2011 International Conference on Modeling, Simulation and Control
IPCSIT vol.10 (2011) © (2011) IACSIT Press, Singapore

A Proposed Probabilistic Model for Risk Forecasting in Small Health Informatics Projects

[1]Basit Shahzad, [2]Azween Abdullah, [1]Abdullah S. Al-Mudimigh

[1]College of Computer & Information Science King Saud University, Riyadh, Saudi Arabia
University Teknologi PETRONAS, Malaysia
basit.shahzad@gmail.com, azweenabdullah@yahoo.com, mudimigh@ksu.edu.sa

Abstract. The history of software project failures is very old. In recent past huge emphasis has been given to investigate the possible causes of failure. In this regard an empirical study was conducted in 1996 on a group of 13,000 projects and it was concluded that almost 25% project were either delayed or fail projects . The project failure in any organization and in any environment dents the firms so badly that occasionally it becomes harder for them to survive. This means that delayed and failed projects not only put the name and reputation of the firm at question but also the firm's revenues, jobs of the employees, moral and prospects to gain future projects is also greatly damaged. With the advancement of technology and ease in development of small and medium level systems has strongly encouraged the development of information systems and the Health specific Information Systems (IS), generally referred to as Health Information System (HIS) or Medical Information System (MedIS). The science that governs the development of MedIS or HIS is known as Medical Informatics (MI) that has emerged as a consolidated branch of IS during last thirty years. In precise terms (MI) "is primarily intended to provide solutions for problems of processing data, information and knowledge in medicine and healthcare" . There are a lot of risks in developing the projects that are related to the MI. The research is focused to provide the ability to forecast the risks in MI by categorizing the project in small, medium and large scale and they by applying the probabilistic model calculating the outcome value for each risk factor and ultimately mapping it back to the project factors to reduce the risks by adjusting the values assigned to them. This paper identifies and helps in suggesting the ways to resolve the risks in the software development life cycle.

Index Terms: Risk forecasting, Software Risk Management, Software resource balancing, Medical Informatics risk forecasting.

1. Introduction

The history of software project failures is very old. In recent past huge emphasis has been given to investigate the possible causes of failure. In this regard an empirical study was conducted in 1996 on a group of 13,000 projects and it was concluded that almost 25% project were either delayed or fail projects [5]. The project failure in any organization and in any environment dents the firms so badly that occasionally it becomes harder for them to survive. The cost failure depends on the cost of project under consideration, the huge the cost and time of the project under consideration the huge impact it will have if fail. This means that delayed and failed projects not only put the name and reputation of the firm at question but also the firm's revenues, jobs of the employees, moral and prospects to gain future projects is also greatly damaged. These are the parameters that a firm can't compromise on. They have to ensure that they can use the best tools and technologies in the market to enforce the smooth execution of the project and also that the project schedule is followed in order to make the project a success story, while the consequences of a failure can be huge.

Despite the fact that the failures are huge in the software development environment, the software community has learned a lot from these failures and has come of with more strong defense line to deal with the problems that may lead to the software failures. Not only the software development infrastructure has improved but also the tools to use this infrastructure to develop software have also improved. The advancement in implementation technology has increased the hopes of individuals working on the software

development and quite a few new area have been introduced since then. With the advancement of technology and ease in development of small and medium level systems has strongly encouraged the development of information systems and the Health specific Information Systems (IS), generally referred to as Health Information System (HIS) or Medical Information System (MedIS). The science that governs the development of MedIS or HIS is known as Medical Informatics (MI) that has emerged as a consolidated branch of IS during last thirty years. In precise terms (MI) "*is primarily intended to provide solutions for problems of processing data, information and knowledge in medicine and healthcare*" [4]. There are a lot of risks in developing the projects that are related to the MI. This research is not expected to lean into medical infrastructure, employee issues and medical treatments. So, the focus of this research is on the software development life cycle that helps in producing the Health Informatics software. [8] The research is focused to provide the ability to forecast the risks in MI by categorizing the project in small, medium and large scale and they by applying the probabilistic model calculating the outcome value for each risk factor and ultimately mapping it back to the project factors to reduce the risks by adjusting the values assigned to them[6,7].

2. Related Work

Recently some projects are being conducted in academia and industry that are quite relevant to my research.

2.1 Trial DB

Trial DB is a customized web based which is used for storage and management of the data about patients used in clinics. Trial DB relies on the libraries that contain information about every data element and the classes to which the elements belong to. The libraries are used to provide data and procedures for the Customers Report Form (CRF) which is produced by the web services. The project addresses the typical record management and the means to access the information. The information in this project can be helpful in guiding about the physical, logical and sequence information of procedures that may guide us in this research [1].

