Evaluating and Improving Risk Analysis Methods for Critical Systems

Sardar Muhammad Sulaman

Doctoral Dissertation, 2019
Department of Computer Science
Lund University

ISBN 978-91-7753-950-6
ISBN 978-91-7753-951-3
ISSN: 1404-1219
LU-CS-DISS: 2019-02
Dissertation 62, 2019

Department of Computer Science
Lund University
Box 118
SE-221 00 Lund
Sweden

Email: Sardar@cs.lth.se
`http://www.cs.lth.se/sardar_muhammad_sulaman`

Typeset using LaTeX
Printed in Sweden by Tryckeriet i E-huset, Lund, 2019

To my sons, Nameer & Ifad

ABSTRACT

At the same time as our dependence on IT systems increases, the number of reports of problems caused by failures of critical IT systems has also increased. Today, almost every societal system or service, e.g., water supply, power supply, transportation, depends on IT systems, and failures of these systems have serious and negative effects on society. In general, public organizations are responsible for delivering these services to society. Risk analysis is an important activity for the development and operation of critical IT systems, but the increased complexity and size of critical systems put additional requirements on the effectiveness of risk analysis methods. Even if a number of methods for risk analysis of technical systems exist, the failure behavior of information systems is typically very different from mechanical systems. Therefore, risk analysis of IT systems requires different risk analysis techniques, or at least adaptations of traditional approaches.

The research objective of this thesis is to improve the analysis process of risks pertaining to critical IT systems, which is addressed in the following three ways. First, by understanding current literature and practices related to risk analysis of IT systems, then by evaluating and comparing existing risk analysis methods, and by suggesting improvements in the risk analysis process and by developing new effective and efficient risk analysis methods to analyze IT systems.

To understand current risk analysis methods and practices we carried out a systematic mapping study. The study found only few empirical research papers on the evaluation of existing risk analysis methods. The results of the study suggest to empirically investigate risk analysis methods for analyzing IT systems to conclude which methods are more effective than others. Then, we carried out a semi-structured interview study to investigate several factors regarding current practices and existing challenges of risk analysis and management, e.g., its importance, identification of critical resources, involvement of different stakeholders, used methods, and follow-up analysis.

To evaluate and compare the effectiveness of risk analysis methods we carried out a controlled experiment. In that study, we evaluated the effectiveness of risk analysis methods by counting the number of relevant and non-relevant risks identified by the experiment participants. The difficulty level of risk analysis methods and the experiment participants' confidence about the identified risks were also investigated. Then, we carried out a case study to evaluate the effectiveness and efficiency of existing risk analysis methods, Failure Mode and Effect Analysis (FMEA) and System Theoretic Process Analysis (STPA). The case study investigates the effectiveness of the methods by performing a comparison of how a hazard analysis is conducted for the same system. It also evaluates the analysis process of risk analysis methods by using a set of qualitative criteria, derived from the Technology Acceptance Model (TAM). After this, another case study was carried out to evaluate and assess the resilience of critical IT systems and networks by applying a simulation method. A hybrid modeling approach was used which considers the technical network, represented using graph theory, as well as the repair system, represented by a queuing model.

To improve the risk analysis process, this thesis also presents a new risk analysis method, Perspective Based Risk Analysis (PBRA), that uses different perspectives while analyzing IT systems. A perspective is a point of view or a specific role adopted by risk analyst while doing risk analysis, i.e., system engineer, system tester, or system user. Based on the findings, we conclude that the use of different perspectives improves effectiveness of risk analysis process. Then, to improve the risk analysis process we carried out a data mining study to save historical information about IT incidents to be used later for risk analysis. It could be an important aid in the process of building a database of occurred IT incidents that later can be used as an input to improve the risk analysis process. Finally, based on the findings of the studies included in this thesis a list of suggestions is presented to improve the risk analysis process. This list of potential suggestions was evaluated in a focus group meeting. The suggestions are for example, risk analysis awareness and education, defining clear roles and responsibilities, easy-to-use and adapt risk analysis methods, dealing with subjectivity, carry out risk analysis as early as possible and finally using historical risk data to improve the risk analysis process. Based on the findings it can be concluded that these suggestions are important and useful for risk practitioners to improve the risk analysis process.

The presented research work in this thesis provides research about methods to improve the risk analysis and management practices. Moreover, the presented work in this thesis is based on solid empirical studies.

ACKNOWLEDGEMENTS

This research work was funded by the Swedish Civil Contingencies Agency (MSB) under a grant for PRIVAD, Programme for Risk and Vulnerability Analysis Development.

To this point there are so many people who have stood by me in the ups and downs and so many inspiring people that I have met during the years that have passed. First of all, I would like to express my sincere gratitude and thanks to my supervisor, Prof. Martin Höst, for his great support, valuable feedbacks, discussions, continuous support and encouragement. Thanks for always being ready to take discussions and for guiding me. I am also thankful to Dr. Kim Weyns, who was my co-supervisor during the first two years of my PhD studies, for his support and critical feedbacks. I also would like to thank Prof. Per Runeson, my co-supervisor after the first two years of my PhD studies, for his great support and fruitful discussions.

I owe a huge THANK to all the people at the department of Computer Science for providing an inspiring and motivational work environment. In addition, I would like to thank my fellow researchers, especially those in the Software Engineering Research Group for sharing valuable experiences and supporting me towards my PhD.

My last, but not least, thank goes to my family. I am very thankful to my mother, Shafqat Sultana, and brothers (Usman, Ali Ufan, Hamayun) for their love, support and countless prayers. A very special thank to my wife, Humera, for her unconditional love, care and support in pursuing my PhD. A final thanks to my sons, Nameer & Ifad, for being my motivation for everything.

Sardar Muhammad Sulaman

Popular Science Summary

What if a person is sick and goes to a doctor, and the doctor cannot access that person's health record due to a malfunction in the health record system? What if a person has an emergency and wants to call the police but cannot call because of a communication system failure?
Yes, IT systems are very critical for the society and we need to make them more reliable. We present research results for improving the risk analysis process to make sure these kinds of severe failures should not happen.

IT systems have become an essential part of our modern society. This evolution has not only created new opportunities, but also new threats to our society. The presence of IT systems everywhere, e.g., in communication, transportation, health, education, and emergency services, has made us dependent on them for our daily life. This is the case both for individuals and organizations, both private as well as public. However, at the same time as the usage of, and dependence on, IT systems increases, the number of reports of problems caused by failures of critical IT systems has also increased.

One of the common aspects of these failures is the trust in systems that are not sufficiently dependable. The core of the problem is not that these systems suddenly become unreliable, but that we have become critically dependent on a wide variety of systems without analyzing whether they are dependable enough and what the consequences could be of possible failures.

To prevent critical systems from causing problems for the organizations dependent on them, risk analysis and risk management are necessary activities. Analysis of IT risks is getting more and more important because of the increasing complexity of IT systems and our dependence on them. In some countries, e.g., Sweden and the US, governmental authorities are obliged by law to regularly conduct risk and vulnerability analyses of the critical processes and operations. The US Department of Homeland Security issued national strategy documents for the protection of physical and cyber infrastructures that make risk and vulnerability assessments mandatory.

As almost all societal critical processes and operations are dependent on IT systems, this dependency requires a detailed risk analysis or management of IT systems. Risk analysis and management is a process that identifies and assesses risks and introduces countermeasures to reduce risks to an acceptable level. It is a necessary activity that protects an organization's ability to perform their critical processes and activities. A risk management process is a systematic and structured way of “forward thinking” that provides a framework to make more effective decisions about an organization or system. It helps decision makers to make well informed and prioritized decisions by choosing best from different available options.

There exist a number of risk analysis methods for analyzing IT systems but they are not empirically evaluated to any large extent. These methods are based on different theories, e.g., on reliability theory and system theory. These methods are also different in their way of analysis, e.g., top-down and bottom-up. While analyzing IT systems it is hard to decide which method that should be used that is sufficiently effective and efficient.

Therefore, there is a need to investigate existing risk analysis methods empirically.

The work carried out in this research project mainly focuses on improving the risk analysis process for IT systems in large-scale organizations. The presented work is carried out to improve risk analysis processes in the following three different ways. First, we review current literature and practices related to risk analysis of critical IT systems. Second, we evaluate and compare existing risk analysis methods. Finally, we suggest improvements in the risk analysis process and develop new effective and efficient risk analysis methods to analyze critical IT systems.

For the first part of the research objective, understanding current practices of risk analysis, we carried out a literature review. It summarizes the existing risk analysis methods and the main empirical research that has been conducted in the area of risk analysis for IT systems. Based on the findings of the literature review we conclude that there is a need for empirical investigation of risk analysis methods for analyzing IT systems by conducting case studies and controlled experiments. Furthermore, we also investigated how practitioners working with risk analysis are carrying out risk analysis and what challenges they are currently facing. By having more knowledge about this we can determine how well they are analyzing critical IT systems and how we can further improve risk analysis methods.

The second part of the research objective gives a basis for evaluation and comparison of different risk analysis methods. Since there exist different risk analysis methods, it is difficult to know which method should be used in a given situation. It is not clear what measures should be used to evaluate or compare these methods. Therefore, different measures were investigated that can be used to evaluate and compare risk analysis methods. Based on the results we conclude that the effectiveness and efficiency of risk analysis methods can be evaluated and compared by counting the number of relevant and non-relevant risks identified by the participants or risk analysts. Moreover, the ease of use is a suitable attribute to evaluate effectiveness and efficiency of risk analysis methods. The time efficiency is also a suitable attribute for evaluation and comparison of different risk analysis methods. The results of this research work also show that different risk analysis methods can be evaluated by comparing hazards and risk types identified by these methods. Five hazard types were defined to analyze the identified hazards to evaluate Failure Mode and Effect Analysis (FMEA) and System Theoretic Process Analysis (STPA) methods. Furthermore, the comparison and evaluation of the analysis process of these methods can also be used. In this research work, we evaluated the FMEA and STPA analysis methods by analyzing and comparing their analyses process by using a set of qualitative criteria.

The third part of the research presented in this thesis is to improve the risk analysis process. For this the use of different perspectives has been suggested and empirically assessed. A perspective is a point of view or a specific role adopted by risk analysts while doing risk analysis, i.e., system engineer, system tester, or system user. We also proposed the idea to identify and save historical information about IT incidents that can later be used for risk analysis to improve the risk analysis process.

List of Publications

This dissertation and the research work presented here is composed of an introductory section and six papers (I-VI). The introductory section is partly based on the licentiate thesis [X]. The introductory section gives an overview of the risk analysis and management field in which the work has been carried out during my PhD studies and a brief summary of the main contributions. The second part consists of six included papers that constitute my main scientific contributions.

List of Included Publications

I **A Review of Research on Risk Analysis Methods for IT Systems**
Sardar Muhammad Sulaman, Kim Weyns, Martin Höst
In *Proceedings of the 17:th International Conference on Evaluation and Assessment in Software Engineering (EASE'13)*, Porto de Galinhas, Brazil, pages 86-96, 2013.

II **Risk Analysis and Management of IT Systems: Practice and Challenges**
Sardar Muhammad Sulaman, Martin Höst
In *Proceedings of the 15:th International Conference on Information Systems for Crisis Response and Management (ISCRAM), Rochester, US, pages 831-840, 2018.*

III **Perspective Based Risk Analysis – A Controlled Experiment**
Sardar Muhammad Sulaman, Krzysztof Wnuk, Martin Höst
In *Proceedings of the 18:th International Conference on Evaluation and Assessment in Software Engineering (EASE)*, London, UK, pages 47:1-47:10, 2014.

