribute towards better healthcare services. In the future, we will extend the work to cope with practical issues identified during the implementation of the system prototype.

Acknowledgment

The authors would like to thank Universiti Teknikal Malaysia Melaka (UTeM) for their support especially to Faculty of Information Technology and Communication for their support and facilities along this research study. This research is sponsored by research grant FRGS (RACE)/2012/FTMK/SG05/01/1 F00155.

References

[1] Clio, "Blog Data Accessibility , Security , and Privacy," 2008. [Online]. Available: http://www.goclio.com/blog/2008/10/data-accessibility-security-and-privacy-part-i/. [Accessed: 14-Oct-2013].

[2] G. D. Smith, "the accessibility," *Br. Med. J.*, vol. 308, pp. 1519–1520, 1994.

[3] A. O. Adesina, K. K. Agbele, R. Februarie, A. P. Abidoye, and H. O. Nyongesa, "Ensuring the security and privacy of information in mobile health-care communication systems," *S. Afr. J. Sci.*, vol. 107, no. 9/10, pp. 1–7, Sep. 2011.

[4] T. Hara, "Effective replica allocation in ad hoc networks for improving data accessibility," in *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213)*, 2001, vol. 3, pp. 1568–1576.

[5] L. Yin and G. Cao, "Balancing the tradeoffs between data accessibility and query delay in ad hoc networks," *Proc. 23rd IEEE Int. Symp. Reliab. Distrib. Syst. 2004.*, pp. 289–298, 2004.

[6] D. Gorea and V. Felea, "Semantic Integrity Control in the Database Layer of an e-Health System Functional and architectural perspective of Telemon e-health system," *Int. J. Comput. Commun. Control*, vol. III, no. Proceedings of ICCCC, pp. 162–167, 2008.

[7] V. Topac and V. Stoicu-Tivadar, "Patient empowerment by increasing information accessibility in a telecare system.," *Stud. Health Technol. Inform.*, vol. 169, pp. 681–5, Jan. 2011.

[8] B. E. Purnama and S. Hartati, "Convenience and Medical Patient Database Benefits and Elasticity for Accessibility Therapy in Different Locations," *Int. J. Adv. Comput. Sci. Appl.*, vol. 3, no. 9, pp. 54–60, 2012.

[9] S. Engineering, S. Committee, and I. Computer, "IEEE Standard for Software and System Test Documentation," *IEEE Std 829-2008*, vol. 2008, no. July, pp. 1–150, 2008.

[10] M. Badurowicz, "Two-dimensional barcodes in mobile applications," *Actual Probl. Econ.*, no. 132, pp. 346–350, 2012.

[11] T. A. N. Shiang-yen, L. Y. Foo, and R. Idrus, "Application of Quick Response (QR) Codes in Mobile Tagging System for Retrieving Information about Genetically Modified Food," in *Proceedings of the 9th WSEAS international conference on Data networks, communications, computers, DNCOCO*, 2010, pp. 114–118.

[12] C. Becker and C. Bizer, "DBpedia Mobile : A Location-Enabled Linked Data Browser," pp. 6–7.

[13] R. Brooks, "13 CREATIVE WAYS TO USE QR CODES FOR MARKETING," 2011. [Online]. Available: http://www.fastcompany.com/1720193/13-creative-ways-use-qr-codes-marketing.

[14] M. Limpag, "Leon Kilat : The Tech Experiments Use QR codes , Google Docs to set up free inventory system Related posts :," *max.limpag.com*, 2013. [Online]. Available: http://max.limpag.com/article/qr-code-google-docs-inventory/.

[15] M. Zainuddin, D. Baswaraj, and S. M. Riyazoddin, "Generating SMS (Short Message Service) in the form of Quick Response Code (QR-code)," *J. Comput. Sci. Inf. Technol.*, vol. 1, no. 1, pp. 10–14, 2012.

[16] J. Murkute, H. Nagpure, H. Kute, and N. Mohadikar, "Online Banking Authentication System Using QR-code and Mobile OTP," *Int. J. Eng. Res. Appl.*, vol. 3, no. 2, pp. 1810–1815, 2013.

[17] P. L. Dolan, "The latest health care marketing tool : QR codes," *American Medical News*, 2011. .

[18] G. Charlton, "The pros and cons of QR codes," *Econsultancy*, 2011. [Online]. Available: http://econsultancy.com/my/blog/7884-the-pros-and-cons-of-qr-codes.

Improving Large Database Performance Using SQL Query Rewriting in MES

Nor'azah Md Khushairi[1]
Software Engineering Department, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM),
norazah_khushairi@silterra.com[1]

Nurul A.Emran[2]
Software Engineering Department, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM),
nurulakmar@utem.edu.my[2]

Mohd Mokhtar Mohd Yusof[3]
Software Engineering Department, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM),
mokhtaryusof@utem.edu.my[3]

***Abstract*— Semiconductor fabrication is a complex manufacturing operation managed by manufacturing execution system (MES). In MES, maintaining database performance is a challenging task as the volume of data produced and used everyday is big. The reports need to be produced frequently in shortest period of time possible. Delay in retrieving data used to produce the reports affect the management of manufacturing operations. To assist database administrators and application developers to tune database performance, it is important for them to understand the suitable approaches to deal with the problem. Motivated by database performance problem faced by a semiconductor fabrication company called as Silterra, in this paper, we will propose a novel SQL Query rewriting technique to deal with the problem. The methods used for database performance tuning as seen in the literature will also be presented. The usefulness of the technique will be evaluated through some sets of experiments using real fabrication data within MES architecture setup hosted by Silterra. The contribution of the work is a database performance tuning model based on query rewriting that will be of benefit for big data users community.**

***Index Terms*— query rewriting, database performance tuning.**

I. INTRODUCTION

In modernization of technological era, the amount of data has been exploding in many application domains such as healthcare, public sector, retail and manufacturing. For example, in healthcare, electronic medical records consist of patient's family history, illnesses and treatment, which able to improve preventive care and disease treatment. In some cases, doctors take images with machines like a computerized axial tomography (CAT) or magnetic resonance imaging (MRI). These images can help map out patient's problems and help save lives. The need to retain long term active data or even permanent are increasing [1]. Data are accumulated and transformed to big dataset before they can be stored in a database. Big datasets can cause overhead to Database Management System (DBMS) and lead to database performance issues [2]. Database performance is a crucial issue, which can decrease the ability of the DBMS to respond to queries quickly.

Poor database performance cause negative consequences such as in financial, productivity and quality of the businesses in many application domains. In this research, we will focus on database performance issue in manufacturing domain. This research will base on a real case study in a semiconductor fabrication factory, Silterra Malaysia Sdn Bhd.

Semiconductor fabrication processes are complex, which are characterized by a diverse product mix that is changing over time, having re-entrant process flows; different process types; different types of disruptions [13]. Most of the processes are delicate, for example the process of dicing the wafers must be carefully monitored as the wafers are thin and fragile. Even a tiny scratch may scrap the wafer. These delicate processes require close monitoring, which is beyond human's capability. In addition, processing of 40,000 to 50,000 work-in-progresses (WIP) usually takes 50 to 70 days, 300 to 400 equipments and 300 to 900 steps to complete [3,4,5]. Manufacturing Execution System (MES) is used to manage WIP, equipment automation, material control system (MCS) routing and material transfer within an automated material handling systems (AHMS). Huge amount of data are recorded automatically or semi automatically in multiple databases during fabrication process where the data become the input to monitor lot movements in the semiconductor fabrication plant, or shortly FAB.

MES database composed of a collection of subsystems, each with a specific task that will interact with the user application programs and the database as shown in Fig. 1 below. These data will be retrieve in timely manner to produce meaningful reports. Complexity of the processes in FAB industry contributes to huge database because every single transaction needs to be recorded. Transactions of wafer are characterized by a diverse product mix, re-entrant process flow, different processes types and different disruption [13].

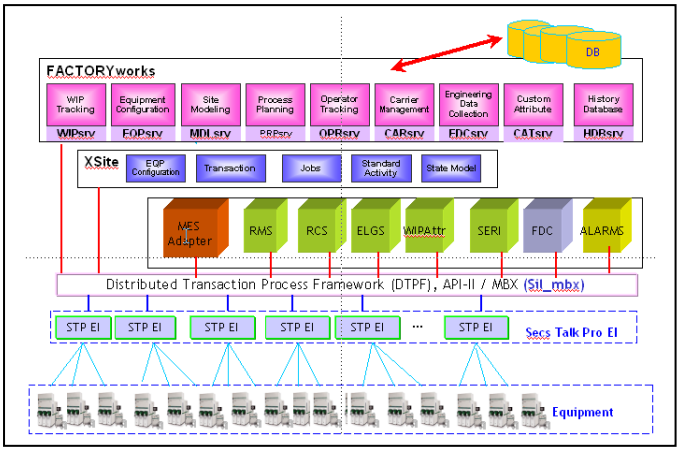

Fig. 1. Silterra MES Architecture

MES Reports acts as a main source to retrieve data of various applications such as FACTORYworks, Xsite, Equipment Interface (EI), AHMS, SPACE, SERI and others. Reports are used to monitor the movement and activities in the FAB. It plays an important role for the management to manage the FAB operation, achieve the daily output target and reduce cycle time. For example if the reporting system not available even for one hour, the line managers will have a difficulty to understand the output they can meet. In Silterra, one hour without reports can cause 2000 wafer moves lost which estimates equivalent to a few thousand USD move lost. In addition, during wafer processing real-time reporting is important to know how many process wafer have go through specially for equipment that need to run multiple processes. As consequences, without real time reporting user can misprocess the wafer.