2.2 Influence of Studies on Students' Health (ISSH)

The project is developed to establish types of education in relation to stress related diseases. The project is based on the survey that contains 25 questions which is statistically compared between the students of the countries involved. The results showed that appropriate changes need to be made in medical curricula in order to improve the quality of education and student life; there is a need for stress management at medical facilities. This project is based on the collection of data from the students about the mental stress that they can have. The data collection and analysis techniques used in this project can be of great help for usage in my research [2].

2.3 Information Management with Privacy, Lineage, Uncertainty, and Security

The project address the information linage issue (from where the data came and the process that act on data) to determine if the information produced from this process is certain, useful and trustworthy. The project focuses on extending the prior research on data lineage and workflows to include Web information, and processing chains used in data fusion applications. The project analyzes interactions among privacy, lineage, uncertainty, and security requirements to exploit shared efficiencies, highlight trade-offs, and support the user's decisions in interpreting and trusting the data. This project focuses to address the authenticity of data by working on the data and process. As data used in actual risk management has to be precise and accurate the techniques used in this project can be extremely helpful in this regard.

2.4 Analysis and design of decision making system.

The project focuses on the vision to make it possible that the patient specific diagnostics and medical data is made available to the doctor in every circumstance by using intelligent modeling and real time analysis. The project assumes that statistical techniques for modeling and data integration can hardly support the critical decision making process therefore the existing techniques require modifications before new techniques are derived. The project is focused to exploring the techniques for data analysis and

understanding which are different from the existing techniques [6,14,15]. The project is focused to propose new techniques for exploring and understanding new techniques for analysis and understanding of the healthcare data. The techniques developed for data analysis can be helpful in gathering the problems/risks specific to medical informatics.

3. Research Questions

In light of the discussion above in section I and II, it is logical to identify some questions that should be addressed in this research. Following research questions are identified in this regard:

A. How to develop a detailed probabilistic model that can identify, support, manage and compare the risks in medical informatics domain?
B. How to identify a relationship between the type of project, project resources and the impact of risk produced as an ultimate outcome of the module?
C. How to prepare the detailed list of risks in medical informatics to be applied in newly developed probabilistic model to produce overall impact for any risk factor under medical informatics system?
D. How to increase the trust of practitioners, developers and project managers on the risk assessment and comparison of risk handling tools in medical informatics?

4. Proposed Solution

The proposed strategies are expected to result in a list of risk factors in the domain of Health Informatics, that may be of diversified nature. The risks involved can be operational, logical, procedural, managerial, technical or even environmental. As described in figure 1, the scope of this research is not limited to just identifying the risks in medical informatics, but it is also bound to provide the support of a probabilistic model to assess and handle the risk factors for different projects related to health informatics. The flow for the implementation of this research is presented in the model in figure 1 and is described in the proceeding section[12,13]. The ultimate objectivity will be implemented with the help of following modules of the model.

The mandate of this research is to identify the risks in healthcare informatics domain and once they have been identified they have to be rationalized as well. The scale of projects for which the risk is being considered is also to be taken into consideration as large projects have high orientation and exposure towards the risks. Once the risks have been identified with the scale of the project they will be passed through a probabilistic model to transform the results into an equation, that: against some given values will be able to identify the overall threat level to the system. The reverse of that equation will guide us that by reducing that threat value what impact is made on allocation of resources and a re-adjustment of resources may be recommended and justified at that time.

It is also worth mentioning at this point that this research will hardly dig into the problems that are specific to patients, medicine or the infrastructure, but it will focus the risks that are faced in development and maintenance of the health-informatics projects. So the identification of risks, problems, process and procedure will all be focused to identifying and improving the software development process for the healthcare projects.

4.1 Identification Module (IM)

As the success of risk management strategies depends on the identification of more and more legitimate risks, I shall be interested to observe more and more risks, so that a better management can be made in this regard. The collection and identification of risk factors itself is a broad area of research and involves a huge empirical and experimental study. This research is bound to consider both qualitative and quantitative measures to identify as many risk factors as possible in diversified areas around the scope of this project. In order to implement the qualitative and quantitative measures to collect the risk factors the focus will be the risk registers, questionnaires, interviews, experience and intuition of the risk analyst. Other development team members will also be consulted about their viewpoint about the risk factors and the handling of risk factors as well. This research is expected to focus different types (at least three) of sub-systems within the MI

and therefore the associated risks have to be different as well based on the nature and recourses of the project. The model will provide the user with a list of possible risk factors for any specific project and for the specified parameters of the project (that will govern that if the project is small, medium or large). Thus the risk factors for same type of project may vary from each other depending on the recourse allocation for the project, which will determine that in which category the project is expected to fall w.r.t. its scope.

The scope determination parameters, although, somewhat known, yet will be supported by the extensive literature survey to provide solid ground for the foundations of further research in the domain of medical informatics.