IV **Comparison of the FMEA and STPA safety analysis methods – A case study**
Sardar Muhammad Sulaman, Armin Beer, Michael Felderer, Martin Höst
Software Quality Journal, "Online First", DOI 10.1007/s11219-017-9396-0, 2017.

V **Identification of IT Incidents for Improved Risk Analysis by Using Machine Learning**
Sardar Muhammad Sulaman, Kim Weyns, Martin Höst
In *Proceedings of the 41:st Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, Funchal, Madeira, Portugal, Pages 369-373, 2015.

VI **A Method for Assessing Resilience of Socio-Technical IT-systems**
Finn Landegren, Sardar Muhammad Sulaman, Peter Möller, Martin Höst, Jonas Johansson
In *Proceedings of the European Safety and Reliability Conference (ESREL)*, Glasgow, UK, pages 2199-2206, 2016.

In this dissertation the included papers are referred as [I], [II], [III], [IV], [V] and [VI].

Contribution Statement

Sardar Muhammad Sulaman is the first author of all included papers except Paper VI. He was the main inventor and designer of the studies I-V, and was responsible for running the research processes. He also conducted most of the writing. In Paper VI, Sardar Muhammad Sulaman contributed in designing the study objective, formulation of research questions and also participated in focus group meetings with experts to collect data.

Paper I

The systematic mapping study reported in Paper I was co-designed and carried out with Dr. Kim Weyns and Prof. Martin Höst, but Sardar Muhammad Sulaman wrote the majority of the paper.

Paper II

The research work presented in Paper II was planned and co-designed with Prof. Martin Höst. Sardar Muhammad Sulaman carried out all the interviews for data collection except one, which was in Swedish and carried out by Prof. Martin Höst. Furthermore, Sardar Muhammad Sulaman transcribed and analyzed all the collected data for the study and then he wrote majority of the study with assistance from Prof. Martin Höst.

Paper III

The experiment in Paper III was conducted by Sardar Muhammad Sulaman and Dr. Krzysztof Wnuk. The study was co-designed with Prof. Martin Höst, although

Sardar Muhammad Sulaman was responsible for the design, carrying out the experiment, collection of data, and analysis of the collected data. Sardar Muhammad Sulaman wrote the majority of the paper, with assistance from Dr. Krzysztof Wnuk and Prof. Martin Höst.

Paper IV

Paper IV is a continuation of a study [VII], which started as a project report of a PhD course (Safety Critical Software-Intensive Systems). The previous study [VII] was mainly designed and carried out by Sardar Muhammad Sulaman. In Paper IV, Sardar Muhammad Sulaman was responsible for carrying out the hazard analysis using System Theoretic Process Analysis (STPA) and he wrote the majority of the paper. The second author, Armin Beer, was responsible for carrying out the hazard analysis using Failure Mode and Effect Analysis (FMEA) and also for writing the relevant parts in the study. Furthermore, Dr. Michael Felderer and Prof. Martin Höst contributed by discussions and feedback about the application of hazard analysis methods and by reviewing hazard analysis results.

Paper V

The work presented in Paper V was co-designed with Dr. Kim Weyns and Prof. Martin Höst, however Sardar Muhammad Sulaman carried out all the experimental work and also wrote the majority of the paper.

Paper VI

In Paper VI, the first author of the study, Finn Landegren, designed and developed the simulation model and analyzed the collected data. However, Sardar Muhammad Sulaman participated in designing the study objective, formulation of research questions and also participated in focus group meetings with experts to collect data about IT systems and core network to assess their resilience. Sardar Muhammad Sulaman also contributed in writing and reviewing the paper.

List of Related Publications

The author of this dissertation has also contributed to the following publications. However, these publications are not included in this thesis.

VII **Hazard Analysis of Collision Avoidance System using STPA**
Sardar Muhammad Sulaman, Taimoor Abbas, Krzysztof Wnuk, Martin Höst
Short paper in *Proceedings of the 11:th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, Penn State University, Pennsylvania, USA, pages 424-428, 2014.

VIII **Development of Safety-Critical Software Systems Using Open Source Software – A Systematic Map**
Sardar Muhammad Sulaman, Alma Oručević-Alagić,
Markus Borg, Krzysztof Wnuk, Martin Höst, Jose Luis de La Vara
In *Proceedings of the 40:th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, Verona, Italy, pages 17-24, 2014.

IX **Mapping and Scheduling of Dataflow Graphs – A Systematic Map**
Usman Mazhar Mirza, Mehmet Ali Arslan, Gustav Cedersjö,
Sardar Muhammad Sulaman, Jörn W. Janneck
In *Proceedings of the 48:th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, pages 1843-1847, 2014.

X **Improving Risk Analysis Practices in Governmental Organizations**
Licentiate Thesis, Software Engineering Research Group (SERG), Department of Computer Science, May 2015, Lund University, Sweden.

XI **An Industrial Case Study on Measuring the Quality of the Requirements Scoping Process**
Krzysztof Wnuk, Markus Borg, Sardar Muhammad Sulaman
In *Proceedings of the 17:th International Conference on Product-Focused Software Process Improvement (PROFES)*, Trondheim, Norway, pages 487-494, 2016.

CONTENTS

Introduction **1**

1 Concepts and Definitions 3
2 Related Work 8
3 Research Overview 13
4 Research Method 15
5 Summary Of The Included Papers 20
6 Synthesis 24
7 Validity Assessment 36
8 Conclusions and Future Work 38
Bibliography 41

Included Papers **49**

I A Review of Research on Risk Analysis Methods for IT Systems **51**
1 Introduction 52
2 Related Work 53
3 Methodology 53
4 Validity Assessment 60
5 Results 62
6 Discussion 67
7 Conclusions 69
Bibliography 76

II Risk Analysis and Management of IT Systems: Practice and Challenges **79**
1 Introduction 80
2 Background and Related Work 80
3 Research Methodology 81
4 Analysis and Results 85
5 Discussions 92

6 Conclusions 93
Bibliography 94

III Perspective Based Risk Analysis - A Controlled Experiment 97
1 Introduction 98
2 Related Work 98
3 Experimental Design 99
4 Experiment Execution 109
5 Results 110
6 Analysis 111
7 Validity Evaluation 114
8 Discussion 116
9 Conclusions and Future Work 117
Bibliography 119

IV Comparison of the FMEA and STPA safety analysis methods – A case study 123
1 Introduction 124
2 Background 125
3 Related Work 132
4 Case Study Design 134
5 Results 138
6 Analysis 144
7 Validity Evaluation 152
8 Discussion 156
9 Conclusions and Future Work 158
Bibliography 160

V Identification of IT Incidents for Improved Risk Analysis by Using Machine Learning 165
1 Introduction 166
2 Related Work 167
3 Research Methodology 168
4 proposed IT Incident Identification Method 169
5 Results and Discussions 172
6 Conclusions and Future Work 174
7 Acknowledgement 175
Bibliography 176

VI A method for assessing resilience of socio-technical IT-systems 179
1 Introduction 180
2 Related Work 181
3 Research Methodology 182
4 Simulation Model 183

5 Case Study 186
6 Result 188
7 Discussion 192
8 Conclusions 194
Bibliography 195

INTRODUCTION

IT systems have become a critical part of our modern society. This evolution has not only created new opportunities, but also new threats to our society. The presence of IT systems everywhere has made us dependent on IT systems for our daily life. This is the case both for individuals and organizations, private as well as public organizations. However, at the same time as the usage of, and dependence on, IT systems increases, the number of reports of problems caused by failures of critical IT systems has also increased [56]. The complexity of socio-technical IT systems and our dependence on them is increasing day by day. More complex IT systems contain more interacting components and sub-systems, which in turn increases the probability of serious failures [49]. Moreover, failures in these complex safety-critical systems are often results of multiple interacting decisions and errors [46].

One common aspect of these failures is the trust in systems that are not sufficiently dependable. The core of the problem is not that these systems suddenly become unreliable, but that we have become critically dependent on a wide variety of systems without analyzing whether they are dependable enough and what the consequences could be of a possible failure [56]. To prevent critical systems from causing problems for the organizations dependent on them, risk analysis is a necessary activity. Analysis of IT risks is getting more and more important. In some countries, e.g., Sweden and the US, governmental authorities (central or local) are obliged by law to regularly conduct risk and vulnerability analyses of the critical processes and operations [67, 68, 73]. The US Department of Homeland Security issued national strategy documents [73] for the protection of physical and cyber infrastructures that make risk and vulnerability assessments mandatory. As almost all societal critical processes and operations are dependent on IT systems, this dependency requires a detailed risk analysis or management of IT systems.

Risk management (RM) is a process that identifies and assesses risks, and introduces countermeasures to reduce risks to an acceptable level. It is important to protect an organization's ability to perform their critical processes and activities along with its assets. A risk management process is a systematic and structured way of 'forward thinking' that provides a framework to make more effective decisions about an organization or system. It helps decision makers to make well informed decisions. Management of risks helps in increasing opportunities and decreasing threats to an organization or system [30, 31].

A well-structured risk management methodology can help an organization's management to identify appropriate measures for providing the mission specific control capabilities. ISO 31000 [31] suggests a few principles for effective risk management in an organization. For example, risk management should be an integral part of an organization. It should be a part of the responsibilities of management and an integral part of all organizational processes. It should be an embedded activity of an organization's culture and practices. It should be dynamic, iterative and responsive to changes. An organization should allocate appropriate resources for the risk management. Finally, an organization should develop and implement strategies to improve their risk management maturity with its other aspects. Risk management requires some input, which usually is in the form of historical information about different incidents, expert opinions, and user's or employee's feedback, observations, and experience. Based on such information, risk analysts and managers can forecast or predict potential risks or events that have negative effects. Input to risk management is not always accurate and is based on that one can not foresee accurate future risks.

The increasing complexity of IT systems and our dependence on them put additional pressure on the effectiveness of risk analysis methods. The complexity, size, and heterogeneity of today's IT systems demand for effective and efficient risk analysis methods [46]. There exist a number of risk analysis methods for analyzing IT systems but they are not empirically evaluated to any large extent [I]. While analyzing IT systems it is hard to decide which method is sufficiently effective and efficient and should be used. Therefore, there is a need to investigate existing risk analysis methods empirically. The work presented in this thesis mainly focuses on improving the risk analysis process for IT systems. This research work addresses the main research objective in the following three different ways. First, by understanding current practices related to risk analysis of IT systems and how they can be improved then by investigating different measures that can be used to evaluate and compare existing risk analysis methods and finally, by suggesting improvements in existing risk analysis methods and developing new effective and efficient risk analysis methods to analyze critical systems.

The outline of this thesis is as follows. Part I consists of the introduction section that presents an overview of the risk analysis and management field with a brief summary of the research methods and main scientific contributions presented in this thesis. Part II presents the papers that are included in this thesis.

In part I, Section 1 presents some basic terms and concepts with their interpretations used in this thesis. Section 2 presents related work in the field of risk analysis and management. Section 3 presents the research objective with the research questions. Section 4 discusses the research methods used. Section 5 presents the summary of the included papers in this thesis. Next, Section 6 synthesizes the results of the research carried out in this thesis. Section 7 assesses the validity threats. Finally, Section 8 concludes the results of this thesis and presents agenda for the future research work.