Reports will be scheduled automatically at various time frames and frequencies. High frequency of query submissions and data retrieval is an overhead that affects the speed of query responses. This scenario leads to competition of system resources (both hardware and software resources), maintaining that requires high database performance level in this domain, which is unfortunately non-trivial.

A. Research Problems

As MES has a complex architecture with multiple modules and databases with large sizes, one of the challenges is to respond to endless queries made by MES users. These queries are issued to produce real time monitoring reports for the managements and customers. MES reports that consists of information about WIP status and history movements starting from the product creation until the end of it. Data from various sources must be gathered and combined before a complete report produced by issuing set of queries. This information is crucial for the management personnel to run the operation and to improve the cycle time in the FAB. On the other hand, the database performance monitoring helps (Database Administrator) DBAs and application developers to locate and identify the causes of bottlenecks due to poor database performance. SQL query is one major problem that affects MES reporting task and database performance as a whole.

Herodotou and Babu pointed out that an attempt to improve a poorly performing execution plan produced by the database query optimizer is a critical aspect of database performance tuning [7]. Identifying and fixing the poor SQL performances can help to solve the performance problem as stated by Darmawan et al. [14]. Therefore, there is a need to fix problematic SQL queries by understanding the problem and re-writing them. It is expected that, queries that undergone the 're-write' process will retrieve accurate results within an acceptable time rate given tight CPU or IO constraints within MES environment.

B. Significance of Study

The result of this research will be of benefit in the industries with complex system architecture (such as MES) and where poor database performance is intolerable. By improving database performance through SQL queries rewriting, we can minimize the 'cost' that must be paid by those industries due to poor SQL queries. Specifically, the techniques proposed should support Silterra (a manufacturing industry that hosts complex FAB) to improve its MES reporting function. In the next section, literature survey on the techniques proposed to improve database performance will be presented.

II. LITERATURE REVIEW

There are several factors that affect database performance such as database settings, indexes, memory, CPU, disk speed, database design, and application design [8]. Several methods have been used to improve the database performance as shown in Fig. 2. Most of the methods are performed by DBAs while the remaining ones are performed by application developers. DBAs need to perform a regular inspection and database performance diagnosis. This process identifies the top ten activities that cause the problem with the detail information. DBAs need to know exactly what are the causing performance problems before they can apply a corrective method.

One major factor that can degrade database performance is SQL queries. SQL queries can be complex enough to run and they rely heavily on optimization techniques to response in timely manner [15]. Given the indicator of poor database performance (such as slow query responses), the task of database performance tuning involves deciding (and selecting) which method works best to resolve the problematic SQL queries.

Queries (in form of SQL statements) are embedded in the application source codes to enable interaction with the databases in retrieving data sets of interest. During the application development, developers usually focus on writing SQL queries to retrieve correct set of data from the database, without considering how the queries may affect database performance. Consideration must be made on how well the queries behave during the actual implementation of the applications as well especially when the database grows bigger in the future.

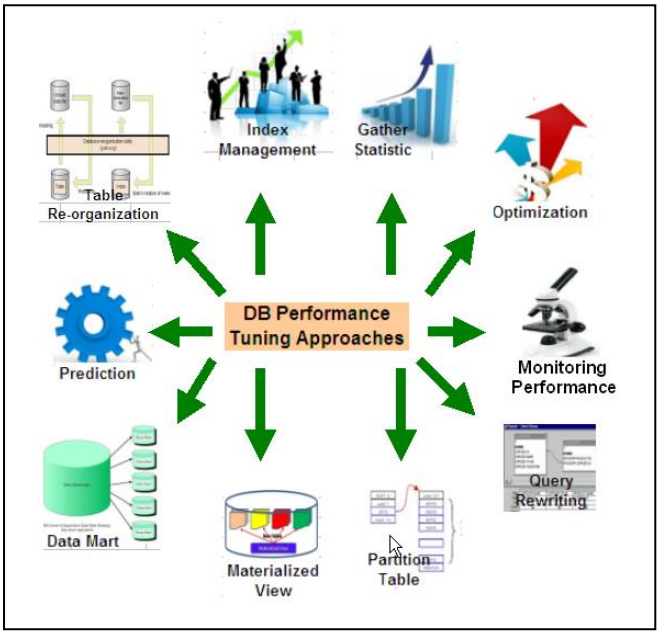

Fig. 2. Methods of Database Performance Tuning

In the literature, several methods have been proposed to deal with poor database performance caused by SQL queries. Firstly, query optimization is applied to support decision on the most efficient way to execute SQL statements [21]. In general, database optimization involves calculating the best possible utilization of resources needed to achieve a desired result such as minimizing process time without impacting the performance of any other system resource. It can be accomplished in three levels whether hardware/operational, database and or application (by optimizing source code or SQL queries [9].

Within DBMS tuning hierarchical level as shown in Fig. 3, SQL statement tuning is one critical aspect of database performance tuning [7]. Application developer will start with this level of tuning with the helps from the DBA. Application level can impose stronger influence the on database as compared to other levels in the hierarchy [6,9].

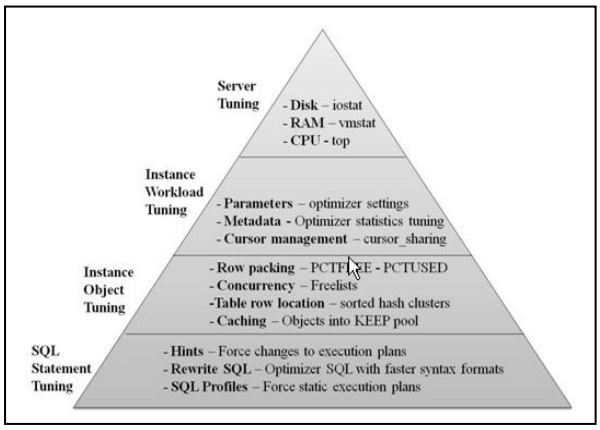

Fig. 3. DBMS tuning hierarchical level [10]

The second method is performing table partition to improve the database performance. In this method, tables are split into smaller parts for easier data maintenance and faster data retrieval in order to overall manageability of database system [20]. Table partitioning contributes to better performance where data selection and data purging can be done easily. However, this method causes maintenance overhead for DBA in the case involving large amount of partitioned tables. In addition index creation will be very slow if hash partition is used.

The third method is to perform monitoring performance to determine the possible problems, to locate the root causes and to provide recommendations for correcting them. Subsequently, problematic SQLs need to be identified for rewriting to improve performance.

SQL is a declarative language, which can be written in many forms to get the correct query results, however it produce different execution plan and response times. SQL query rewriting consists of the compilation of an ontological query into an equivalent query against the underlying relational database [23]. This process will improve the way data being selected and able to improve the performance drastically. However it can be a troublesome job to change hardcoded queries. Furthermore, queries that are not tested thoroughly could cause delay.

For the rest of the methods, please refer to Table 1, where we define each of the methods based on literature. Furthermore we discuss the advantages and limitations performing this method in database performance tuning.

The work presented above highlighted the need to deal with problematic SQL queries in order to improve database performance. However, several questions as listed below are open problems especially in real-time environment that worth answering:

1) How to identify problematic SQL?
2) How to rewrite the identified problematic SQL?
3) How to measure the results of query re-writing?

Motivated by the crucial need to improve database performance that hosts big data in the manufacturing domain, we therefore set to answer the above questions in this research.

