Software Delivery Risk Management: Application of Bayesian Networks in Agile Software Development

Ieva Ancveire1, Ilze Gailite2, Made Gailite3, Janis Grabis4
1, 2, 3, 4 Riga Technical University

Abstract – The information technology industry cannot be imagined without large- or small-scale projects. They are implemented to develop systems enabling key business processes and improving performance and enterprise resource management. However, projects often experience various difficulties during their execution. These problems are usually related to the three objectives of the project – costs, quality and deadline. A way these challenges can be solved is project risk management. However, not always the main problems and their influencing factors can be easily identified. Usually there is a need for a more profound analysis of the problem situation. In this paper, we propose the use of a Bayesian Network concept for quantitative risk management in agile projects. The Bayesian Network is explored using a case study focusing on a project that faces difficulties during the software delivery process. We explain why an agile risk analysis is needed and assess the potential risk factors, which may occur during the project. Thereafter, we design the Bayesian Network to capture the actual problem situation and make suggestions how to improve the delivery process based on the measures to be taken to reduce the occurrence of project risks.

Keywords – Agile software development, Bayesian networks, project risk management, project success, software delivery.

I. INTRODUCTION

Since the advancement of agile methods in software development processes, they have been considerably applied to projects of different application domains. Although their application has become more popular because of the demand for developing flexible information technology solutions faster, there are still several problems to be solved and explored [1], [2].

While traditional software development models are based on precisely and qualitatively defined requirements and subject to strict regulations, agile is not fully defined [3], [4]. Nowadays, there is no definite consensus on the need for risk management within the agile approach. This has led to believe that a software risk analysis is irrelevant using this approach. Many follow the approach to manage risks through the natural sprint progression after they have manifested into issues [5]. Nevertheless, irrespective of the methodology used in software development, there is a possibility for project to fail [6]. If the reasons for unsuccessful project outcome using traditional project management are incomplete and changing requirements, lack of user involvement, resources and executive support caused by communication and interaction issues between developers and stakeholders, then in agile multiple software development problems arise due to incomplete understanding of its principles that include collaboration between business and IT as well as organisation of work. It is noted that an agile approach does not guarantee for project to be successful and risk management should be considered a significant remedy in project management and software development [4]–[7].

Project risks are usually identified maintaining a list; thus, their mutual interaction is rarely investigated. Many different quantitative evaluation methods can be used for research, seeking reasons for project success or failure. One of the methods that should be mentioned is a Bayesian Network (BN). Prosperous project management, progress and closing are important to both business and technology representatives [6]; thus, the purpose of the present research is to investigate agile project risks, their interaction and dependency using BN in order to reveal project success factors.

The reminder of the paper is organised as follows. Section II summarises the related work. Section III describes the method used for further research on agile risk management. Section IV reports a case study and issue identification. Section V presents the results applying the developed model on actual information technology project. In conclusion, Section VI provides with a summary of results obtained and opinion on modelling suitability for agile development.

II. RELATED WORK

Serviceability and high-quality are meaningful characteristics of software. In order to deliver information technology solutions more successfully, many studies and papers have been written, and various models and methods have been developed. Speaking of these different techniques, many have been applied to software project planning, management and development.

BNs have found numerous applications in software development, especially, in relation to iterative development practices. They are viable means for assessment and prediction in the case of uncertainty. Generally speaking, this approach has been applied to fields such as investigation of software influencing factors, management process overview, implementation of project management methodology, release planning, risk assessment, decision-making support, quality assurance and prediction.

The impact factors that affect the agile development cycle should be evaluated constantly in order to ensure high-quality software on time. For these purposes, a model has been introduced for agile software development release planning and project health monitoring [2]. However, these influencing factors can be accompanied by issues. A procedure that
describes how to detect a variety of problems in software development processes through a conceptual network has been explored and discussed in literature as well. This model has been successfully applied to agile project management and planning using Agile-Scrum methodology, managing various processes and achieving the success of the solution [4]. Complementing the above-mentioned considerations, a pattern is developed that assists a project manager to implement this method in the company. It provides the responsible person with information about the potential problems and issues that can arise during the project or through the management process. It helps manage the business and development teams to improve chances of project success [7]. In order to integrate an effective risk management process, a model has also been developed for software process risk assessment using BN. There is a study related to the decision-making in risk management systems and their necessity, showing that they support the process and the occurrence of risks, assessing both complexity and dependence [8]. Despite the importance of software quality, the management of it is difficult. To ease this process, an approach has been introduced in order to assess and predict software quality, using activity-based quality models [12].