1 Concepts and Definitions

This thesis uses a few concepts that are interpreted differently in different contexts. To avoid confusion, their exact definitions used in this thesis are clarified in this section.

1.1 Risk, Uncertainty, Likelihood and Consequence

In pre-industrialization societies there were no risks because during that time uncertainties of everyday life were expressed, predicted and managed by religious and magical beliefs e.g. fate, providence and luck [60]. In reality there were risks but people did not think in terms of risk. In the medieval time, these beliefs were considered a part of cosmology in which all earthly events and fate of every individual were depicted as a symbolic representation of the will of God [23, 60]. From around the turn to the 17th century, rapid development in social, economic and intellectual life changed the way of thinking about uncertainty, future prediction by laying the grounds for risk notion. In mid of 17th century, the term risk was used first time by bringing uncertainty and time into a quantifiable relation. The insurance domain (guarding one against uncertainty) developed and evolved in parallel with the term risk. The main idea behind insurance was to protect individuals from the threats of newly predictable events that were not possible to predict before [17]. The term risk was used to predict that something *might* happen instead of something *will* happen [60].

Today, *risk* is a commonly used term and everyone thinks and talks about risk in their daily life. We all analyze risks in our daily life, for example while crossing roads and driving, but that analysis is not systematic. Sometimes the word *risk* is used to describe the *likelihood* of an event, for example “there is a risk of rain today” but in the risk management context, *risk* is the likelihood of an event combined with its potential impact.

There is no general definition for risk. According to ISO 31000 [31] *risk* is the effect of *uncertainty* on *objectives* and an effect is a positive or negative deviation from what is expected. *Uncertainty* (or lack of certainty) is a state or condition that involves a deficiency of information and leads to inadequate or incomplete knowledge or understanding. In the risk management context, uncertainty exists whenever knowledge or understanding about an event, consequence, or likelihood is inadequate or incomplete. *Risk* is defined in the Merriam-Webster dictionary[1] as the “possibility of loss or injury” and *hazard* as a “source of danger”. Hazard, therefore, simply exists as a source [35].

[1] http://www.merriam-webster.com/dictionary

The definitions used in this thesis are that, *risk* is the chance that an undesired/negative event *might* happen with some consequences or impact. *Hazard* is a situation with potential danger to people, environment, or material. *Failure* is the inability of a component or system to perform its intended function [46]. *Likelihood* is the chance that something might happen. It can be determined, measured or expressed subjectively or objectively (quantitatively, qualitatively or semi-quantitatively). *Consequence* is the outcome of an event and has an effect, positive or negative, on objectives or assets. A single event can be a cause of many consequences with both positive and negative effects on organization's objectives [31].

Risk analysis can be performed during development of a system or at any time afterwards. In the ideal situation, the risk analysis should be repeated each time major changes occur in the system or in the environment in which the system is used. This thesis mainly focuses on risk analysis methods for *operational* IT systems that do not include risk analysis methods used to analyze the project management risks in software development projects.

For managing IT system risks, one important step is to define the *scope* of the system. The *scope* of the system contains the identification of *system boundaries* along with the components and the information that constitute the system.

1.2 Risk Management

Risk management is a coordinated set of activities that are used to direct an organization to control risks that can affect its ability to achieve objectives [31]. It is considered as an important element of good management and improved decision making for an organization, especially regarding its delivery of critical services to the society. More or less each department of an organization identify and manage their potential risks iteratively, however sometimes more rigorously and systematically and sometimes less [5]. Furthermore, risk management provides assurance that a critical system in an organization has met its stated safety and security properties, and that the system can be depended upon to deliver its intended services in a safe manner. According to the ISO 31000 standard [31], risk management consists of five steps as shown in Figure 1. The first three steps (context establishment, risk assessment, and risk treatment) are carried out in series but steps 4 and 5 (monitoring and review, and communication and consultation) are continuously carried out in parallel during the whole risk management process.

The first step of risk management is *context establishment*, which helps an organization to define its objectives with external and internal parameters. The scope or boundaries of the organization or system being analyzed is/are also defined in context establishment. External or internal parameters are the definition of the key drivers, external stockholders, goals and objectives of risk management activities, responsibilities, scope, risk assessment methodologies, relationship with other projects or systems, and required resources etc. Then, the risk criteria are defined at this stage which are used later to evaluate the significance of identified

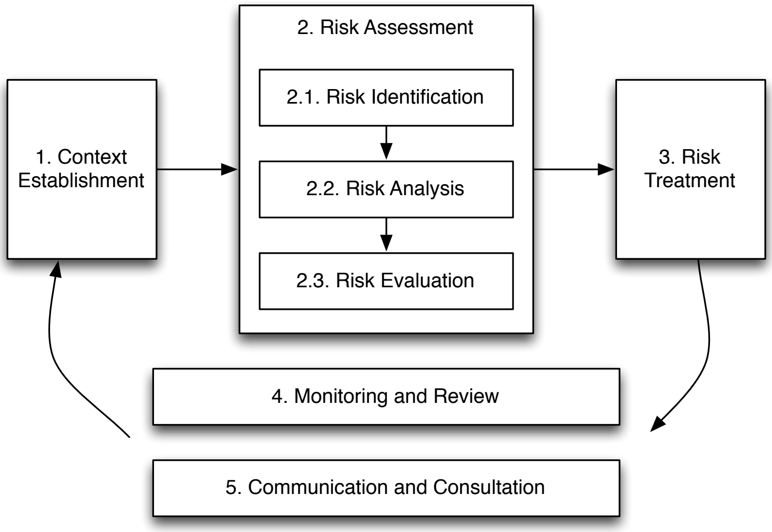


Figure 1: The risk management process according to [31]

risks. The defined risk criteria must be based on the objectives and the resources of organization. The risk criteria consist of several factors [31]. The most important factors are the nature and type of risks with their causes and consequences that can occur and how they will be measured. Then the estimation (qualitative or quantitative) of the likelihood and levels of identified risks are defined with the definition of acceptable risks.

Risk assessment is an important step of the risk management process. It usually deals with analysis of a system with existing measures or controls and anticipates the weaknesses present in assessed organization or system [31]. Normally, risk assessment consists of risk identification, risk analysis and risk treatment activities. *Risk identification* helps risk analysts to identify potential risks with their sources. Risk identification involves several activities, such as identification of assets, existing controls, threats, and vulnerabilities. It is recommended that in this step all risks should be identified and considered whether or not their sources are under the control of organization [31]. As an input it requires detailed information about organizations or systems being analyzed. Different people with different perspectives and knowledge should also be involved in the whole risk management process, especially in this activity [III]. *Risk analysis* helps to develop understanding of the identified risks. Later this understanding is used as an input to risk evaluation and risk treatment. It takes into account the causes and sources of potential risks with their consequences. It helps to analyze identified risks by determining consequences and their likelihood. Risk analysis can be done both qualitatively and quantitatively. By using the output of risk analysis, different levels are assigned to all identified risks and then prioritized risks are treated first. The *risk evaluation* step evaluates identified and analyzed risks for their treatment, i.e., which risks have highest priority to be treated. In this step the levels of identified risks must be compared with the risk criteria established earlier during the context establishment process. Based on the comparison further decisions are made to treat risks.

The *risk treatment* activity helps to modify or treat identified risks by selecting and implementing new controls and countermeasures from available different options. Selection of an appropriate treatment or control depends on the balance between the required cost and effort for implementation and the gained benefits from it. Here, it is important to consider all direct and indirect costs and benefits to estimate right balance. Risk management as a whole process addresses the balance between the cost and benefit to increase opportunities. There are different risk treatment options to select from, such as mitigating a risk by not starting that specific activity or by removing a risk source, modifying a risk by changing its likelihood and consequence, sharing a risk with third parties, and retaining the risk by an informed decision [31].

The *risk monitoring and review* activities are carried out in parallel with all other activities of risk management. They ensure that the selected risk treatments are effective. The risk management process should be traceable and for that it should be recorded or documented. *Communication and consultation* with all, internal and external, stakeholders are important activities for the whole risk management process. Therefore, a detailed plan for communication and consultation should be developed at an early stage with the context establishment.

1.3 Types of Risk Analysis

There are mainly two types of risk analysis methods, *quantitative* and *qualitative* [I]. However, there also exists one more type that is the combination of both quantitative and qualitative types, which is called *semi-quantitative*.

Quantitative analysis expresses the probability and consequences of the identified risk as numerical results. This makes it possible to calculate the relationship between loss prevention and cost associated with proposed countermeasures. Often it is difficult to use quantitative risk analysis because it is hard to estimate the exact probability and loss associated with each risk in numbers.

Qualitative analysis, on the other hand, uses descriptive values such as ‘high’, ‘medium’ or ‘low’ to express the probability and consequences of each risk. Both types of risk analyses are widely used for different types of systems, and in some cases they can be used together.

Semi-quantitative analysis, which is intermediary risk analysis technique that classifies the probability and consequences by using quantitative categories such as ‘financial loss between 10.000 USD and 100.000 USD’ or ‘less than once per

100 years'. They do not require the exact estimates needed for a quantitative risk analysis, but offer a more consistent approach than qualitative risk analysis.

1.4 IT Systems

IT System is a combination of hardware, software, data-bases, infrastructure and IT support organized to facilitate decision making in an organization. Hardware includes physical components such as hard drives, processors, and input and output devices. Software consists of the operating system, compilers and applications. Infrastructure means communication channels such as wireless connections, network cables and telephone lines. Databases save interrelated data used by different application software. Finally, IT support consists of help facilities provided for the proper functioning of IT system, such as IT support personnell, manuals, documentation or trainings [74]. It can be defined as:

"An information system is a set of interrelated components that collect, process, store, and distribute information to support decision making, coordination and control in an organization. In addition, it also helps management to analyze problems and visualize complex subjects" [44].

Critical IT systems are systems that provide or support critical services to the society, e.g., water supply, power supply, transportation, and failures of these systems have serious and negative effects on society. The Swedish Civil Contingencies Agency (MSB) has given examples of critical services or infrastructures [40]:

- Telecommunication
- Data communication
- Electrical power supply
- Health care
- Water supply and district heating
- Provision of fuels
- Transport and distribution
- Police services, emergency management
- Financial services
- Critical governmental services

Afore-mentioned services or infrastructures are directly or indirectly dependent on IT systems and failures of these systems have direct negative effects on society. The consequences of IT system failures could be stoppage or disruption of the functions of critical services, i.e., transportation services, data communication, emergency services, etc. These critical services are dependent on each other. For example, if one service (transportation or electrical power) stops working it will directly or indirectly affect other services (emergency services, postal, data communication, water supply, etc.). Therefore, risk analysis of these IT systems is very important for proper functioning of societal critical services.