TABLE 1: LIST OF METHODS FOR DATABASE TUNING

Methods	Literature	Advantages	Limitation
Gather statistic	For Oracle DB. It relies on up to date statistic to generate the best execution plan[19].	Updated statistics helps optimizer to select perfect execution plan for a query. For Oracle DB. It relies on up to date statistic to generate the best execution plan	Can be resource consuming. Must plan accordingly before executing
Index Management	Indexes are optional structures associated with tables and clusters that allow SQL statements to execute more quickly against a table [21].	Index created columns helps queries to select using index instead of doing full table scan, which is usually expensive.	DML statements can be slow if there is a lot of indexes on the table
Table Re organization	to improve the performance of queries or data manipulation language (DML) operations performed against these tables[22].	All data blocks will be moved to be together and prevent fragmentation which can cause slowness	Usually time consuming and needs downtime
Prediction	Estimating the Space Use of a Table, estimating the Space Use of an Index and Obtaining Object Growth Trends [21].	Able to predict the problem before it happened	Sometimes the prediction are not accurate because the data consumed is not increase in sequence
Data Mart	A data warehouse that is designed for data to be collected directly from various sources [21],	The database will be grouped according to schemas and departments. Easy to do maintenance and will improve the database performance.	Only applicable to schemas which are less than 100GB. Its not optimum to use data mart for bigger database.
Materialized view	A materialized view provides access to table data by storing the results of a query in a separate schema object [21].	Fast synchronization between source and target. Data can be refreshed on preferred method.	Complex queries on MV tables perform badly especially if there are joins with other tables
Partition Table	Splits the table into smaller parts that can be accessed, stored and maintained independent of one another [20].	1. Improve performance when selecting data 2. Easily for data pruning.	Hard to do maintenance for DBA's on partitioned table if it involves lots of partitions in a table. Index creation will be very slow if using hash partition
Query Rewriting	Query rewriting consists of the compilation of an ontological query into an equivalent query against the underlying relational database [9,10,23].	Improve the way data are being selected. By adding hints in the SQL, sometimes enhance the performance of individual queries.	Can be a troublesome job to change hardcoded queries. Queries that are not tested thoroughly could cause slowness.
Monitoring Performance	To determine possible problems, locates the root causes and provides recommendations for correcting them [21].	Able to identify the root cause of the problem.	Only DBA able to do monitoring.
Optimization	Query Optimization is the process of choosing the most efficient way to execute a SQL statement [9,10,21].	Helps queries to run faster. Data retrieval can be improved. Parameter tuning and memory tuning enable the database to perform in optimum level.	Must be done with supervision and wrong parameter set can cause database to go down or perform badly. Memory leak is also possible.

III. AIM OF INVESTIFATION

The objectives set out for this research are:

1) To propose a technique of SQL query rewriting for database performance tuning in real-time environment.

2) To develop an algorithm of real-time SQL query re-writing for database performance tuning.

3) To determine the technical issues those are barriers to implement successful real-time query rewriting.

IV. RESEARCH METHODOLOGY

SQL query tuning involves lots of understanding in database technology, DBMS, application design, SQL query and new knowledge in this field. Tuning task need high level of expertise to manage it, time consuming and challenging [7,18].

Herodotou and Babu stated that, generally DBA tuning steps start with collecting monitoring data to diagnose the problem. Secondly based on the observation, need to form hypotheses on the potential fixes to the problem. Then continue to execute the plan to refine and confirm the hypothesis. After that analyse the result and if not satisfied need to do trial and error and repeat the process [7]. They propose a technique to automate experiment driven SQL tuning as shown in Fig. 4.

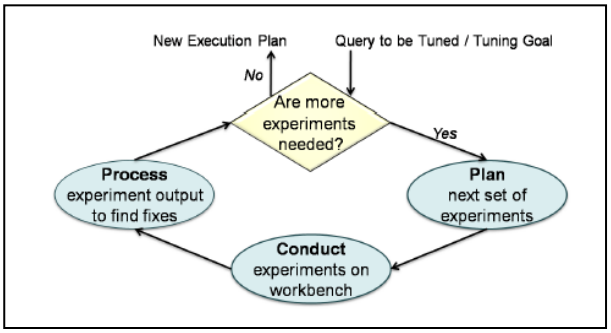

Fig. 4. Experiment Driven Workflow that abstract SQL tuning [7]

Bin Cao in his research used explain plan to evaluate the queries and perform performance analysis among alternative query plan [16]. Meanwhile, Bruno and Chaudhuri used experiment based method as shown in Fig. 5 to develop an 'alerter' that helps determining when a physical design tool should be invoked [17]. The process can be categorized in three stages that are monitoring, diagnose and recommendation.

Based on the literature and observation, we adopt experimental-based research method in order to achieve the objectives set for this research. By definition, experimental-based method component are participants, element and factors, instrument and material used, procedures and measures effectiveness, statistical analysis and experimenter [12].

We design the stages of this research using experimental approach as shown in Fig. 6. The description of each stage is depicted in Table 2.

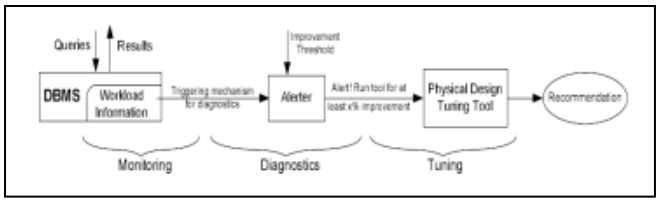

Fig. 5. Experiment based flow for physical design [17]

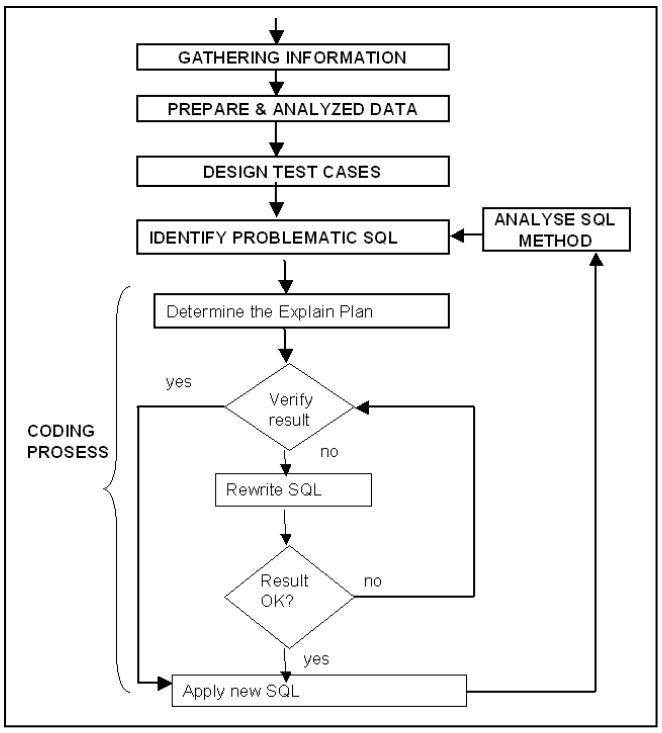

Fig. 6. The Proposed Experimental-based Research Stages

TABLE 2: EXPERIMENTAL-BASED RESEARCH STAGES DESCRIPTION

Stage	Description
Gathering Information	Consists of : Literature Review; Identify the cause of problematic SQL; Identify the tuning process; study the implication between the tuning process.
Prepare & Analyze data	Collect monitoring data to diagnose the problem..
Design Test Cases	Create test cases for different type of problem.
Identify problematic SQL	Script will be developed to collect the SQL execution statistics and to get the list of low efficiency SQL.
Coding Process	Involve steps as below: 1.To determine existing execution plan and understand the problem. 2. Verify the result and hypothesis potential fix to the problem 3. Rewrite the SQL and test The process will be repeated until get the optimum result.
Analysis SQL Method	The result will be analyzed either to improve the resource allocation or execution plan. It will be measured to analyze the relative performance.

In the proposed methodology the first stage is gathering information, which consists of literature review; identify the cause of SQL poor performance problem; identify the SQL tuning process, which can improve the database performance and study the implication between the tuning processes. In this stage we also will study the new knowledge and discovery in this field.

Secondly, the process to set up the experiment environment with large database size to evaluate queries. Then, continue by the process of collecting the monitoring data to diagnose the problem. The collected data need to be analyse and need to understand the plan selected by the optimizer and the reason behind it.

Designing test cases follows this. According to Degaville, SQL tuning involved identifying high load or top SQL statement, attempt to improve execution plan produced by optimizer and finally implement corrective action [17]. In this stage we also need to illustrate the SQL queries and come out with different types of test case for the experiment. These three steps are repeated until the overall system performance reaches the satisfactory level or no more statement can be tuned.

In the forth stage, a script will be developed to collect the SQL execution statistics into the database repository in order to pinpoint problematic SQL. The scripts need to be run at the application peak time to identify low efficiency SQL and which SQL use high usage resource and low.

Based on the statistic reports, the SQL will be tuned by rewriting the SQL or by modifying the database design. In then coding process, it involves three steps that need to be followed and repeated until get the satisfied result. Firstly, need to determine existing execution plan, analyse and understand the problem. Secondly, verify the result. Based on the results and observation, hypothesis potential solution to the problem. Thirdly, rewrite the SQL based on the hypothesis and test the SQL. The process will be repeated until we get the optimum result and confirm the hypothesis.

Finally, the result will be analyzed either to improve the resource allocation or execution plan. The new SQL, need to be tested in the real environment to check the validity and recognize the limitation.

To validate the usefulness of the proposed technique, real datasets collected from Silterra MES will be used and the experiment will be implemented within the system architecture and database hosted of the company.

V. Expected Outcomes

This research will contribute to a new way of large database performance tuning using real-time query re-writing technique within manufacturing industry. Fig. 7 shows the current DBMS pool environment where a lot query transactions go into (and out from) DBMS at the same time. This leads to overhead to the DBMS and DBAs. We would like to reduce the current traffic towards DBMS pool as the traffic shown in Fig. 8. A novel query rewriting technique will be applied to this environment, which is expected to provide a better satisfactory level system performance relative to the current scenario.