Our challenge is to implement and model the software delivery process and consequent risk interdependence in practice, by constructing a BN. The goal is to show how the constructed network could help in specific project risk management. We intend to investigate which risks significantly affect the quality of software delivery, analyse what benefits the model can provide and prove that the application of BN reflects the real situation of the problem domain in projects using agile.

III. BAYESIAN NETWORKS

A BN also called a belief network is an acyclic graph structure that describes relations between variables, containing different causes and effects to ensure formal description of problem domain [9]. It consists of two major parts: (1) graphics including nodes and values of the occurrence of particular event; (2) conditional probability tables covering dependencies between the graph nodes (see Fig. 1). Probability table is defined and built for each node of the network. It is based on predecessors. Dependency values between nodes could be settled as Boolean, arrangement or integral meaning truthfulness, order and grouping or range [2], [8], [10].

Examining node E (see Fig. 1), probability tables for a network are constructed taking into account predecessors of particular node finding whether events are met or not.

![Expanded Bayesian network model with conditional probability tables.](image)

The probabilities are controlled by the Bayesian formula, where $A$ – hypothesis or probability of a particular event, and $B$ – evidence or condition of related node [11].

\[
P(A|B) = \frac{P(B|A)P(A)}{P(B)}
\]

Let us find the probability for true event $E$ to prove Bayesian formula (1), assuming that related events $A$ and $B$ are also true ($A, B = 1$). The probability to find is $P(E = 1, A, B = 1)$. There is a necessity to use probability tables for solving this situation. If $A, B$ of the network are true (see Fig. 1), but $C$ is unknown, both scenarios are necessary to be described, meaning that the end probability for node $E$ equals the sum of both situations ($C = 0$ and $C = 1$).

\[
P(E = 1|A = B = 1) = P(E = 1, C = 0|A = B = 1) + P(E = 1, C = 1|A = B = 1)
\]

\[
P(E = 1)|A = B = 1|P(E = 1|C = 0) + P(E = 1|C = 1) = (100 - P(4))P(5) + P(4)P(5)
\]
project risks that affect project delivery in each software delivery iteration.

IV. CASE STUDY

In order to demonstrate that the risk analysis is important and substantial in any project affecting its success and failure, BN application possibilities will be justified in a real agile software development situation where the project faces problems associated with low-quality software deliveries due to negligent risk identification. To explore these opportunities, it is important to find answers to the following questions: (1) Does the dependency analysis using BN help reflect the real situation of the software delivery process? (2) How does BN help improve software delivery process? (3) What are the key risks associated with the need to manage and monitor them through the agile project delivery process? (4) Is the BN suitable for agile project risk management? (5) Whether and how does the quantitative risk analysis using BN ensure project success?

A five-step procedure is used to analyse the case. The software delivery process is analysed in the first step. That includes risk identification and reasoning about their dependencies. In the second step, the results of the first step are used to construct a network for quantitative evaluation for each risk factor. The constructed graph structure must be validated and experiments with it must be carried out as well, looking at various situations provided for comparison with the previous software delivery outputs, resulting in those situations where attention is required. Then experimental, forecast and actual results should be compared to obtain conclusions whether this kind of risk identification and analysis is needed to improve the project management, solve problems, meet and satisfy the customer's demands. Finally, opinion on this matter should be provided, explaining how difficult such a risk analysis is carried out and how suitable it is for agile project management.

Justification will be based on a project for the communications and telecommunications industry. The goal is to develop an information system that provides tangible and intangible resource management, implements different kinds of information collection into a single solution, ensures the increase in business value and reduces the paper flow in the company. This solution also provides integration of two databases, eight system modules and six subsystems. Further work is based on the development data of one subsystem.

The success of the project, regardless of the selected design methodology, is described by a number of factors influencing the project. The project under review has seen some problems which are opposite to generally defined agile, including financials, requirements, scope, project team, planning and organisation of work, system architecture and feedback.

Project development is regulated by a fixed price, meaning that all changes are implemented without the involvement of additional funds. Therefore, customer requirements tend to change, causing situations when client and supplier are unable to agree on the necessary functionality. Frequent changes in the project scope are also indicators that can affect a successful go-live of the project. There is a large amount of functionality to be developed and an excessive focus on details of the initial development may result in the loss of target. Although the development team is small, the team members and the client have different levels of expertise, personal interests, commitments and support in problem resolution. Agile also requires good planning and organising work. In this case, time and work planning have not been determined by the project management method, in which the compliance and knowledge of it are required but not observed. There is no complete understanding of the system architecture and technologies used as well. This adversely affects the outcome of the project. Deficiencies arising from servers, applications and database must be taken into account to avoid environmental gaps and standard software vulnerability. Finally, feedback has a significant influence on project progress. If this is missing, then project progress and service cannot be delivered or are at a struggle, taking into account communication problems between team members and the client.