1.5 Crisis Management

Crisis management is a systematic process that deals with the preparations and response to a crisis situation. The Swedish Civil Contingencies Agency (MSB), defines a crisis situation[2] as [74],

> *"an event that affects many people and threatens the basic values and functions of society. A crisis[3] is a condition that can not be handled with normal resources and organization. Resolving a crisis requires coordinated action by several actors."*

Crisis management is normally divided into four main activities such as, mitigation, preparedness, response, and recovery. The mitigation and preparedness activities of crisis management are carried out before the happening of a crisis situation. The response and recovery activities of crisis management are carried out during or after a crisis situation. The *mitigation* activity attempts to reduce the likelihood and/or consequences of unwanted/undesired events. The *preparedness* activity deals with the development of an emergency plan. These two activities involve risk and vulnerability analysis to estimate likelihood and/or consequences of unwanted events that help to develop an emergency plan. The *response* activity of crisis management consists of emergency actions and resources that help to mitigate or decrease the effects of crisis on society. After a crisis situation, the *recovery* activity deals with the restoration of society to its normal or desired situation [34, 74].

2 Related Work

This section presents research literature related to the work carried out in this thesis. This thesis mainly presents research in three parts i.e. 1) understanding current research literature and practice about risk analysis, 2) evaluation and comparison of different risk analysis methods, and 3) improvements in the risk analysis process.

[2] https://www.msb.se/en/About-MSB/Crisis-Management-in-Sweden/

[3] http://www.krisinformation.se

2.1 Existing Risk Analysis Standards, Methods and Practices

There exist different national and international high-level frameworks for information technology risk management and assessment. Such frameworks have for example been published by the International Organization for Standardization (ISO), such as ISO/IEC 27005 [30] and ISO/IEC 27002 [29], by national governmental organizations, such as the National Institute of Standards and Technology (NIST) [70] or the British Central Communication and Telecommunication Agency (CCTA) [9], by non-governmental organizations such as Club de la Sécurité del' Information Français (CLUSIF) [50] or by research institutes such as the Carnegie Mellon Software Engineering Institute (SEI) [3]. A detailed comparison of some of these frameworks is conducted by ENISA [14] and Syalim et al. [72].

There also exist a number of low-level risk analysis methods for technical systems in general or for IT systems in particular [I]. Some of the most well-known methods are Fault Tree Analysis (FTA) [15], Failure Mode and Effect Analysis (FMEA) [52] and Hazard and operability study (HAZOP) [58]. Some of the frameworks mentioned above specifically recommend one or more of these risk analysis methods. FTA, FMEA, and HAZOP risk analysis techniques are considered the most commonly used.

FTA is a top-down risk or hazard analysis method. It uses a deductive approach and carried out by repeatedly asking: how can this (a specific undesirable event) happen? and what are the causes of this event? It consists of a logical diagram that shows the relation between the system components and their failures. Ericson [15] presented a review of the research performed on FTA with its advantages and shortcomings.

FMEA is a risk and hazard analysis technique that can be applied as both a top-down and a bottom-up approach [52]. The top-down approach (usually function oriented) is mainly used in an early design phase before deciding the whole system structure. The bottom-up approach is used when a system concept has been decided. Moreover, as a bottom-up approach, FMEA can augment or complement FTA and identify many more causes and failure modes. Grunske et al. [22] introduced an extension to conventional FMEA, named probabilistic FMEA. It has the advantage of formally including rates at which component failures can occur. This method helps safety engineers to formally identify if a failure mode occurs with a probability higher than its tolerable hazard rate.

HAZOP is a qualitative risk analysis technique commonly used in the planning phase in system development. It identifies risks by analyzing how a deviation can arise from a design specification of a system. It is used to identify the critical aspects of a system design for further analysis. It can also be used to analyze an operational system. A multi-disciplinary team of 5 to 6 analysts lead by a leader usually carries out the HAZOP analysis. The HAZOP team identifies different scenarios that may result in a hazard or an operational problem, and then their causes and consequences are identified and analyzed [49].

Leveson [46] proposed a hazard analysis technique, named System Theoretic Process Analysis (STPA) that considers safety as a control problem rather than a component failure problem. It focuses on analyzing the dynamic behavior of system and therefore provides significant advantages over the traditional hazard analysis methods. STPA is a top-down method, just like the FTA method. However, STPA uses a model of the system that consists of a functional control diagram instead of a physical component diagram [47].

2.2 Evaluation and Comparison of Risk Analysis Methods

There exist some studies that have evaluated and compared different risk and hazard analysis methods. For example, Stålhane and Sindre [71] performed a comparison of two safety analysis methods, MisUse Case (MUC) method and FMEA. The MUC method was originally proposed for eliciting security requirements [69], but it has also been used for safety analysis. The MUC method was developed by the software community as an alternative to FMEA and HAZOP. Both methods were compared in an experiment to investigate which method is better than the other for identifying failure modes and if one of the methods was easier to learn and to use. The authors concluded that when the system's requirements are described as use cases, MUC is better than FMEA for analyzing failure modes related to user interactions. Furthermore, FMEA is better than MUC for analyzing failure modes related to the inner working of the system. The authors also concluded that MUC will create less confusion and in general be easier to use than FMEA.

Yu et al. [78] compared and discussed three well known risk analysis methods by applying them on a box fan, FMEA, AFMEA (Advanced Failure Mode and Effect Analysis) [16], and FTA. The authors presented the advantages and disadvantages of these methods and concluded their study with an attempt of combining both deductive (top down) and inductive (bottom up) risk/safety analysis methods.

Abdulkhaleq and Wagner [1] performed a controlled experiment with 21 graduate and undergraduate students to compare three safety analysis techniques (FTA, FMEA and STPA) with regard to their effectiveness, applicability, understandability, ease of use and efficiency in identifying software safety requirements at the system level. The authors concluded that STPA seems to be an effective method to identify software safety requirements at the system level. In particular, STPA addresses more different software safety requirements than the traditional techniques FTA and FMEA. However, the authors did not find any statistically significant difference in the applicability, understandability and ease of use of the three techniques. The authors also mentioned that STPA requires more time to carry out an analysis by safety analysts with little or no prior experience.

Ishimatsu et al. [28] compared the STPA hazards analysis results with the FTA analysis results that were used to certify the H-II Transfer Vehicle (HTV). The

HTV is an unmanned cargo transfer spacecraft that is launched from the Tanegashima Space Center aboard the H-IIB rocket and delivers supplies to the international space station (ISS). In the development of the HTV the potential HTV hazards were analyzed using FTA and during the analysis the NASA safety requirements were also considered. After comparison of the results, the authors concluded that STPA identified all the traditional causes of losses identified by FTA and FMEA, but it also identifies additional causes. The additional factors include those that cannot be identified using fault tree analysis, including software and system design as well as system integration.

Fleming et al. [18, 19] analyzed the NextGen In-Trail Procedures (ITP) application by using the STPA analysis method and compared its results with the official NextGen ITP application analysis [62]. NextGen is the next generation of air traffic management systems that contains In-Trail Procedures application. ITP is an application of Automatic Dependent Surveillance-Broadcast (ADS-B) that allows aircraft to change flight levels in areas where current radar separation standards would prevent desirable altitude changes [24]. To summarize, ITP helps to increase operational efficiency and throughput in oceanic airspace [19]. The authors concluded that STPA found more potential causes of the hazards considered (violation of separation requirements) than the traditional hazard analysis performed on ITP [62]. In the comparison, the authors identified 19 safety requirements that were not in either of the two official NextGen analysis documents.

Moreover, Fleming et al. [18, 19] also compared STPA with bottom-up and other top-down analysis techniques. According to the authors, bottom-up analysis methods e.g. FMEA, start by identifying all possible failures. This list can be very long if there are a lot of components and all the permutations and combinations of component failures are considered. However, STPA only identifies the failures and other causes that can lead to a system hazard and does not start by identifying all possible failures. Moreover, in the top-down STPA analysis approach, the analyst can stop refining causes at the point where an effective mitigation can be identified and does not go down any further in detail. The analyst only has to continue refining causes if an acceptable mitigation cannot be designed. That is the major difference between STPA and FMEA (and any other bottom-up technique), which explains the differences in time and effort required [18, 19].

Furthermore, Nakao et al. [55] evaluated STPA in a case study where it was applied on an operational crew-return vehicle design. The authors conclude that with STPA it is possible to recognize safety requirements and constraints of the system before the detailed design.

That is, it is interesting to investigate and evaluate the main differences in traditional and new methods e.g. STPA, and Perspective Based Risk Analysis (PBRA) and also the types of hazards identified by them.

2.3 Improvements In The Risk Analysis Process

To improve the risk analysis process, researchers and practitioners have introduced some improvements in the current risk analysis practices. For example, to tackle the lack of information in early design problem, Johannessen et al. [33] proposed an actuator-based approach for hazard analysis. This approach is a logical approach for an early hazard analysis when only basic or limited information about the system is available. Such an approach is beneficial as major hazards can be identified in an early stage based on their criticality. Gleirscher [20] suggested a framework for hazard analysis for software-intensive control parts of technical systems, and exemplified on a commercial road vehicle in its operational context.

Yoran and Hoffman [77] proposed the Role-Based Risk Analysis (RBRA) method that defines roles and identifies actors before performing risk analysis activities in order to reduce the set of vulnerabilities and controls to those appropriate to a given role. RBRA was presented on an illustrative example from the computer software engineering domain but not experimentally investigated. Leveson [46] and McDermott et al. [52] advocated to involve various perspectives during risk analysis, also from external organizations. The idea of using perspectives is not new, it is always recommended, in almost all risk analysis methods, to have experts with domain knowledge while performing risk analysis. Perspectives were utilized for reading software engineering artifacts with the purpose of improved defect identification [4, 59]. Perspective-based reading was also applied for object oriented design inspections [64], code reviews [41] and usability inspections [79]. Different perspectives, e.g., developers, testers and domain experts are often involved in requirements elicitation. This results in increased quality of elicited requirements and often uncovers new requirements based on various views and perspectives.

In this thesis, we presented an improvement in the risk analysis process as a prototype of a system that saves historical risk information to be used later in risk analysis [V]. Here, the research presented is carried out using the text classification and information filtering techniques. A number of studies have discussed text classification in general and presented results by using different machine learning algorithms. For example, Sebastiani et al. [66] present an overview of different available machine learning approaches for automatic text classification. In the study, the authors discuss different methods, their applications, their effectiveness and recent progress that has been made in the field.

To suggest improvements in the current practices of risk analysis and management for IT systems we carried out an exploratory study [II]. There are some studies that discuss and present the current practices of risk analysis and management in different domains [25, 51, 54] but not for IT systems. For example, Murdock et al. [54] present the lessons learned and best practices for risk management for developing an enterprise-wide risk management framework. Henschel presents the current state of risk management practices in German SME:s through the studies based on questionnaires and interviews [25].

To summarize, most of the presented research on available methods and frameworks for risk analysis and management is of normative nature that guide how one should carry out risk analysis and management. However, it is not sufficiently investigated how different organizations are actually carrying out risk analysis and management. There is a need for empirical investigations about the current practices of risk analysis and management for IT systems in large-scale organizations. This way, the risk analysis process can be improved further by knowing more about the current practices and existing challenges.

3 Research Overview

The research presented in this thesis was carried out as a part of PRIVAD, Program for Risk and Vulnerability Analysis Development, program funded by the Swedish Civil Contingencies Agency (MSB). The overall objective of the PRIVAD program is to develop tools and methods to improve risk and vulnerability assessments at all levels of society. The research objective of this thesis is to evaluate and improve the analysis process of risks pertaining to IT systems in large-scale organizations. To address the main research objective, the research in this thesis is carried out in the following three ways.