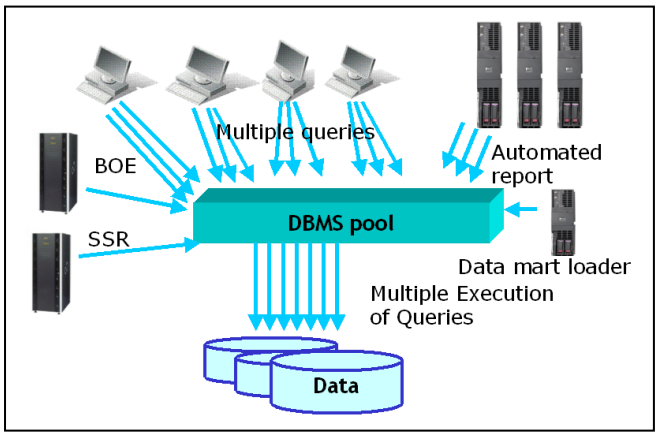

Fig. 7. Current DBMS pool environment

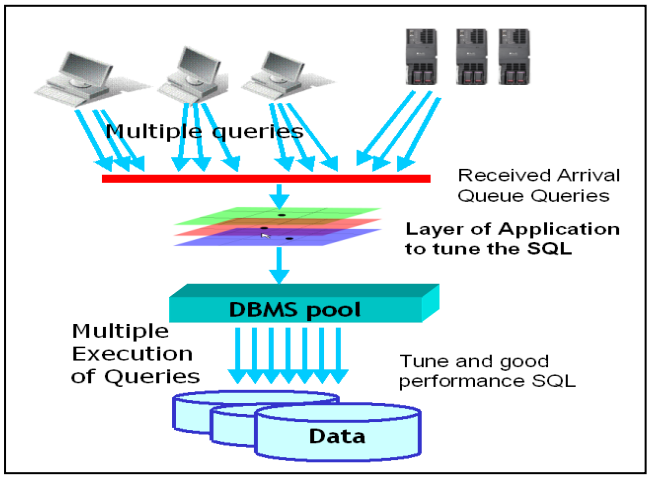

Fig. 8. Expected DBMS pool environment

VI. Summary

Motivated by database performance problem caused by SQL queries and the gaps identified in the literature in this domain, we proposed a novel technique for SQL query rewriting. In general, the result of this research will contribute to the database community as well as big data users where database performance is an issue. The usefulness of the technique proposed will be validated in a real manufacturing industry.

References

[1] H.Zhou, Z.Yao and H.Xiao, "The Design Research on the Storage Structure for Large Amount of Data of the Database, Proceeding of the International Conference on Advance Mechatronic System," Zhengzhao. China,2011, pp384-388.

[2] Adam Jacobs, The pathologies of big data, Communications of the ACM, v.52 n.8, August 2009

[3] SA Rahim, I.Ahmad and MA Chik, "Technique to Improve Visibility for Cycle Time Improvement in Semiconductor Manufacturing,Proceeding of IEEE International Conference on Semiconductor Electronics (ICSE), 2012 , pp. 627-630.

[4] Puvaneswaran Balakrishna, MA Chik, Ibrahim Ahmad and Bashir Mohamad, "Throughput Improvement in Semiconductor Fabrication for 0.13µm Technology," Kota Kinabalu.Malaysia, 2011.

[5] P. Balakrishna, MA Chik, I.Ahmad and B. Mohamad, "Throughput Improvement in Semiconductor Fabrication for 0.13µm Technology," in RSM2011 Proceeding.,Kota Kinabalu, Malaysia, 2011

[6] Beeler, D., Rodriguez, I., Optimizing your Database Performance, BMC Software ,2002

[7] H. Herodotou and S. Babu. Automated SQL Tuning through Trial and (Sometimes) Error. In Proc. of DBTest '09. ACM, 2009.

[8] Connolly, T.M.,& Begg, C. E, SQL: Data Manipulation in A.D. McGettrick (3rd Ed). Database Systems: A Practical Approach to Design, Implementation and Management, . (pp 110-155). Harlow: Addison Wesley, 2002

[9] Leccotech, Microsoft T-SQL Performance Tuning Part 3: Query Optimization Strategies. Quest Software Inc ,2005.

[10] Trezzo and Joseph C, Oracle PL/SQL Tips & Technique, Osborne:MsGraw-Hill, 1999.

[11] DK. Burleson, "Advanced Oracle SQL Tuning-The Definitive Reference, Oracle in Focus" 2nd ed,, Rampant, 2010

[12] M.Rao and S.Shah, Experimentation A Research Methodology, 2002.

[13] T.Ponsignon and L.Mönch, "Architecture for Simulation-Based Performance Assessment of Planning Approaches in Semiconductor Manufacturing," Neubiberg. Germany,2010.

[14] B.Darmawan, G.Groenewald, A.Irving, S.Henrique and M.Snedeker, Database Performance Tuning on AIX, IBM International Technical Support Organization, 2003, pp.291.

[15] Li, Dandan, Lu Han, and Yi Ding. "SQL Query Optimization Methods of Relational Database System." Computer Engineering and Applications (ICCEA), 2010 Second International Conference on. Vol. 1. IEEE, 2010.

[16] Cao, Bin. "Optimization of complex nested queries in relational databases." Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on. IEEE, 2006.

[17] Bruno, Nicolas, and Surajit Chaudhuri. "To tune or not to tune?: a lightweight physical design alerter." Proceedings of the 32nd international conference on Very large data bases. VLDB Endowment, 2006.

[18] Ling Hu , Kenneth A. Ross , Yuan-Chi Chang , Christian A. Lang , Donghui Zhang, QueryScope: visualizing queries for repeatable database tuning, Proceedings of the VLDB Endowment, v.1 n.2,2008

[19] Cao, Wei, and Dennis Shasha. "Tuning in action." Proceedings of the 16th International Conference on Extending Database Technology. ACM, 2013.

[20] Herodotou, Herodotos, Nedyalko Borisov, and Shivnath Babu. "Query optimization techniques for partitioned tables." Proceedings of the 2011 ACM SIGMOD International Conference on Management of data. ACM, 2011.

[21] Oracle, Oracle Database Administrator's Guide 11g Release 1 (11.1), 2013.

[22] L.Hobbs and P.Tsien, Oracle Database 10g Release 2 Online Data Reorganization & Redefinition, 2005.

[23] G. Gottlob, G. Orsi, and A. Pieris. Ontological queries: Rewriting and optimization. Data Engineering ICDE, 2011 IEEE 27th International Conference, Hannover.

Test Case Generator for Embedded Real Time Software

Nabilah Kamarul Bahrin and Radziah Mohamad
Faculty of Computing
Universiti Teknologi Malaysia
81310 UTM Skudai, Johor, Malaysia.
nabilahkb@gmail.com, radziahm@utm.my

***Abstract*— Real Time Embedded System (RTES) is widely applied nowadays, moreover in security and highly demand area. Real Time Embedded systems need an assurance in software quality and for that; software testing is nowadays getting more attention in order to ensure the quality of software. Currently, with wide practice of embedded software in the field that claims high reliability and security, it makes the software testing facing increasingly severe challenges. Unified Modeling Language (UML) is widely applied and emerging in software testing area. Lately, the importance of UML models in designing test cases has been well recognized. By using UML models, it enabled the testing to be run in early phase which is also can help can help developer to check software failure at early stage, which can save time and money.**

Keywords- Software Testing; Real Time Embedded System; Sequence Diagram; Test Case Generator

I. INTRODUCTION

Real Time Embedded System (RTES) is widely applied nowadays, moreover in security and highly demand area, such as in engineering field (B. Zhang and X. Shen, 2011). This type of software is extremely important and a bit different from normal usage type of software which it is also classifies as a safety critical software (P. Krishnan *et.al*, 2012). Without realizing it, we actually use and need embedded system in our daily life. Cell phones, washing machine, toaster, radio or even our vehicle, are using embedded system. RTES are also widely used in entertainment devices, vehicle and household applicants (G. Karsai, 2010).

Software testing is one of the best methods in order to prove the correctness and check system validity especially to safety critical system, like embedded system. Testing has its own procedure and process activities to be followed for example test planning and control, test analysis and design, test implementation and execution, evaluating exit criteria and reporting and last is test closure activities. A system must be tested effectively and comprehensively before giving it to the user (M. Malekzadeh and R.N. Ainon, 2010). By running software testing, developer can ensure the efficiency and the correctness of the system (Shanthi & Mohankumar, 2012). A minor failure of the system may put the safety of human in risk and also can lead to the major disaster. It is very important to run software testing thoroughly before real world implementation. Therefore, systematic and correctness of the chosen technique is very important and highly necessary to ensure software quality and assurance (L. Wang, 2004).

Nowadays, researcher are interested in studying in test case generators using a various techniques reading from UML (Unified Modeling Language), for example, use case, state diagram and sequence diagram. With test case generator, testing can be done in earlier phase which may help developer to test the system earlier before build and implement the system. Thus, developer can avoid realizing the bug or faults later, which is more risk and involve more money and demanding time to repair the bug or faults. UML currently widely accepted in system development and it continuously evolves in order to satisfy the user (M. Popovic and I. Velikic, 2006). UML shows achievement in test case generation for example test case generations based on state diagram which it allow test cases to be available in earlier phase in software development cycle (P.samuel , 2008). Thus, by getting test cases earlier, testing the system will become more affective. Currently, most of the method that are being suggested by the researcher will produce a number of test cases that sometimes will not be required to test the

software. However, as mentioned M.Khandai et al. (2011), they are planning to apply test case prioritization technique which can reduce cost and time in order to generate the effective and quality of software test cases.