![Software delivery process](Fig. 2. Software delivery process.)
Software delivery process (see Fig. 2) is initiated by the client’s request. With the help of messaging, the request reaches the supplier’s project manager whose role is to examine the incoming request and decide whether a new conception is needed or it is possible to immediately start the development. Regardless of the decision, the project manager creates an issue and delegates it to a developer or a system analyst for further work. The system analyst is responsible for functional clarification, issue modification and submission to the developer. Next activities involve the development, unit testing and elaboration set-up in the test environment where the system analyst performs acceptance tests. Another task for the system analyst is to create test scenarios for the client. This is always done during the acceptance test phase. When all the aforementioned activities are completed, the project manager forms the delivery and sends it to the customer. In turn, when the delivery is received, the client performs acceptance tests. In case of successful completion of the tests, the client installs the new functionality in the production environment. If the acceptance tests are unsuccessful, the client creates new requests. If the number of returned requests is greater than 50% of the requests included in the delivery, it is classified as unsuccessful.

Software delivery to the customer consists of the following: (1) Mandatory delivery – the new system functionality and change requests are delivered to the customer once a week; (2) Additional delivery – pre-supplied software fixes are delivered once or multiple times a week, based on issues found during the acceptance tests; (3) Exceptions – if there is no scheduled change or new functionality development, pre-supplied software fixes or high priority fixes can be delivered as part of mandatory delivery.

The amount of functionality, change requests and the number of fixes are variable and can be different for each iteration (see Table I) due to various factors that occur between delivery packages and existing delivered software.

![Table I: Data from Delivery Journal](image)

### V. RESULTS

**A. Risk Identification**

Project risks could be defined as uncertain events or conditions that hold a positive or negative effect on at least one of projects objectives – time, cost, scope or quality. A risk factor is any attribute, characteristic or exposure that increases the likelihood of project to succeed or fail. Even though risk management is not mandatory in agile projects, there are multiple risk factors that could be defined for iterative development projects related to integrated risk management [11], [14], [15]. Risk factors for agile projects usually are unclear objectives, ineffective stand-up meetings, inadequate prioritisation of requirements [14], etc. We use different risk factors for agile software delivery risk identification (see Table II) to formally describe the main problems of the project (see Section IV).

![Table II: Risk Factor Analysis](image)
B. The Construction and Validation of BN

The software delivery model is constructed on theoretical restrictions mentioned in Section III, the software delivery journal containing delivery data for the last six months (see Table I) and risk factor analysis explaining the actual situation of the project (see Table II). In addition, an important step is also to make a quantitative risk analysis. In our case, we use probability matrices for each node of the network taking into account experience and daily observations on the occurrence of risk factors. Fig. 3 represents the constructed network describing software delivery. The formal description of BN results in the probability of each event, taking into account the direct descendants. Bayes Server\(^1\) software is used to evaluate the risk model. After determining probabilities of all events, the probability of unsuccessful delivery \(P_u\) was evaluated as \(P_u = 64.94\%\) (i.e., project success probability was 35.06\%). To validate the evaluation results, \(P_u\) was compared to data recorded in the delivery journal (see Table I).

\(^1\) Bayes Server is an analytics software covering various fields, such as Machine Learning, Data Science, Artificial Intelligence and Time Series Analysis. It is used to make predictions, extract insight, and perform decision support using models built from data and expert opinion.
Additionally, the number of requests returned by the client versus the total number of requests was compared on a weekly basis. Fig. 4 represents the total number of supplier’s requests and number of successfully and unsuccessfully delivered requests. Fig. 5 shows unsuccessfully delivered requests as a percentage of number of requests delivered. The percentage of unsuccessfully delivered requests over 27 weeks is 65.35%. That indicates that the actual percentage of the unsuccessful deliveries is close to the estimated probability $P_u$ and the risk model can be perceived as adequate for the given problem situation.

Fig. 4. Delivery statistics.

Fig. 5. Unsuccessfully delivered applications.