First, the research is carried out to understand current literature and practices related to risk analysis of critical IT systems. There exist a number of risk analysis methods for technical systems consisting of mechanical parts. However, the failure behavior of information systems is typically very different from mechanical systems. Therefore, risk analysis of IT systems requires different risk analysis techniques, or at least adaptations of traditional methods. This means that there is a need to understand what types of methods are available for IT systems and how they can be improved.

Then, the research is carried out to evaluate and compare existing risk analysis methods. As there exist a number of risk analysis methods, there are still a number of uncertainties when it comes to what risk and hazard analysis method to apply in a given situation. By comparing and evaluating different existing risk analysis methods empirically it can help in deciding which method should be used in a given situation.

Finally, the research is carried out to suggest improvements in the risk analysis process and to develop new effective and efficient risk analysis methods to analyze IT systems. As we know, the dependence on IT systems in large-scale organizations is very crucial. Failures of these systems or services have serious and negative effects on society. In general, governmental organizations are responsible for delivery of these services to society. Therefore, analyzing risks of IT systems and later mitigating identified risks decreases potential threats that are faced by the society.

3.1 Research Questions

The research objective is divided into the following more detailed research questions:

RQ1: What is the current state of the risk analysis research and practice?

RQ1.1: What risk analysis methods and approaches exist for analyzing IT systems? Is there any empirical research that compares or evaluates existing risk analysis methods?

RQ1.2: What are the current practices of risk analysis and management for IT systems in large-scale public organizations and what are the main challenges in carrying out risk analysis?

RQ2: How can we evaluate the effectiveness and efficiency of a risk analysis method?

RQ2.1: How can we evaluate and compare different risk analysis methods and what comparative attributes of the risk analysis methods should be used for this?

RQ2.2: How can we assess the resilience of critical IT systems and networks that can help to determine how dependable a typical system or network is?

RQ3: How can we improve the risk analysis process?

RQ3.1: Can the use of different perspectives in risk analysis improve the risk analysis process?

RQ3.2: How can we identify and save historical information about IT incidents to improve the risk analysis process?

RQ3.3: How can the risk analysis practices be improved in large-scale governmental organizations?

RQ1 summarizes the existing risk analysis methods and the main empirical research that has been conducted in the area of risk analysis for IT systems. Furthermore, it will also give an idea about how practitioners working with risk analysis are carrying out risk analysis and what are the challenges that they are currently facing. By having more knowledge about this we can determine how well they are analyzing critical IT systems and how we can further improve risk analysis methods.

RQ2 gives a basis for evaluation and comparison of different risk analysis methods. Since there exist different risk analysis methods, it is difficult to know which method should be used in a given situation. When it comes to evaluation and comparison of exiting risk analysis methods to know which method is better then it is not clear what measures should be used to evaluate or compare these methods. Therefore, it is important to investigate different measures, both quantitative and qualitative, that can be used to evaluate and compare different risk analysis methods. It is interesting to investigate because it can help to determine how dependable an IT system is.

RQ3 is addressed by suggesting improvements in the risk analysis process in the following ways. To improve risk analysis and management, the use of different perspectives has been suggested [2,6,30,46,52,70,77] but it is not empirically assessed. Therefore, it is important to empirically assess the potential of different perspectives in risk analysis and management processes. After this, another potential improvement is suggested to answer RQ3 by an evaluative investigation. Then, based on the investigation of the current practices of risk analysis and management of IT systems (RQ1) a set of potential improvements are presented and validated to answer RQ3.

4 Research Method

The research presented in this thesis is based on *empirical research*, which is a way to obtain knowledge through observation and measurement of a phenomenon. The research questions in empirical research are related to the class of knowledge questions, i.e. the questions are focused on the observable and measurable state of the world [12]. This section gives an overview of the research methods applied in our empirical studies.

Research studies have different research objectives and aims, and there is not a single research strategy that fits them all. Runeson et al. [63] list four purposes for research in software engineering, adapted from Robson [61]:

- **Exploratory** – discovering what is happening, pursuing new insights, and generating ideas and hypotheses for future research.
- **Descriptive** – characterizing the current status of a phenomenon or situation.
- **Explanatory** – seeking an explanation for a phenomenon, often in the form of a casual relationship.
- **Improving** – attempting to improve certain aspects of a phenomenon, and to evaluate the effect of improvement proposals.

The papers included in this thesis are either exploratory or of improving and evaluative nature.

In *exploratory* research the aim is to understand, with more or less prejudice, a specific phenomenon [61]. It is typically carried out in early stages of research projects and tries to achieve initial understandings of a phenomenon, usually from rich qualitative data [13]. Exploratory research is commonly used to find research gaps and to guide further research. It helps to design future studies with their data collection methods and sample selections. Papers I and II are both of exploratory research strategy type. Paper I summarizes the risk analysis methods for IT systems presented in the scientific literature and Paper II explored the state of practice of risk analysis for IT systems in large scale public organizations.

In *improving* research the aim is to improve the current state of practice of a phenomenon. An important part of improving research is the evaluation. In *evaluative* research part the aim is to assess the effects and effectiveness of innovations, interventions, practices etc. [61]. It involves a systematic collection of data, which can be of both of qualitative and quantitative type, and a rigid analysis and interpretation. Paper III presents a new risk analysis method, Perspective Based Risk Analysis (PBRA) that is of improving research strategy type, which is evaluated by a controlled experiment. Paper IV presents an evaluation and comparison of two risk analysis methods, Failure Mode and Effect Analysis (FMEA), and System Theoretic Process Analysis (STPA). Paper V proposes a method for automatic identification of relevant text that is of improving research type and then it presents an evaluation of the proposed method. Furthermore, Paper VI presents an evaluation of a method for assessing resilience of IT systems.

4.1 Empirical Research Methods

When designing an empirical study, the researcher needs to find an appropriate balance between the level of control and the degree of realism [63]. Studying phenomena in a real-world setting means less control of the involved variables, and often there are several confounding factors to conclude casual relationships. When isolating real-world phenomena on the other hand, e.g., by increasing the control i.e subjects in lab environments (controlled experiments), there is a risk that the software engineering aspect under study becomes less representative of real-world setting [13]. Here, the researcher needs to have an appropriate balance.

The data collected in an empirical research study are either of quantitative or qualitative type. Quantitative data constitute numbers obtained from measurements, and generally their purpose is to answer questions about the relationships between variables. On the other hand, qualitative data involve words, descriptions, pictures, etc. While quantitative data provide ‘exactness’, qualitative data instead offer ‘richness’ [63], which helps a researcher to understand a phenomenon beyond numbers. In software engineering research, qualitative data are often collected using interviews. Analysis of quantitative data is based on statistics, and analysis of qualitative data is carried out using categorization and sorting. Moreover, analysis of qualitative data is based on the researcher’s interpretation, therefore special measures are needed to mitigate biased conclusions. However, quantitative and qualitative data are of different nature. Therefore, research studies that use both quantitative and qualitative data reach the strongest conclusions by providing both ‘exactness’ and ‘richness’ [63, 65]. It is also called triangulation that is collecting data from multiple sources to improve validity [13].

The research process of empirical studies can be characterized as *fixed* or *flexible* design [61, 63, 76]. Fixed designs are pre-specified and require enough pre-understanding of a phenomenon to know what to do, and how to measure it, already when the study is initiated. Studies relying on quantitative data are often of a fixed design. Flexible designs on the other hand, allow the study design to evolve while data is collected. The collection and analysis of data is intertwined, and both research questions and data sources may be adapted to the circumstances of the study.

Empirical research studies, both fixed and flexible, can be carried out using different research methods. Easterbrook et al. [13] identified five research methods as the most important and applicable in software engineering:

- **Experiments** – testing hypotheses by manipulating independent variables and measuring the effect on dependent variables.
- **Case studies** – investigating a contemporary phenomenon with in its real-life context.
- **Surveys** – identification of population characteristics by generalizing from a sample.
- **Ethnographies** – understanding how a community of people make sense of their social interactions.
- **Action research** – attempting to solve a real world problem by intervention, while simultaneously studying the experience of solving the problem.

In this thesis we did not use the ethnography and action research methods. The other three methods *‘experiment’* used in Paper III, *‘case studies’* used in Paper IV-VI and *‘survey’* used in Paper II are further described in the remainder of this section. Moreover, this thesis uses the *‘qualitative survey’* research method instead quantitative survey as discussed by Jansen [32]. This thesis also uses the *‘focus group’* research method for validation of the suggested improvements for the risk analysis process that is also briefly described in the remainder of this section. Furthermore, Paper I presents a systematic mapping study that is carried out by using the *‘systematic literature review (SLR)’* research method, which is also described briefly later in this section.

An *experiment* (or controlled experiment) is a commonly used research method in software engineering research to investigate the cause-effect relationships of different methods, techniques or tools. In experimentation research different treatments are applied to, or by, different subjects, while other variables are kept constant, and the effects on outcome variables are measured [75].

A *case study* in software engineering is conducted to understand a phenomenon within its real-life context. Such a study draws on multiple sources of evidence to investigate one or more instances of the phenomenon, and the research method is especially applicable when the boundary between the phenomenon and its context cannot be clearly defined. According to Runeson et al., the case under study can be any contemporary software engineering phenomenon in its real-life setting, e.g., a group of people, a process, a specific role, or a technology [63]. Case studies are often conducted to explore a phenomenon, but they can also be confirmatory, i.e., designed to test existing theories [13]. In a case study, researchers also distinguish between the case(s) and the unit(s) of analysis [76]. In a holistic case study, the case is studied as a whole. On the other hand, in an embedded case study multiple units of analysis are studied within a case, e.g., different development sites, teams, or roles in an organization. Moreover, a case study can be characterized as a single-case study, or a multiple-case study if two or more cases are studied within different contexts [63].

The basic idea of the *survey* method is to collect information from a group of people by sampling individuals from a large population. Since the method relies on sampling, it is typically carried out by first planning and then carrying out the study according to the plan (i.e. a 'fixed design' according to Robson [61]). According to Easterbrook [12] the survey method is applied when characteristics of a broad population of individuals need to be identified. The survey data needs to be collected from a representative sample of a well defined population. However, in this thesis (Paper II) the qualitative interview research method [7] in the form of a *qualitative survey* based on the discussions by Jansen [32] is used. According to Jansen [32], the qualitative survey analyses the diversity of member characteristics within a population as opposed to the statistical survey that analyses frequencies in member characteristics in a population.

The *focus group* sessions have a moderator in a similar way to semi-structured interviews, i.e. predefined questions were used as a guiding framework in leading the discussions. All participants of a focus group session are allowed to freely express their opinions within the scope of the targeted topics and researchers can ask follow-up questions. The suggestions for improving the risk analysis process presented in the introduction section of this thesis are validated in a focus-group meeting. According to Kontio et al. [39], the focus group method is an effective method to obtain qualitative insights and practitioners feedback. The focus group research method has been applied to obtain qualitative insights and feedback of practitioners e.g. [26].

A *Systematic Literature Review (SLR)* is a secondary study aimed at aggregating a base of empirical evidence. It is inspired by evidence-based medicine. SLRs rely on a rigid search and analysis strategy to ensure identification of a comprehensive collection of evidence related to a specific question [37]. A variant of an SLR is a Systematic Mapping Study, a literature study designed to identify research gaps and direct future research [38, 57]. Paper I is a systematic mapping study about the risk analysis methods for analyzing IT systems.