II. OBJECTIVE

Based from the problem being state above, the objectives of this research are situated as the following:

1) To investigate the features of existing test case generation

2) To propose test case generation framework.

3) To validate propose test case generation framework.

III. RELATED STUDIES

Based on R. Mitsching (2011), testing is crucial moreover in the world of safety critical system, which reflect to RTES. J. Grossmann (2009) also said that in order to cope with RTES requirement and complexities of the system, the developer was forced to ensure the quality of the system in functional or even in non-functional requirements of the system. In addition, when we involve in RTES, nonfunctional requirements are also play an important role, especially in time related input-output behavior (J. Grossmann, 2009). Currently, researcher are concerning in doing automatic test case generation based on UML. Report by (Swain, 2010), UML models are one of the important resource for test case design and thus can reduce testing cost towards improving software quality. Automatic test generation that based on UML received a lot of attention from researcher. For example, in (Sarma, 2007) work, they use sequence diagram as a source of test case generation. From the sequence diagram, they transform it into a graphical representation, called sequence diagram graph where the information of test case generation is stored in each node. Different from work done by (Nebut *et al.* , 2006) , they used use case diagram to generate test cases. There are also a researcher who combined a few model from UML, to generate test cases, for example (Biswal *et al.*, 2008) who is using the combination of activity diagram, sequence diagram and class diagram to generate test cases.

IV. RESEARCH METHODLOGY

Table I show the process of the research and expected output from each objective. Currently author in the 2nd objective, which is developing transformation diagram from sequence diagram to petri nets model. Figure1 show the framework of the research which relate to the objectives and the whole research process.

TABLE I. Research Process

Research Objectives	Research Questions	Research Methods	Research Output
1. To investigate the features of existing test case generation.	What are the methods/ techniques that currently used to perform test case generator?	Study the previous researchers and figure out any limitation / problem of testing.	-UML (Sequence diagram) as an input for test cases. - Current test case generation has limitation on prioritization
2. To propose test case generation	How can we generate test cases using proposed enhancement of	Run experiment using selected case study, to the tools that has been choose,	-UML model, transform to Petri Nets (Formal Method)

framework.	the framework?	with the proposed method /techniques.	-From Petri Nets Model, generate test cases, and apply prioritization rule.
3. To validate propose test case generation framework.	How does the proposed method /techniques can be validate and verify the effectiveness of test case generator?	Not yet identified	Validation and verification report towards proposed method / techniques of test case generator.

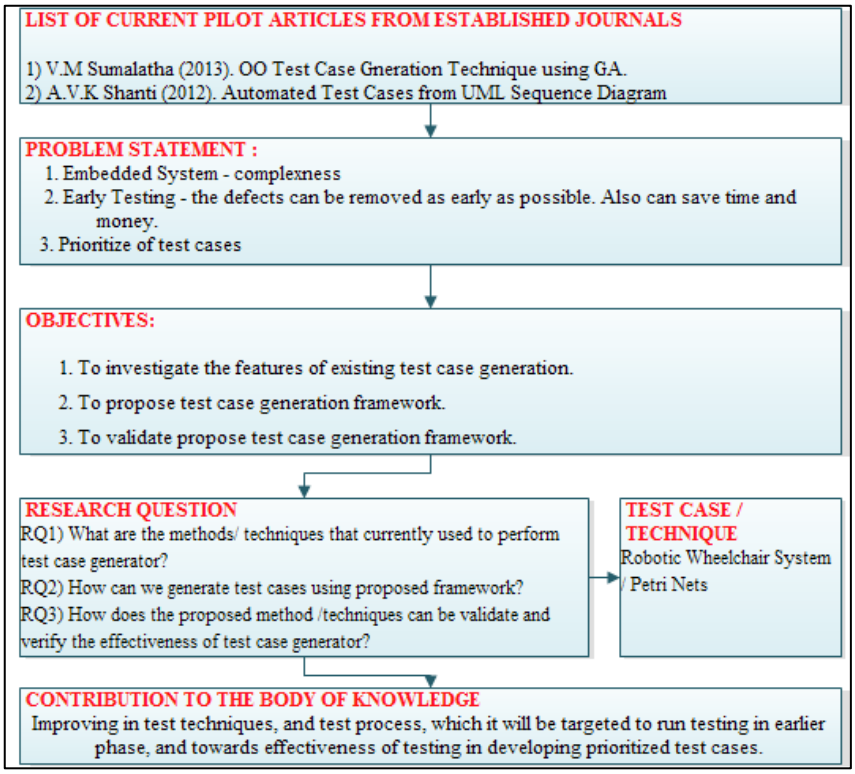

Figure 1: Framework Research

V. EXPECTED RESULT

Our study is to proposed a framework for generate test cases from embedded system and also to prioritize the test cases. Expected result is the model of transforming from sequence diagram to petri nets, and then extract it to test cases. From the element of MARTE, we will prioritize the test cases.

VI. CONCLUSION AND FUTURE WORKS

Currently, there is a lacking on current work on the limitation of showing how they can prove the test cases that they auto generate can be used effectively. This is because; there are issues that sometimes current generated test cases are not relevant for the software testing. From the studies also shows that, not all of the test case generation can be applied to RTES, as the system has their own criteria in order to obtain efficiency in the system. Our future research is to find the best method that can be effectively applied to RTES.

ACKNOWLEDGEMENT

We would like to thank Ministry of Education Malaysia (MOE) for sponsoring the research through the FRGS grant with vote number 4F165 and Universiti Teknologi Malaysia for providing the facilities and support for the research. In

addition, we would like to extend the gratitude to our lab members in EReTSEL Lab of Faculty of Computing, Universiti Teknologi Malaysia for their ideas and support throughout this study.

REFERENCES

B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering – A systematic literature review," Information and Software Technology, vol. 51, no. 1, pp. 7-15, Jan. 2009.

B. Zhang and X. Shen, "The effectiveness of real time embedded software testing," Reliability, Maintainability and Safety (ICRMS), 2011 9th International Conference, China: pp.661-664, 2011.

Biswal, B. N., Nanda, P., & Mohapatra, D. P. (2008). A Novel Approach for Scenario-Based Test Case Generation. *2008 International Conference on Information Technology*, 244–247. doi:10.1109/ICIT.2008.43

C. Nebut, , Fleurey, F., Le Traon, Y., & Jezequel, J.-M. (2006). Automatic test generation: a use case driven approach. *IEEE Transactions on Software Engineering*, *32*(3), 140–155. doi:10.1109/TSE.2006.22

C.-S. Koong, Lai, H.-J., Chang, C.-H., Chu, W. C., Hsueh, N.-L., Hsiung, P.-A., Shih, C., et al. (2010). Supporting Tool for Embedded Software Testing. 2010 10th International Conference on Quality Software, 481–487. doi:10.1109/QSIC.2010.34

C. P. Vudatha, S. Nalliboena, S. K. Jammalamadaka, Duvvuri, B. K. K., & L. S. S. Reddy, (2011). Automated generation of test cases from output domain of an embedded system using Genetic algorithms. 2011 3rd International Conference on Electronics Computer Technology, 216–220. doi:10.1109/ICECTECH.2011.5941989

G. Karsai, "Evolving Embedded Systems," Vanderbilt Univ., Nashville, TN, USA, vol.43, pp. 34-40, 2010.

H.Qin, C. Zheng "A Embedded Software Testing Process Model," International Conference on Computational Intelligence and Software Engineering, 2009. CiSE 2009, p.p 1-5, 2009.

J. Grenning, "Applying test driven development to embedded software" , pp. 20-25, vol.10, issue.6, (2007).

J. Grossmann, D. Serbanescu, & I. Schieferdecker, (2009). Testing Embedded Real Time Systems with TTCN-3. 2009 International Conference on Software Testing Verification and Validation, 81–90. doi:10.1109/ICST.2009.37

K. A. Sedek, S. Sulaiman, and M. A. Omar, "A systematic literature review of interoperable architecture for e-government portals" Proceeding of the 5th Malaysian Conference in Software Engineering (MySEC) 2011, pp. 82-87, 2011.

K. G. Larsen, & M. Mikucionis, (2005). "Testing real-time embedded software using UPPAAL-TRON: an industrial case study on Embedded software" Retrieved from http://dl.acm.org/citation.cfm?id=1086283

L. Wang, "Issues on software testing for safety-critical real-time automation systems," The 23rd Digital Avionics Systems Conference 2004 (DASC 04), vol.2, 2004.

M.J. Karlesky, W.I Bereza, C.B. Erickson, "Effective Test Driven Development for Embedded Software", Electro/information Technology, 2006 IEEE International Conference, pp. 382-387, 2006.