C. Risk Analysis

To investigate how to reduce the probability of $P_u$, the data of 1st week of August are used as a basis, where three fixes, one change request and five new system functionalities are delivered. To limit the possibility of unsuccessful delivery, there is a need to evaluate the controllable risks of delivery process and the possible prevention methods, which would allow managing software delivery more effectively. By researching the problem situation, one of the risks that can be controlled is unavailability of project owner with the occurrence of 37.56%. To reduce the influence of this risk on $P_u$, actions that can be done is encouraging Project Owner (further – PO) to be available for team interactions and surrogating PO to interact with PO on the client side [14]. The requirement conflict risk with the occurrence probability of 37.92% can also be controlled. To control it, such actions as facilitating group discussions, showing demonstrations to the client, recording the minutes of meetings with the customer, performing user acceptance tests after each build and team members visiting end-user location can be done [14]. Software delivery risk can also be reduced by test Data Management control, the occurrence probability of which is 60%. This risk
has a direct impact on inappropriate test scenario regulation – 53.43%. Thus, uncertain requirements affect the quality of test scenario accuracy. In cases when full amount of test data is not available, it is needed to automate test data management by using appropriate tools [14]. Test scenario quality can be improved by also having onsite a representative from the vendor at the client location and using exploratory testing as it needs less documentation [14]. In turn, the occurrence probability of reorganisation of project team risk (initially 80.33%) can be reduced by decreasing employee rotation. It can be controlled with majority of team members being core members and assigning a mentor to new team members to speed them up [14]. Such a solution is also suitable for controlling different levels of expertise risk, thereby reducing the 100% occurrence probability. In case of risk occurrence there is a need to involve additional human resources, to encourage growth in team size, which is one of the success factors of agile. It determines that relationship is the experience level of the project team. It is recognised that project team members with greater experience and background in project-based work are more adept at completing their assignments, working together collaboratively, and performing tasks efficiently [6]. The most suitable solutions to release management and deployment risk (36%) reduction are frequent and periodic release dates, keeping infrastructure for component integration flexible, keeping functionality flexible if dates are constrained [14]. Standard Software Vulnerability (52%) and Environmental Gaps (40%) are controllable risks, but full prevention of these is only possible in the project planning phase. The team should choose the SW and HW tools after careful tool study and refer to existing guidelines for tool selection for engineering activities [14]. During the delivery in August, there was an enhanced control of semi-automated delivery risk by developing guidelines for delivery installation process and enhancing communication with involved parties in organisation communication channels. After executing the experiment, it was recognised that $P_a$ percentage can be reduced to 58.85% (see Fig. 6). Additionally, it is possible to mathematically calculate the effect of this risk when it occurs. The calculations show that the control of this risk has reduced $P_a$ occurrence by 11.31%.

VI. CONCLUSION

Since the success of any project lies in its management and control, project main objectives can be achieved following the instructions. Although usually a risk analysis is not carried out in agile projects, it may have a significant impact if the project is complex, large-scale and faces numerous problems during the development.

In the present research, we have used a Bayesian Network concept to display an agile project software delivery process in order to seek out whether such an analysis accompanied by suitable actions ensures and increases the service quality. The results have shown that it is possible to limit the failure percentage by monitoring the controllable risks and their impact factors. In addition, direct ancestor controlling constitutes the highest improvement of event under scrutiny.

This kind of risk factor management provides a complete and accurate representation of the actual situation in the project, especially when there is a clear understanding of the key problems associated with risks and their dependencies. Accompanied by statistical and quantitative data, the model is usable for various kinds of experiments related to investigation and forecast purposes. It ensures demonstrative representation of main risk factors and can assist in the decision-making process.

Practicality of this kind of modelling in agile projects is still a moot point. Despite the fact that the graph design process is relatively simple, conditional probabilities are hard to get. It takes time of daily observations and experience for correct probability determination. A knowledge management system, guidelines and a model verification analysis could assist for ensuring the accuracy of data entered. Nonetheless, such models can be used repeatedly describing processes that are standardised, implying that they can be reused within the organisation for projects to come.

REFERENCES


Using Interpretive structural modeling, in Indian apparel retail chains and proposal of risk prioritization model


E-mail: ieva.anceire@edu.rtu.lv

Ilze Gailite holds a B. Sc. (2013) degree in Information Technology from Riga Technical University, Latvia. Since 2015 she has been working as a Junior Software Tester / Systems Analyst at Autentica, Ltd. Her areas of interest are project management, IS testing, test management and system analysis.

E-mail: ilze.gailite@edu.rtu.lv

Made Gailite obtained her Bachelor Degree in Computer Management and Computer Science from Riga Technical University, Latvia, in 2013. Since August 2015, she has been working as a Software Engineering Associate at Accenture Latvia. Her main areas of interests are enterprise resource planning, SAP, software development, data analysis and cloud solutions.

E-mail: made.gailite@edu.rtu.lv

Janis Grabis holds a Doctoral Degree and is a Professor at Riga Technical University (Latvia) and the Director of the Institute of Information Technology. His main research interests lie within the application of mathematical programming methods in information technology, enterprise applications and system integration. He has published more than 80 scientific papers, including a monograph on supply chain configuration. He has led a number of international and national projects and has participated in five projects in collaboration with the University of Michigan-Dearborn (USA) and funded mainly by industrial partners, such as SAP America and Ford Motor Company.

E-mail: grabis@rtu.lv