4.2 Classification Of The Included Papers

This thesis mainly contains *exploratory* and *improving (or evaluative)* empirical research, based on studies using systematic mapping study, qualitative survey, controlled experiment, case study and focus group research methods. An overview of the research methods used in the included papers with their research purposes is shown in Table 1.

Paper I presents a systematic mapping study that is carried out as exploratory research. A mapping study reviews a broader topic and classifies the primary research papers in that specific domain. It has high level (generic) research questions and include issues such as which sub-topics have been addressed, what empirical methods have been used. In general, it helps to find what research has been done in a specific topic area by providing an overview of the literature in that topic area [36]. On the other hand, the goal of a systematic literature review (SLR) is to analyze and aggregate the base of empirical evidence [37]. An SLR has specific research questions (related to outcomes of empirical studies) that can be answered by empirical research. It also has a focused scope and uses a stringent search strategy. Moreover, the quality evaluation of the results is very important for an SLR. Finally, unlike mapping studies, in SLR the found results are aggregated to answer specific research questions (for more details see [36] table I).

Paper II presents research that is carried out by using a qualitative interview research method [7] in the form of a qualitative survey, based on the discussions by Jansen [32]. According to Jansen [32], the qualitative survey analyses the diversity of member characteristics within a population as opposed to the statistical survey that analyses frequencies in member characteristics in a population. Based on this, Paper II focuses on the diversity in experiences of the interviewees regarding risk analysis and management rather than frequencies of their opinions.

Paper III presents results of improving and evaluative research carried out based on a controlled experiment as research method. *Experiments* (or controlled experiments) are used in software engineering research to investigate the cause-effect relationships of different methods, techniques or tools.

Paper IV presents the results of a case study that is carried out as evaluative research. It presents an evaluation and comparison of two well-known and widely used risk analysis methods, System Theoretic Process Analysis (STPA) and Failure Mode and Effect analysis (FMEA), by using qualitative measures. The used primary data was of third degree [45], which is taken from the published literature and system description of qualitative nature.

Paper V presents a case study that is carried out as improving research. The research in Paper V was initiated by an idea of automatic identification of IT incidents reported in online news sources that can later be used for risk analysis.

Table 1: Research purpose and methods used in the included papers

Work	Research purpose	Research method
Paper I	Exploratory	Systematic mapping study
Paper II	Exploratory	Qualitative survey
Paper III	Improving and Evaluative	Experiment
Paper IV	Evaluative	Case study
Paper V	Improving and Evaluative	Case study
Paper VI	Evaluative	Case study

This way, by having historical information of already happened IT incidents, risk analysis and management practices can be improved.

Paper VI presents a case study that is carried out as evaluative research. In this study, a simulation based method was evaluated by applying it on a real IT system with its core network to assess its resilience. The method for evaluation of resilience had previously been employed on an electricity network [42] and found to be beneficial for assessing resilience of electricity.

5 Summary Of The Included Papers

This section summarizes the main contributions of the work carried out in this thesis. The detailed results and conclusions can be found at the end of this thesis (appended papers).

Paper I: A Review of Research on Risk Analysis Methods for IT Systems

We present a systematic mapping study on risk analysis methods for IT systems. A mapping study identifies research gaps and clusters of evidence in order to direct future research. In an initial database search carried out on 23 May 2012, 1086 unique articles were identified. Then 57 out of 1086 papers were identified as relevant for this study. The same search query was used again on 12 April 2017 to update the mapping study by finding the newly published relevant literature between 2012-2017. It retrieved 363 articles. Then 6 out of 363 articles were identified as relevant for this study. The total number of identified articles relevant for this study became 63.

The main results of this study show that most of the discussed risk analysis methods are qualitative and not quantitative, and that most of the risk analysis methods that are presented in these papers are developed for IT systems in general and not for specific types of IT system. It is found that most articles focus on proposing new methods, frameworks and models for risk analysis. Only few papers focus on already available, and thereby maybe already known, methods.

Based on the findings of this mapping study a number of areas for further research are identified. There is a need to evaluate already available methods. This can for example be carried out as studies where different types of methods are compared in controlled experiments. We did not find many articles comparing available risk analysis methods, which is one reason that we argue there is a need for this kind of research. We also believe that there is a need to further investigate the whole risk management process in longer case studies, where actual cases of risk management are investigated in practice.

Paper II: Risk Analysis and Management of IT Systems: Practice and Challenges

We used a semi-structured interview method [61] to investigate the current practices of risk analysis with experts having responsibilities of risk analysis of IT systems in large-scale public organizations. In total 5 persons responded that they were willing to participate in the study. Then, the interviews were conducted over a period of about four months and recordings were transcribed for further analysis. After this, a set of codes based on the research and interview questions was first produced and then iteratively updated during the analysis.

Risk analysis is important for safety-critical IT systems and services in both public and private organizations. A number of risk analysis frameworks have been defined and they are adapted by different organizations to different extent. However, the actual practices and the challenges of risk analysis in these contexts have not been fully explored. In this study, several factors were investigated regarding the current practices and existing challenges of risk analysis and management e.g., its importance, identification of critical resources, definitions of roles, involvement of different stakeholders, used methods, follow-up analysis. Furthermore, this study also investigates existing challenges in the current practices of risk analysis. A number of challenges are identified, e.g., that risk analysis requires competence both about the risk analysis procedures and the analyzed system, and to follow-up and repeat a risk-analysis that is conducted.

Paper III: Perspective Based Risk Analysis – A Controlled Experiment

In this paper, we present the results from a study designed to experimentally assess the potential of perspectives in risk management and therefore further experimentally explore the suggestions given in previous work [2, 6, 30, 46, 52, 70, 77]. In this paper we investigate the effectiveness of Perspective-Based Risk Analysis (PBRA) compared to Traditional Risk Analysis (TRA). Involving perspectives into risk analysis brings a potential to increase the efficiency of the risk analysis and confidence in the identified risks. A controlled experiment was designed and carried out. 43 subjects performed risk analysis of a software-controlled train door system using either PBRA or TRA. We measured the efficiency of the methods by counting the number of relevant and non-relevant risks and we used a questionnaire to measure the difficulty of the methods and the confidence of the subjects in the identified risks. In the experiment results some potential benefits of using perspective-based risk analysis are uncovered and confirmed. We found that PBRA helps to identify more relevant risks than TRA. In particular, it was discovered that PBRA is more effective than the traditional method and identifies more relevant risks.

Paper IV: Comparison of the FMEA and STPA safety analysis methods – a case study

In this paper, we compare two existing risk analysis methods, Failure Mode and Effect Analysis (FMEA) and System Theoretic Process Analysis (STPA). This study compares both methods by comparing the hazards type identified by these methods. Five hazard types were defined to analyze the identified hazards by both methods. These are component interaction, software, component failure, system type, and human error hazards. Then, we compared the causal factors for identified hazard by both methods. Furthermore, we also evaluated the analysis process of both methods by using a set of qualitative criteria, derived from the Technology Acceptance Model (TAM) [10, 11]. For this, steps of both methods are mapped to each other to find the common steps. Then, we compared the mapped common steps of both methods based on the qualitative criteria derived from TAM. Here, it should be noted that this study does not aim at comparing both methods quantitatively, but instead to understand the differences through a qualitative analysis. That is, we investigate both methods qualitatively by analyzing hazard analysis results and the process of analysis.

It can be observed that almost all types of hazards that were identified in the study were found by both methods. That is, both methods found hazards classified as component interaction, software, component failure and system type. The findings of this study regarding the identification of causal factors of the identified hazards reveal that STPA is better than FMEA. Here, a potential reason to result in more complete result is that STPA has a structured process to follow in doing the analysis and to identify causal factors. To summarize, FMEA takes the architecture and complexity of components into account, whereas STPA is stronger in finding causal factors of identified hazards. It can be concluded that none of the methods in this study was effective enough to find all identified hazards and their causal factors, which means that they complemented each other well in this study.

Paper V: Identification of IT Incidents for Improved Risk Analysis by Using Machine Learning

We present a prototype solution of a system that automatically identifies information pertaining to IT incidents, from texts available online on Internet news sources, that have already happened. This way IT incidents can be saved automatically in a database and the saved information can be used as an input to risk analysis. By having an overview of already occurred IT incidents, the risk analysis process can be improved, which is an essential activity for development and operation of safe software-intensive system. However, historical data about such unwanted events is not easily accessible and it is not available at a single place.

In this study for the proposed prototype solution, two datasets were manually classified. One dataset was used for training and the other dataset was used for evaluation. In this study 58% of the texts that potentially can contain information about IT incidents were correctly identified from an experiment dataset by using the presented method. It is concluded that the identifying texts about IT incidents with automated methods like the one presented in this study is possible, but it requires some effort to set up. This way, by having historical information of already happened IT incidents, risk analysis and management practices can be improved.

Paper VI: A Method for Assessing Resilience of Socio-Technical IT-systems

A simulation based method for resilience of IT networks is evaluated. Simulation-based methods that consider supply network topology as well as system responsible for repairing supply networks have previously been used and found to be beneficial for assessing resilience of electricity and water distribution systems [43]. Here, IT system and networks are studied as a socio-technical system, broadly understood as a system whose functionality is dependent on technical as well as organizational sub-systems. The aim of this study is to test if such a method is applicable for assessing resilience of IT-systems, meaning that: 1) it is possible to use based on available data, in this case gathered mainly through interviews with focus groups, 2) the results are relevant for users, owners and maintainers. This study was conducted in the following three main steps and they were carried out in a sequence. In step 1, data was collected from the case organization. In step 2, the model was used with the data that was collected in the previous step. In step 3, results from applying the model were presented to the organization and the researchers actively asked for information about the usefulness of the results.

The method was tested in a case study on the IT-network of one department of Lund university as well as the university core network. Results show that the method is applicable for the studied IT-network and that we can obtain the resilience metrics sought for. It is found that the method can enable system owners to see if and for what levels of strain they are presently reaching their desired targets concerning system resilience. Regarding the relevance of the method, feedback from system experts indicates that the method might primarily be useful for IT-systems whose failure would result in large economic losses (e.g. IT-systems of financial organizations) or lead to loss of health or safety (e.g. IT-systems of governmental organizations and hospitals).

6 Synthesis

This section summarizes the main results in relation to the research questions and included studies. Moreover, the main validity threats to the results are discussed for each study. More detailed descriptions of the results with their validity assessment for each study can be found in the respective papers. Figure 2 shows the relationship between the research objective, research questions and the included papers.

RQ1: What is the current state of the risk analysis research and practice?

Regarding RQ1, we investigated the current state of the risk analysis research that is explained below in the sub research questions.

RQ1.1: We investigated what risk analysis methods and approaches exist for analyzing IT systems. Moreover, we also investigated if there is any empirical research that compares or evaluates existing risk analysis methods. RQ1.1 has two parts and both parts are addressed by Paper I. In Part I, existing risk analysis methods or techniques for analyzing IT systems were identified and investigated. 57 studies were identified in the mapping study that present different, existing or new, risk analysis methods. A majority of the identified studies focus on presenting new risk analysis methods. The main focus of this part is on types of IT systems for which risk analysis methods are presented and also types of risk analysis methods (quantitative or qualitative).