M. Khandai, "A novel approach of test case generation for concurrent systems using UML Sequence Diagram," Electronics Computer Technology (ICECT), 2011 3rd International Conference, vol.1, pp.157 – 161, 2011.

M. Malekzadeh , R.N. Ainon, "An automatic test case generator for testing safety-critical software systems," Conference on Computer and Automation Engineering (ICCAE), 2010 The 2nd International, pp. 163- 167, 2010.

M. Popovic, I. Velikic, "A generic model-based test case generator," Conference on 12th IEEE International Conference Engineering of Computer-Based Systems, 2005 (ECBS '05), pp. 221-228, 2005.

P. Iyenghar, (2011). Test Framework Generation for Model-Based Testing in Embedded Systems. 2011 37th EUROMICRO Conference on Software Engineering and Advanced Applications, 267–274. doi:10.1109/SEAA.2011.48

P. Krishnan , R. Venkatesh, P. Bokil, T. Muske, and V. Suman, "Effectiveness of Random Testing Embedded System," Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, pp. 5556–5563, 2012.

P. Samuel, Mall, R., & Bothra, A. . (2008). Automatic test case generation using unified modeling language (UML) state diagrams, *2*(2), 79–93. doi:10.1049/iet-sen

R. Mitsching , C. Weise, D. Franke, T. Gerlitz, & S. Kowalewski, (2011). Coping with Complexity of Testing Models for Real-Time Embedded Systems. 2011 Fifth International Conference on Secure Software Integration and Reliability Improvement - Companion, 128–135. doi:10.1109/SSIRI-C.2011.27

Sarma, M. (2007). Automatic Test Case Generation from UML Sequence Diagrams, 60–65. doi:10.1109/ADCOM.2007.68

Shanthi, A. V. K., & Mohankumar, G. (2012). A Heuristic Approach for Automated Test Cases Generation from Sequence Diagram using Tabu Search Algorithm, *85*(4), 534–540.

Swain, S. K., Mohapatra, D. P., & Mall, R. (2010). Test Case Generation Based on Use case and Sequence Diagram, (1), 21–52.

W.Yichen & W.Yikun. (2011). Model-Based Simulation Testing for Embedded Software. 2011 Third International Conference on Communications and Mobile Computing, 103–109. doi:10.1109/CMC.2011.132

Y.Yin, B.Liu, "A Method of Test Case Automatic Generation for Embedded Software," International Conference on Information Engineering and Computer Science, 2009. ICIECS 2009, pp.1-5, 2009.

Y. Yin, Z. Li, & B. Liu, (2010). Real-time Embedded Software Test Case Generation Based on Time-extended EFSM: A Case Study. 2010 WASE International Conference on Information Engineering, 2, 272–275. doi:10.1109/ICIE.2010.160

Y. Yongfeng, & L. Bin, (2009). "Test Cases Generation for Embedded Real-time Software Based on Extended UML Technology and Computer" doi:10.1109/ITCS.2009.23

A Study on Code Peer Review Process Monitoring using Statistical Process Control

Alhassan Muhammad Abubakar
Faculty of Computing,
Universiti Teknologi Malaysia,
81310, Skudai, Johor
Email: mohanajeeb66@yahoo.com

Dayang N. A. Jawawi
Department of Software Engineering,
Faculty of Computing,
Universiti Teknologi Malaysia,
81310, Skudai, Johor
Email: dayang@utm.my

***Abstract*— Software development process (SDP) and Software products are like two sides of a coin. We cannot achieve one without another. Today, in our software industries, monitoring software process is very challenging. Many problems of software process monitoring are hampering the quality of our software products. In this paper, we plan to address the problem of process instability and causes of process anomalies. These problems can be addressed using one of the powerful statistical techniques known as statistical process control (SPC). Also, control chart would be used in our study as it has been proved to be one of the suitable tools of SPC in monitoring process stability. As we know, the more defects we found during SDP, the less quality of the software product. Therefore, this study considers defect density as the metric to be use due to its significance in determining product quality.**

***Key words*— Software development process, Statistical process control, Control Charts, Defect density.**

I. INTRODUCTION

Software development process has been defined by many authors within software engineering domain. These definitions are the same in respective of the type software product, architecture of the products as well as the tools used in developing the product. Basically, software development process involves the arrangement of people (mostly software engineers) with related facilities or tools for development of good quality software product. Due to the high intensive or involvement of people in SDP, it is possible to have so many factors that can contribute to the process instability. For example, lack of proper arrangement of facilities can have a great impact on the process behavior.

According to the authors in [2], causes variations are inevitable in any process. These variations may be either as a result of common cause's variations or special (assignable) cause's variations. The common causes may be as a result of common interactions between human and material resources that are involved in the process. In other words, the normal variation of process that exists because of normal interactions among different components (people, machines, methods and so 'on) of the process is referred to as common cause variation and sometimes is also called controlled variations. Assignable causes of variations on the other hand, may be as a result of events that are not included in the process. This event changes the process behavior, have significant impact on the process performance and the quality of the product. Also, this type of variation is referred to as uncontrolled variations [9]. Since common causes of variations are also referred to as controlled variations, then identifying and eliminating uncontrolled variations (assignable causes of variations) is one of the successful paths towards achieving process control and stability. The technique that we can use to achieve control and stable process is called statistical process control.

The idea of SPC was initiated by Shewart in 1920's [2]. During that time, he was working or conducting research in Bell Telephone Laboratories on how to improve process quality and lower costs. This leads to the founding of SPC. Based on his studies, SPC can be used to monitor and control process using various statistical tools. Among these tools are: pareto charts, histogram, control charts and so 'on. This study focus on control charts as it has proved to be effective statistical tool for detecting out-of-control situation. That is to say, many researchers in the field of software engineering performed several studies with this statistical process control (SPC) technique and proved to be effective for process control and improvements. In line with this, this study would be additional effort by contributing with the statistical idea of ensuring software process stability.

However, the remainder of this paper is organised as follows; section 2 contains the motivation of our study. Section 3 presents related work. In section 4; we describe the problem statement of this study. Section 5 describes our research methodology. Section 6 presents our initial result. In section 7; we discussed about the problem we encountered in this study and section 8 contains summary and future work.

II. MOTIVATIONS

Today, the world is changing to a digital world. Many activities of our daily life are carried out with support of information technology or software systems. In our global

world, various sectors contribute their quota towards improving our economy and other sectors. Therefore, software engineering sector should be considered in terms of economy improvements, educational improvements and so ‘on.

Basically, appreciating the status and contributions of software systems to support human activities is one of the motivations of this study. To support and improve the use of software systems in all our activities, we can contribute significantly by monitoring SDP. This will enable software engineers to investigate process behavior over time and quickly react to any detected process anomalies so as to enhance the quality of their products. However, many researchers share their studies on using SPC in the field of software engineering community.

III. RELATED WORK

Software organizations use standard models of process improvement purposely to improve their software development process, and to be sure that certain level of quality has been achieved during the development of the software [1], [6], [11]. This will enhance the product quality and bridge the gap between customer and software engineers. Example of these models is capability maturity model (CMM) which was later superseded by capability maturity model integration (CMMI). This study focused on CMM because, CMM is more specific on process assessment and it implicitly directs software organizations (at CMM level/stage 4) to use quantitative techniques such as statistical process control to quantitatively managed their software development process.

In line with this, many researchers share their idea or experiences on using SPC for software process improvement. Recently, in the studies of [6], control chart of SPC were used to improved process performance of software development life cycle. During their study, they focused on requirement process and they visualized how statistical process control can be effectively and efficiently applied to the processes in order to achieve good quality software products. Additional effort was done by researchers in [11], [2], [16], they emphasized that in order to increase software process efficiency and effectiveness so as to have good quality product; we need to enhance our process measurements skills. Good measurements are suitable for SPC implementations and enable us to evaluate process performance.

Moreover, in the work of [22], they use SPC to monitor software reliability. But, monitoring software reliability is not a very simple activity. These authors make a good effort in using SPC on inspection and review process so as to improve software reliability and quality. The authors used control chart which is one of the important tools of SPC to monitor the time between failures, when the time between failures fall outside control limits specified by these authors, then causes of variations may be presents and need to be addressed so as to ensure software reliability. Similarly, in the study of [1], SPC can be use for continuous software process improvement. He proposes an SPC-based approach that implements SPC technique, and used it for software process improvements. Also, the author in [19]; discovered his experience of using SPC in software development process. He conducted a case study on Hitachi Software Engineering (HSE) in Japan. After he conducted several use of control charts to investigate HSE processes e.g peer review process, he discovered many process anomalies through the use of control charts. He concluded that SPC is a very useful quantitative tool that can be use in software development process.

Similar work was done by researchers in [15], [18], [16]; they used control charts to monitor software quality evolutions for defects. They used the chart to find out the quality evolution of Eclipse and Gnome systems. They measure software quality by of these two open source software systems by number of defects and use control charts on changes in defect over time. They clearly found process deviations of these two systems through the use of control charts and concluded that it very challenging to keep software quality under control during evolutions. But on the other hand, they believed that when control charts are effectively used, software organizations can monitor quality evolutions over a long period of time.