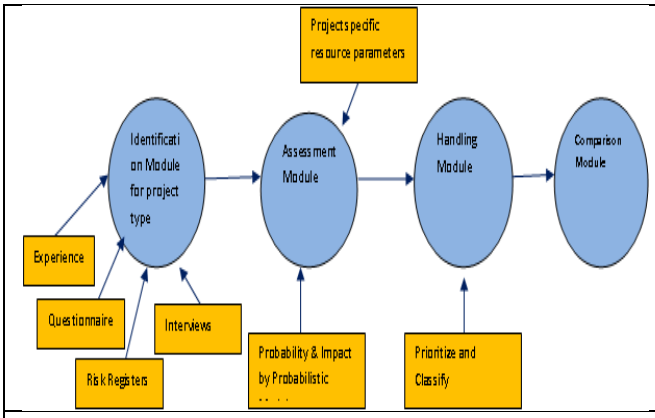

Fig 1: Model for execution of research

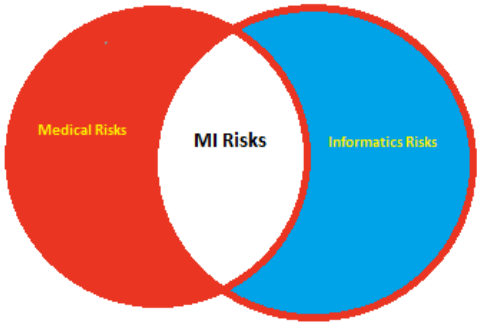

Fig 2: Identification of risks for Medical Informatics

4.2 Risk Assessment Module (RAM)

After identification of risk factors, they will be prioritized with respect to their severity level on the Medical System development process. This module is bound to suggest and implement the most probable impact and probabilities for each risk factor by applying the probabilistic model that will yield its force from the literature not only for its construction and existence but also for proposing the suitable probabilities and impact. With this the severity levels for each risk shall be calculated based on the probabilistic model.

4.3 Risk Handling Module

A prioritized list of risk factors is the legitimate outcome of the RAM module. The risks will be prioritized based on the overall impact (that will be calculated from the impact and probability of individual risk factor) Risks may be classified to individual phases of software development, in order to enable the project manger providing the specialized management in all phases of software development.

4.4 Comparison Module

This module is responsible to analyze the prioritized list of risk factor from the module Risk Handling Module. The impact of risks can be observed in different projects with different parameters suggesting the impact and severity of risks on the development of MI systems. These results of this module can be analyzed repeatedly by changing the subsequent evaluation parameters to accomplish the balanced risk handling in MedIS.

5. Conclusion

The research provides the possible risk in different scales of the Medical Informatics project, as there is less research on defining scales of project with respect to the resources allocated to them: a survey specific for this purpose establishes the scales. The probabilistic model helps in translating the recourses to the equation for improving or decreasing the impact of risk factors. The research broadly helps in decreasing the risk by forecasting hem and allowing the development team to rationalize the resource allocation.

6. References

[1] http://ycmi.med.yale.edu/TrialDB/

[2] http://www.ifmsa.org/index.php?option=com_content&view=category &layout =blog&id=59&Itemid=70

[3] http://www.mitre.org/news/events/exchange09/8.html

[4] A. Hasman, R. Haux, A. Albert, "A Systematic view on medical informatics", Computer methods and programs in biomedicine, volume 51, pp 131-139, 1996.

[5] *Source: Patterns of software failures and successes, Capers Jones, 1996*

[6] http://www.dcs.qmul.ac.uk/research /radar/core_pages/projects.html

[7] B. Kitchenham and S.L. Pfleeger, "Principles of survey research part 1," *ACM SIGSOFT Software Engineering Notes*, 2002, p. 20.

[8] B. Kitchenham and S.L. Pfleeger, "Principles of survey research part 2," *ACM SIGSOFT Software Engineering Notes*, vol. 27, May. 2002, p. 20.

[9] B. Kitchenham and S.L. Pfleeger, "Principles of survey research part 5," *ACM SIGSOFT Software Engineering Notes*, vol. 27, May. 2002, p. 20.

[10] B. Kitchenham and S.L. Pfleeger, "Principles of survey research part 4," *ACM SIGSOFT Software Engineering Notes*, vol. 27, May. 2002, p. 20.

[11] The Standish group report, http://www.projectsmart.co.uk/docs/chaos-report.pdf , pp 1-8

[12] C.E. Journal, "Resistance factors in software processes improvement," vol. 9, 2006.

[13] Boehm, Barry, and Chris Abts. "This paper is an extension of the work done by Sunita Chulani as part of her Qualifying Exam report in partial fulfillment of requirements of the Ph.D. program of the Computer Science department at USC [Chulani 1998].." *Science* (1998).

[14] Nasir, Mehwish, and Case Study. "A Survey of Software Estimation Techniques and Project Planning Practices." *Computing*, 2006.

[15] Reel, John S, and Trident Data Systems. "Critical Success Factors In Software Projects." *Ieee Software*, no. June 1999: 18-23.