In the second part of RQ1, the focus was on research that compares different risk analysis methods empirically (controlled experiments or case studies) and concludes which methods are more effective. We found that the majority of the identified studies present non-empirical research. This study identified 36 studies presenting analytical (non-empirical) research and 21 studies presenting empirical research (case studies). None of the identified studies present research conducted as surveys or controlled experiments for comparison and evaluation of different methods. This mapping study identified five studies that describe, analyze and compare existing well-known risk analysis methods but they do not present empirical research. Based on this we conclude that there is a need for empirical investigations of risk analysis methods for analyzing IT systems by conducting controlled experiments and case studies.

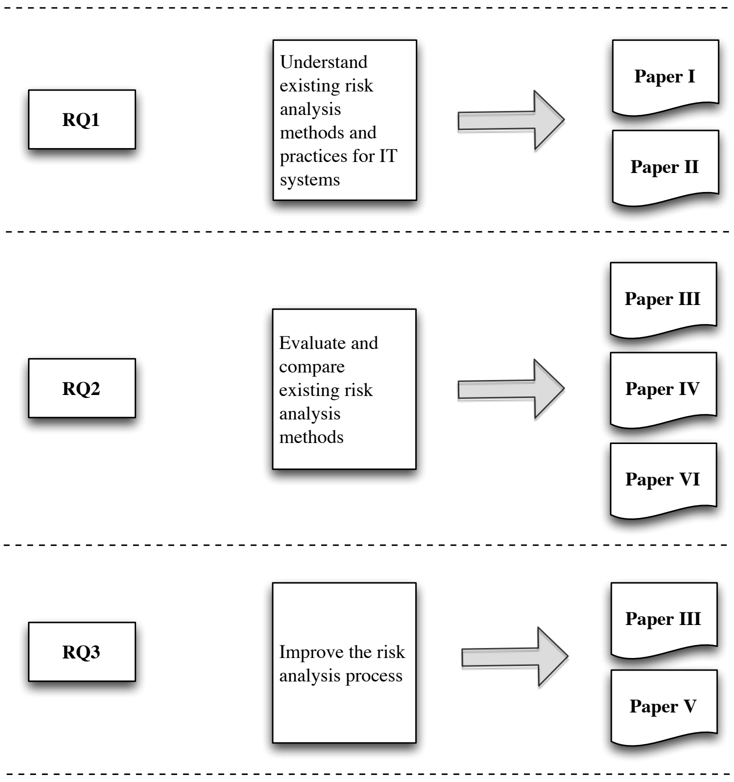

RQ1
Understand existing risk analysis methods and practices for IT systems
Paper I
Paper II
RQ2
Evaluate and compare existing risk analysis methods
Paper III
Paper IV
Paper VI
RQ3
Improve the risk analysis process
Paper III
Paper V

Figure 2: The relationship between the research questions, research objective and the included papers in this thesis

Regarding the types of risk analysis methods, it was found that qualitative risk analysis methods to a larger extent were investigated in empirical research than quantitative methods. Based on this, it could be argued that this is due to lack of easiness in application of quantitative risk analysis methods in practice that require exact statistical information to estimate likelihood and consequences of identified risks. This study has also identified two studies that present semi-quantitative risk analysis methods, which do not require exact statistical information needed for quantitative risk analysis and offer better estimates than qualitative risk analysis methods. Based on this, it can be concluded that there is a need for more research on risk analysis methods or techniques that combine and utilize the benefits of both quantitative and qualitative methods.

RQ1.2: We investigated current practices of risk analysis and management for IT systems in large-scale public organizations and also the main challenges in carrying out risk analysis. It is addressed by Paper II, which investigates the current practices of risk analysis by an interview-based investigation with experts having responsibilities of risk analysis of IT systems in large-scale public organizations.

From this study we conclude regarding the importance of risk analysis in large-scale organizations that it is an important activity for organizations that are dealing with safety-critical systems and services corroborating with the findings of [54]. It is an important activity because most of an organization's security and safety countermeasures or mitigations in different sub-departments and projects are based on risk analysis and assessment. However, the associated challenge with risk analysis is that different employees have different opinions about risks and the risk analysis process. This difference in opinions is probably because of different level of perception, trust, and priorities about the risk analysis process. Therefore, this difference in opinions about importance of risk analysis makes difficult to carry out risk analysis.

Regarding identification of critical assets and services the study suggests that it is an important activity. However, all organizations are not carrying out this activity because they are not fully sophisticated in risk analysis and management as discussed by Henschel [25]. Here, it means some organizations are systematic and follow well-known risk analysis methods and standards but some are in improvement phase and trying to improve their processes and procedures. This kind of activities requires that they should be properly documented to be used later for example in a crisis situation or in follow-up analysis.

Then, regarding carrying out risk analysis practices almost all investigated organizations are analyzing their critical systems or services in the same way from an abstract level. The reason for this could be that these all organizations are of the same nature i.e. governmental organizations dealing with critical services to society. One interesting fact that was found is that these organizations mainly focus on the information or services in risk analysis instead on an IT system. It is information or service that makes an IT system important not the other way around. Here, it can be said that they are using the system level analysis method [53]. We also identified a set of challenges with carrying out risk analysis practice. Firstly, the required competences and skills are a great challenge in carrying out risk analysis in these critical organizations. The findings of this study suggest that lack of knowledge and expertise about doing risk analysis is itself a risk. Moreover, the knowledge about the system context that is being analyzed and its boundary definitions are very crucial as discussed by Lindholm et al. [48]. Secondly, pre-understanding of the risk analysis process is also a challenge while performing risk analysis of safety-critical services. This challenge is very similar to the required skills and competences challenge. However, it is about having good pre-understanding of potential risks, the risk analysis process, and the system being analyzed with its context. On the other hand, required skills and expertise deal with the knowledge of different risk analysis methods or tools and then the knowledge used for defining the system boundaries. The best practices of risk analysis and management identified by Murdock et al. [54] also suggest that the risk management process should start with context establishment that includes organizational objectives, stakeholders, constraints, risk criteria, and other factors.

After this, involvement of different people in risk analysis is investigated and it is found that these organizations involve more than one person with different knowledge in the risk analysis process. Several authors for example, Leveson [46], McDermott et al. [52], and Sulaman et al. [III] advocate to involve various roles or perspectives in risk analysis. Regarding involvement of different people in risk analysis the investigated organizations seem to be mature. The associated challenge (RQ2) with this practice is subjectivity involved in risk analysis. The subjectivity in risk analysis has both positive and negative effects and there is a debate going on whether risk is subjective, objective or some combination of both [8, 27]. However, the findings of Paper II reveal that by involving different roles and knowledge in the risk analysis process eliminates at least the negative effects of subjectivity in risk analysis. Therefore, by involving different competences and roles in the risk analysis process brings significant strength as confirmed in an experimental study [III].

Regarding practices pertaining to the used methods in these large-scale organizations, we found that all the investigated organizations are using well-known analysis methods and standards i.e. ISO 27000, 31000. However, mostly they have adapted these national and international methods and standards according to their specific needs. Furthermore, the analysis carried out in Paper II reveals that the simpler (easy to adapt and use) methods, models and tools for risk analysis are better and being used in the investigated organizations as this corroborates with the findings presented by Murdock et al. [54].

Regarding follow-up analysis, we found that it is not common in all organizations because of the difference in opinions of people involved with the risk analysis process. The main challenge pertaining to follow-up analysis found in this study is how people perceive risk analysis and its priority among other things that is not same in all organizations. If an organization is dealing with very critical services (health care) then they have documents and procedures to carry out follow-up analysis. After every change they carry out follow-up risk analysis for their critical services and systems. However, other organizations investigated in Paper II also have documents and procedures defined to carry out follow-up analysis but they are not carrying out it on regular basis.

RQ2: How can we evaluate the effectiveness and efficiency of a risk analysis method?

RQ2 evaluates the effectiveness and efficiency of the risk analysis methods that is explained in the following sub research questions.

RQ2.1: Here, we investigated how we can evaluate and compare different risk analysis methods. It is addressed by Paper III and IV. It is not easy to compare or evaluate risk analysis methods because of their subjective nature. The risk analysis process is mainly a brainstorming activity that can be performed in different ways by following different methods or frameworks. The main challenge is to find attributes that can be used for evaluation of different risk analysis methods.

Based on the results of Paper III, where we compared and evaluated two risk analysis methods in a controlled experiment, we conclude that the risk analysis methods can be quantitatively evaluated and compared by counting the *number of relevant and non-relevant risks* identified by the participants or risk analysts. Experiments are more suitable for evaluation of different risk analysis methods but the participants should have at least moderate experience of working in industry. Moreover, the experiment participants should have similar level of expertise and experience. This way we can evaluate and compare different risk analysis methods to conclude which method is effective among others. After this, the *ease of use* is another suitable attribute to evaluate effectiveness and efficiency of risk analysis methods. A questionnaire or an interview is the data collection instrument for this attribute. In paper III we also have investigated the *confidence of participants* on their identified risks. The *time efficiency*, investigated in [VII], is also a suitable attribute for evaluation and comparison of different risk analysis methods. However, the measurement of required effort should be done carefully. In [VII], we were not able to measure the effort required of STPA application accurately because the hazard analysis was carried out with interruptions (doing other work).

Based on the results of Paper IV, where we compared two hazard analysis methods, we conclude that the risk analysis methods can also be evaluated and compared qualitatively. For this qualitative evaluation, all the identified hazards were classified into the following five error categories: *component interaction error, software error, human errors, component error, and system error*. These categories were selected and adapted from the previous studies [18, 19, 47]. The main findings regarding the classification of the identified hazards are that FMEA did not find any unique hazard of component interaction and human error type that is not identified by the STPA method. Here, one interesting result is that FMEA identified as many software error type hazards as STPA. It should be noted that the data points in this study are few and the focus of the study is not on quantitative comparison of the methods. However, as noted above, there is almost no difference regarding the identified software error type hazards by both methods. One positive result in favor of STPA, based on the experience of the authors of this study, is that it identified clear software error type hazards because of its keywords ('provided', 'not provided', etc.), which made it simple and easy to identify software error type hazards. There are no common identified hazards of human error type. Apparently, none of the methods could find any human error type hazard in this study. The reason for this can be that the analyzed system does not involve much human input or interaction.

After this, in Paper IV, the analysis processes of both methods, FMEA and STPA, were evaluated and compared based on the following criteria derived and inspired by Technology Acceptance Model (TAM) [10, 11]. *How easy or hard*, *Why was it easy or hard*, *Support by method*, *Confidence about the results*, *Applicability for software*.

Based on the evaluation performed in Paper IV, we conclude that the evaluation and comparison of different risk analysis methods are possible by using the aforementioned criteria. The main findings regarding the evaluation and comparison are that there were no type of hazard that was not found by any of the methods, which means that it is not possible to point out any significant difference in the types of hazards found. However, it can be observed that none of the methods in Paper IV was effective enough to find all identified hazards, which means that they complemented each other well in this study.

RQ2.2: Regarding RQ2.2, we investigated how to assess the resilience of critical IT systems and networks that can help to determine how dependable a typical system or network is. It is answered by Paper VI that assesses resilience of critical IT systems and networks by applying a simulation method. A hybrid modeling approach is used which considers the technical network, represented using graph theory, as well as the repair system, represented by a queuing model. Simulation based methods that consider repair system as well as technical network have previously been used for assessing resilience of electricity and water distribution systems. The main objective of Paper VI was to test if such a method 1) is applicable within the IT-context, 2) is giving relevant results and 3) captures all relevant factors. Below these three questions are discussed.