Furthermore, recently some researchers [8],[22], compare Modified Maximum Likelihood Estimation (MMLE) and half logistic Distribution estimates using SPC in order to find the control limit for assessing software reliability. However, because software reliability measure is a one of the important of software quality attributes, these authors concluded that either exponential distribution or HLD is preferable for the estimations of data that is in the form of inter failure times and SPC is a very powerful quantitative technique that can be used to control the estimates in order to ensure software reliability.

Furthermore, another contribution was made by authors in [20] in which they used SPC to automatically detect performance regression. Performance regression is very important in software development process but often consume time. These authors proposed an approach (control charts) to analyze performance counters across test runs using statistical process control technique of control charts. The result they obtained at the end clearly verifies that the proposed approach can be used to identify performance regression in software systems. Similarly, additional work has been done another researcher [9]; He presents practical application of statistical process control. He found that SPC helped him understand and predict product quality and the development process controlling the quality. He also added that the technique (SPC) provided him with the hard data he needed to justify his result at the end.

Moreover, researchers in [3]; added that statistical and process thinking lead to the use of SPC in order to determine stability of processes used to develop Software. However, because software process improvement is a continuous activity, controlling and monitoring software process is still an issue within software engineering domain.

IV. PROBLEM STATEMENT

Based on the studies we have seen so far, most of them addressed issues of process monitoring using different control charts of SPC. But on the other hand, some of the studies focused on the interpretations of SPC instead of proper monitoring of the software process. Monitoring software process stability remain an open issue that when properly handle, we can enhance the quality of the software product. However, there are some issues such as; process performance deviation that affect software process, by carefully monitoring these issues; we can improve the stability and capability of the process. Therefore, in this section, we describe some problems which we plan to address at the end of our study. These are as follows;

1. Problem of detecting software process deviations as a result of variations.
2. Problem of investigating the causes of variations in software process.

In order to address the above mentioned problems, some silent questions should be considered.

a. Are CMM lower maturity level software industries matured enough to use or apply SPC for process monitoring and control?
b. How can we achieve successful SPC implementation at the CMM lower maturity level organizations?

To responds to these questions effectively, our studies will go further to conduct a case studies at CMM lower maturity level industry. Also, as we seen from the literature in the previous section, SPC was used by many of the researchers and proved to be effective statistical method of process control. Therefore, it can help us to address the above mentioned problems and evaluate whether the process is under control or not by using control charts. When the process is affected with either special or assignable causes of variations, control charts will play a vital role of identifying or detecting these causes through the use control limits. On the other hand, when these causes exceed control limits, the process is said to be unstable and the causes must be identify and eliminated.

V. RESEARCH METHODOLOGY

Statistical process control is a very strong quantitative technique for managing and improving process stability. As we discussed earlier, our studies focused on control charts of SPC as the best tool to be used to carry out our analysis. This is because of its significance in detecting any process anomalies by calculating control limits and plotting the data on a chart to visualized the result. However, Figure 1 represents a simple diagram of control charts.

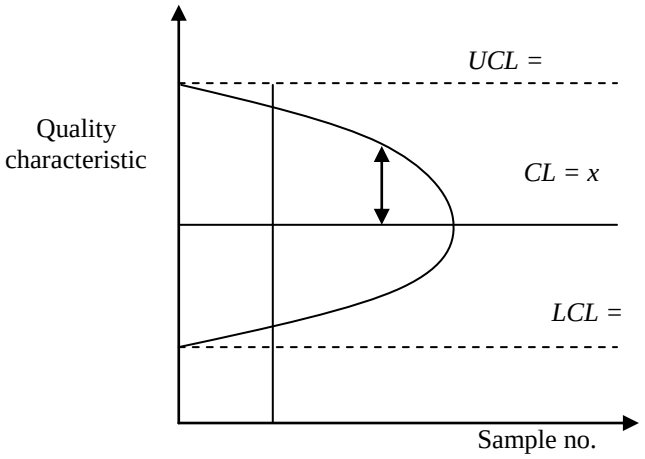

Fig. 1. A simple control chart

Control charts are used in software engineering in either variable data or attribute data. Xbar chart, R-chart, S-chart and XmR charts are all examples of variable data control charts. P-charts, u-charts, c-charts and XmR-charts on the other hand, are all examples of attribute data control charts. This study focused on u-chart because of its significance on our selected metric (defect density). A defect is a flaw in a component or system that can cause the component or system to fail to perform its required function [23]. For example, an incorrect statement or data definition is a defect. This clearly implies that, defects can be found in any phase of the software development life cycle. Therefore, defect density is the ratio of number of defects discovered in a component or system and the size (functional size) of the component or system as expressed below;

In view of the above therefore, the functional size can be source lines of code (SLOC) or lines of code (LOC). However, defect density is usually used in order to measure the quality of the work product. In other words, when the defect density is low during the phases of development, the quality of the product will be high where as when the defect density is very high, it implies that the software product quality will be very low. In line with this, defect density will provide us with the idea or answer of whether the customer will be satisfied with the product or not, how much rework is required and so on. As we stated earlier, attribute control charts such as u-chart can be used to monitor and control defect density in a process.

The u-chart attribute control chart shows how the frequency of defect changes occurred over a period of time.
The procedure for calculating control limits using u-chart is as follows:

(1)

Where by $\bar{\mu}$ is the average rate over all areas of opportunity (ai) and is obtained by using the following formula:

Furthermore, the area of opportunity should be calculated for each sample and can be found by using the following formula:

Whereby, *fi* represents the functional size or (SLOC) for each sample. Similarly, the lower control limit (LCL) as well as the upper control limit (UCL) can be calculated using the following formulas:

(2)

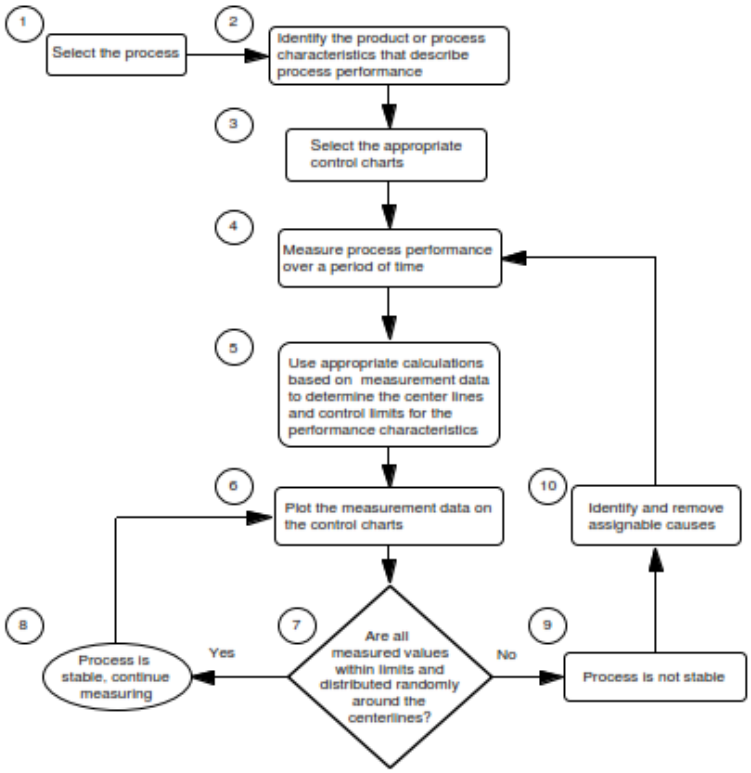

Fig. 2. Step-by-step procedure for evaluating process stability using control charts

(3)

Moreover, the presence of *ai* in the both lower and upper control limits indicates that each limit should be calculated or computed for each sample. By using these control limits, the u-chart is very sensible in detecting any deviation in the process based on the defective measures of the process. Despite the differences of control charts on our data, Figure 2 illustrates the step-by-step procedure of implementing control charts for process monitoring and control.

Based on the above Figure 2, to achieve real successful implementation of this approach in process investigations, three major steps should be considered. These are: collecting data, analyzing data and acting on results. Our study move forward by collecting data from CMM lower maturity level software industry (i.e CMM level 2). Due to the high confidentiality of the data, the name of this organization and the data cannot be stated in our study. In other words, our study will only reveal the result we obtained after collecting and analyzing the data.

VI. INITIAL RESULT

As seen from the literature, the use of control chart within the software domain is gaining more and more attention. The technique is usually used to quantitatively manage or monitor software process especially in CMM level 4 software industries so as to achieve process aims and objectives. In this section, we are going to justify the applicability of SPC for process monitoring and control at CMM lower maturity level software industries. In other words, the result we obtained by implementing or using control chart of SPC in CMM level 2 (repeatable) software industry for process control and improvement especially from the code peer review process point of view. We selected this process because of its significance in SDP. Without this process, many of our software products will lack certain quality attributes that will

leads to customer's dissatisfaction and inability to achieve business goals.

However, the data we collected consists of nine samples or modules that were completed in 2008 by the software organization. These samples have different features regardless of their types, architecture, size, quality attributes and so on. However, in this data we are able to identify the functional size as well as the defects that are associated or found in each sample. Similarly, we calculated the centre line, lower control limit and upper control limit for each sample using u-chart attribute control chart. The result of this data was plotted using the chart as is shown in Figure 3.

As seen from the figure, based on the principles of statistical quality control, sample number 1 fall outside control limits (i.e upper control limits) and therefore the process is not stable or out of control. This clearly implies that the root causes (assignable causes) should be investigated for proper process improvement. In other words, we can say that it is clear the process was not under control when it was started and this indicates that the product will lack certain quality attributes. Moreover, the purpose for any out of control situation should be investigated so as to have stable process that would yield positive result. When we can identify the causes and quickly react to them, therefore the process will definitely change.

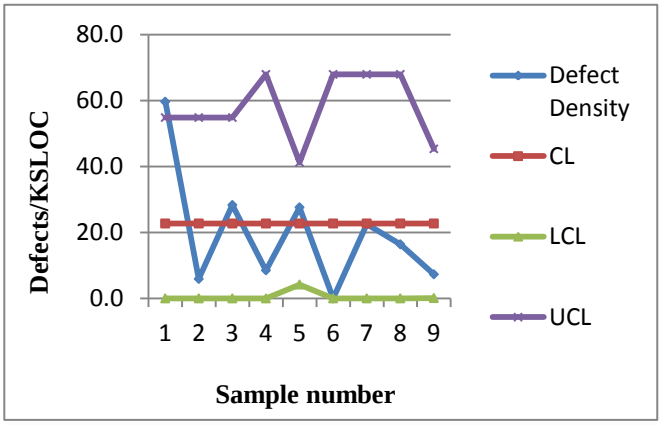

Fig. 3. u-chart in code peer review process

That is to say, detecting and eliminating process anomalies so as to bringing back the process under control would significantly enhance the quality of the software product. This is because; the quality of our software products depends largely on the stability of the SDP involved. Therefore, since we found the special cause of variation that makes the process out-of-control, we can simply remove it and examine the process behavior again. That is to say, we first remove project one and reinvestigate the process behavior by recalculating control limits and plotting new control chart. Figure 4 presents the u-chart of the process after anomaly was removed.

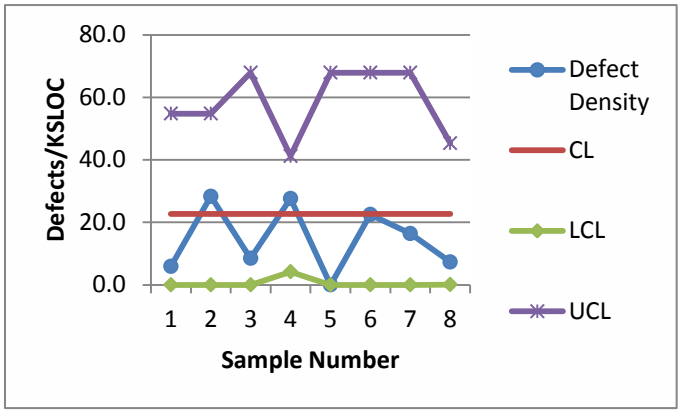

Fig. 4. u-chart control in code peer review process after anomaly was removed.

As we have seen in Figure 3 and Figure 4, we can simply say that the process is now stable (under control) and is defect free to produce quality software product. The high defect rate of the process was removed and the process is now in statistical control. However, sometimes unstable process may occur as a result of so many factors such as: lack of insufficient coding knowledge, violations of various coding rules and so 'on. In such cases, appropriate measures such as adequate staff training and rework will play a significant role in achieving process stability. Finally, based on what we discovered in our study, the data used in the first module is different with the remaining modules. It is strongly important to use control charts for process that have similar data or activities.

VII. PROBLEM ENCOUNTERED

Implementation of SPC in lower level CMM software industries is not a very straightforward activity. In this study, we encountered a problem of process measurement that requires effective SPC utilization in respect of these maturity levels. Improper measurement would yield undesirable results. In order to implement SPC, software engineers should be equipped with various tools or methods that will enable them to carry out effective and efficient software process measurement.

However, addressing this problem would contribute to a large extent on process improvements and product quality enhancements. In line with this, additional effort is required on proposing new methods for effective process measurement especially at the CMM lower maturity levels so as to achieve successful SPC implementations and preparing these industries for higher CMM maturity levels.

VIII. CONCLUSIONS AND FUTURE WORK

In this study, we have seen the idea of SPC on software process control. In responding to our stated questions, question 1 was addressed by detecting process anomaly and we achieved process stability using u-chart attribute data control chart. Question 2 on the other hand, was addressed by discovering the root causes of the process variation. This is as a result of mixing dissimilar process activities as a result of

improper process measurement. Finally, based on the case study we conducted in this research, we can conclude that lower CMM maturity level software industries can utilize SPC for effective process control and monitoring. Our future work is to propose an approach or strategy to achieve successful SPC implementation at the CMM lower maturity level organizations. This will answer our last question in this study.

ACKNOWLEDGMENT

Our sincere appreciations to the entire software engineering research group (SERG), Universiti Technologi Malaysia for their humble support.

REFERENCES

[1] Caivano, D. (2005). Continuous software process improvement through statistical process control. Software Maintenance and Reengineering, 2005. CSMR 2005. Ninth European Conference on, IEEE.

[2] Florac, W. A. and A. D. Carleton (1999). Measuring the software process: statistical process control for software process improvement, Addison-Wesley Professional.

[3] Florac, W. A., et al. (2000). "Statistical process control: analyzing space shuttle onboard software process." Software, IEEE **17**(4): 97-106.

[4] Lantzy, M. A. (1992). Application of statistical process control to the software process. Proceedings of the ninth Washington Ada symposium on Ada: Empowering software users and developers, ACM.

[5] Florak, W. A., et al. (1997). "Practical software measurement: Measuring for process management and improvement."

[6] Pandain, S., et al. (2013). "Control charts for improving the process performance of software development life cycle." International Journal of Research and Reviews in Applied Sciences **14**(2).

[7] Redner, R. A. and H. F. Walker (1984). "Mixture densities, maximum likelihood and the EM algorithm." SIAM review **26**(2): 195-239.

[8] Rao, B. S., et al. (2012). "Monitoring Software Reliability using Statistical Process Control An Ordered Statistics Approach." arXiv preprint arXiv:1205.6440.

[9] Weller, E. F. (2000). "Practical applications of statistical process control [in software development projects]." Software, IEEE **17**(3): 48-55.

[10] Zhou, K.-l., et al. (2012). Improving monitoring accuracy of process based on SPC method. Control and Decision Conference (CCDC), 2012 24th Chinese, IEEE.

[11] Zhang, Y. and L. Hou (2010). Software Design for Quality-oriented Statistical Tolerance and SPC. Information and Computing (ICIC), 2010 Third International Conference on, IEEE.

[12] Wang, Q., et al. (2006). BSR: a statistic-based approach for establishing and refining software process performance baseline. Proceedings of the 28th international conference on Software engineering, ACM.

[13] Tarhan, A. and O. Demirors (2012). "Apply Quantitative Management Now." Software, IEEE **29**(3): 77-85.

[14] Tarhan, A. and O. Demirors (2008). Assessment of software process and metrics to support quantitative understanding. Software Process and Product Measurement, Springer**:** 102-113.

[15] Baldassarre, M. T., et al. (2009). Statistically based process monitoring: Lessons from the trench. Trustworthy Software Development Processes, Springer**:** 11-23.

[16] Cangussu, J. W., et al. (2003). Monitoring the software test process using statistical process control: a logarithmic approach. ACM SIGSOFT Software Engineering Notes, ACM.

[17] Mûller, S. D., et al. (2009). "Managing cultural variation in software process improvement: A comparison of methods for subculture assessment." Engineering Management, IEEE Transactions on **56**(4): 584-599.

[18] Senbergs, J. and B. Misnevs (2006). "Research of a Statistical Process Control in Software Engineering Project." Transport and Telecommunication **7**(1): 70-75.

[19] Komuro, M. (2006). Experiences of applying SPC techniques to software development processes. Proceedings of the 28th international conference on Software engineering, ACM.

[20] Nguyen, T. H., et al. (2012). Automated detection of performance regressions using statistical process control techniques. Proceedings of the third joint WOSP/SIPEW international conference on Performance Engineering, ACM.

[21] Eickelmann, N. and A. Anant (2003). "Statistical process control: what you don't measure can hurt you!" Software, IEEE **20**(2): 49-51.

[22] Satya Prasad, R., et al. (2011). "Software Reliability Measuring using Modified Maximum Likelihood Estimation and SPC." International Journal of Computer Applications **21**(7): 1-5.

[23] Van Veenendaal, E. (2008). Foundations of software testing: ISTQB certification, CengageBrain. com.

Co-organizer

Malaysia Software Engineering Interest Group

Sponsors

Sponsors

YB Encik Sazali Bin Muhd Din @ Muhamad Din

ADUN N13 - DUN Paya Rumput

Sponsors


AMTIS
SOLUTION
Advance Management Technology
In Intelligent System

Sponsors

Sponsors