Regarding *applicability* of the simulation model on IT systems, the preliminary findings of this study show that the method is applicable for the studied IT-network and that we can obtain the resilience metrics sought for. The findings show that disturbance scenarios for which resilience is low can be identified, based on three important resilience indicators: robustness, rapidity and resilience loss. The results of this type can be useful in the process of increasing system resilience. This study further show how system robustness and rapidity change with level of strain. This makes it possible for system owners/operators to see if and for what levels of strain they are presently reaching their desired targets concerning system resilience.

Regarding *relevance* of the results of model, the findings of this study show that the applied model was not found to be relevant to the personnel of the studied system, since this system is not considered to be critical enough. The system experts however thought that the method could be useful when applied to IT systems whose outage could cause either large economic losses or risk to health and safety.

Regarding *completeness* of the applied model, software faults are not considered in the model. It was assumed that the software faults constitute about 20% of the total number of faults in network nodes. These faults are not considered mainly due to their irregular repair time and consequences. Considering this type of faults is a possible topic for future work. Based on the findings of this study it can be concluded that we can assess resilience of a critical IT system by using the proposed simulation model.

RQ3: How can we improve the risk analysis process?

Regarding RQ3, we suggested different improvements for the risk analysis process that are explained below in more detail.

RQ3.1: To answer RQ3.1 we investigated how the use of different perspectives can improve the risk analysis process. It is addressed by Paper III. We propose a new risk analysis method, using different perspectives and roles, and evaluate it in a controlled experiment. In the experiment two methods were compared, Perspective Based Risk Analysis (PBRA) and Traditional Risk Analysis (TRA), to evaluate the use of different perspectives.

The experiment confirms that subjects using PBRA found more relevant risks. We found a statistically significant result that more relevant risks were found by using perspectives than by not using perspectives. We also found a statistically significant result that by using perspectives, the risk analysis becomes more difficult than by not using. We believe that by having different perspectives, the risk analysis becomes more thorough resulting in an in-depth analysis, which makes it more difficult. Moreover, we did not find any statistical difference in the confidence level of the participants with or without using different perspectives. The participants using both treatments, with or without perspectives, were not confident about their identified risks. We believe that the main reason for this lack of confidence was the lack of experience and domain knowledge of the participants. Based on these findings, we conclude that the use of different perspectives greatly improves effectiveness of risk analysis process.

RQ3.2: Here, we investigated how we can identify and save historical information about IT incidents to improve the risk analysis process. It is addressed by Paper V that proposes and evaluates an approach for automatically collecting information about IT incidents from online news sources. To improve risk analysis and management practices, the historical information about already happened incidents is important for the correct estimation of the likelihood of potential risks and their consequences. Based on the findings of Paper V, we conclude that it is possible to identify interesting texts from a large number of potential texts but it requires a substantial effort to set up. We found that it is possible to support the work of identifying texts about IT incidents with automated methods like one presented in Paper V. This means it could be an important aid in the process of building a database of occurred IT incidents that later can be used as an input to improve the risk analysis process.

RQ3.3: Regarding this research question, we investigated how can the risk analysis practices be improved in large-scale organizations. It is addressed by the research work carried out in this thesis. In Paper II, with the investigation of current practices of risk analysis the existing challenges were also investigated. Moreover, All the participants, in Paper II, were also asked to suggest improvements for the risk analysis process and they suggested a few improvements. Furthermore, we present an improvement in Paper III i.e. the idea of using different perspectives in the risk analysis process that can be used for eliminating the negative effects of subjectivity in risk analysis. Then, in Paper V we present an approach to identify historical information about IT incidents. Here, the idea is to use historical information about IT incidents or previously carried out risk analyses to correctly estimate the future risks i.e. their probabilities and consequences.

Based on the findings of the studies included in this thesis and suggestions provided in the investigation carried out for Paper II by the risk practitioners, we created a list of potential suggestions to improve the risk analysis and management process. Then, we evaluated this list in a focus group meeting with the two senior and one mid-level researchers from the risk analysis domain. In the focus group, all the potential suggestions were presented one by one and then the participants were asked to verbally discuss and give their feedback, both positive and negative, on a paper form about each suggestion. After this, the data collected from their feedback was analyzed and synthesized. The following potential suggestions were identified in the research carried out in this thesis (mainly Paper II, III, and V) and then evaluated in the focus group meeting, presented in Table 2.

1. **Risk analysis awareness and education:** The majority of the interviewees suggested that the risk analysis process can be improved by communicating

Table 2: Suggested improvements in the risk analysis process and practices

No.	Suggested improvements
1	Risk analysis awareness and education
2	Defining clear roles and responsibilities
3	Easy-to-use and adapt risk analysis methods
4	Dealing with subjectivity in risk analysis
5	Carry out risk analysis as early as possible
6	Using historical risk data in risk analysis

and making people understand the importance and benefits of carrying out risk analysis instead of seeing it as a hinderance. It suggests organizations to have some training, seminar or sessions for risk education and awareness for their employees to improve organization-wide security and safety thinking. Here, the important fact is to communicate risks and the analysis process in a batter way that should be transparent i.e. showing both good and bad. Moreover, the participants in the study reported in Paper II also suggested that one should develop an environment in the organization where everyone is aware of potential security and safety threats to the organization and its assets. By having more knowledge about these potential threats one can estimate the real benefits of carrying out risk analysis and management. It is also known as a risk aware environment in an organization. One participant mentioned that:

> *To achieve success in the future we will change the way of working. We (central risk analysis unit) will contact dependent institutions, departments and by telling them that we will implement the security and safety thinking in your work by educating everyone.*

One participant in Paper II mentioned that spreading information to the employees and stakeholders about the existing challenges could also improve the risk analysis process. Moreover, it is also important to add risk analysis and management in the organization policies that everyone responsible for an area or project must carry out analysis before initiating it or after introducing an important change.

During the evaluation in the focus group meeting the participants endorsed this suggestion by saying that it is important and all the organizations dealing with safety critical services or systems should have that already. Then one senior researcher (focus group participant) mentioned based on his experience that most organizations struggle with emphasizing the importance of risk analysis and management. Moreover, one of the focus group participants mentioned that it is important to highlight the risk analysis process and its results that how they can be used to improve organization-wide security and safety awareness and also to get money for mitigating actions. Furthermore, it was discussed and suggested that it is a good idea to have a lexicon of risk terminologies with some introductory courses to improve risk analysis awareness.

With the advantages of this suggestion the disadvantages were also discussed in the focus group meeting. The downside of this suggestion is that it requires economical resources and time to implement. There is a need to find a balance between the benefits of risk education and awareness and the cost associated with it for an organization. Moreover, it has a costs and then becomes difficult to prove that it has value for its cost. One more important thing was discussed in focus group that who should be the target of this risk education and awareness in an organization. Because to educate whole organization may be is not possible for an organization or it becomes very costly. Moreover, sometimes it can become difficult to reach a potential person for risk education with very little interest in risk analysis in an organization.

2. **Defining clear roles and responsibilities:** The findings of Paper II suggest that it is important to define clear roles and responsibilities for both normal and crisis situations. According to them, in large-scale organizations that manage several critical services to society it is crucial to prioritize these services and then each service should have a responsible person. The responsible person or unit for each individual service or system can help greatly in analyzing potential risks. Then, the responsible person or unit can also help in aggregating risk information at the central risk analysis unit to take decisions for countermeasure of identified potential risks.

   During the evaluation in focus group meeting the participants discussed that it is good if people are well prepared beforehand and that they know what to do in different situations. It is an important step in defining and planning a risk analysis process. It is good to try to pinpoint responsibilities and to know who is the risk owner that can help in carrying out and guiding the risk analysis process. Finally, it was suggested that in every organization delivering safety critical services there should be a person responsible for each service. The drawback of this suggestion is that again it is expensive in terms of time and money. Moreover, it is also difficult to prioritize it over other things as main responsibles are busy with doing other tasks.

3. **Easy-to-use and adapt risk analysis methods:** The findings also suggest to use as simple risk analysis model or methods as possible. According to them, the term “simple” means methods or models for risk analysis that are easy to use because in practice people working with risk analysis are not technicians or mathematicians. Therefore, as mentioned by one of the participants,

> *We are trying to set an easy work model for risk analysis and management. One important thing is to consider, carry out risk analysis in the easiest way because it is a complex activity.*

That is, the study participants asked for simple, easy to use and adapt, risk analysis models or methods that are easy to use, effective and efficient, i.e. yield complete analysis results.

In the evaluation, it was found that easy to use and adapt (simple) methods are good if they support to increase carrying out risk analyses and the results are more meaningful and useful. In general, easy methods are quite efficient but they are usually not very accurate. On the other hand, hard or difficult methods are very accurate but they are less efficient as they take more time to follow or carry out. Overall, it is important to find a balance between easy methods and the completeness of analysis results yielded from these methods. Furthermore, there should be some education and normative documentation (guidelines, instructions, examples, etc.) regarding these methods that can support to make the methods easy to use. Moreover, it is also important to have some tool support for risk analysis. Finally, it should also be suggested that the quantitative values should be used for consequences and likelihood with uncertainty intervals.

However, the only drawback associated with this is that a risk analysis method should not be too simple because then it can be very efficient but might yield irrelevant risk results.

4. **Dealing with subjectivity in risk analysis:** Based on the findings of Paper III and the improvement suggestions provided by risk practitioners in Paper II we found that one should carry out risk analysis in a group and try to have consensus regarding potential risks and their consequences. This can only be done by including different competencies and perspectives while carrying out risk analysis for IT systems. It also can help to cope with the negative effects of subjectivity in risk analysis. One of the practitioners mentioned that:

   > *people can have different opinions about the probability and consequences levels of a potential risk. Therefore, it is important in this case to discuss it further in a group to decide the values in a certain way.*

   According to the findings of Papers III and II, the negative aspects of subjectivity in risk analysis can be eliminated to some extent by involving different roles and competencies in the risk analysis process.

   In the evaluation carried out in the focus group meeting we found that it is good to involve several perspectives in the risk analysis process. Moreover, carrying out risk analysis in groups is good and also important to communicate risks between people in the organization. It is even more effective if we ask people to have a specific role while carrying out risk analysis and to prepare for it before the analysis meeting.

5. **Carry out risk analysis as early as possible:** The carried out investigations for improving the risk analysis process in Paper II identified an important fact that one should carry out risk analysis as early in a project as possible. Then later a follow-up analysis can be carried out, which can greatly improve the project results. One of the interviewees mentioned that:

   > *We at the central risk unit, in a public large scale organization, are trying to drive other sub-departments or organizations to carry out risk analysis as early as possible and then later carrying out a follow-up analysis.*

   According to the discussions of the focus group meeting it was found that it is important to do the analysis early. Moreover, the most important thing is that risk analysis should be carried out at a planned time instead of postponing it too long. There is a difference between project and operational risk analyses. In projects, risk analysis is usually carried out at a planning stage and on the other hand operational risk analysis is carried out after a fixed interval, which is also called follow-up analysis. Overall, according to the discussions of the focus group meeting this suggestion to improve the risk analysis process is very practical and risk practitioners really consider it while planning and carrying out risk analysis. The only associated difficulty with carrying out risk analysis as early as possible is that it can be a bit difficult to identify risks at an early stage.

6. **Using historical risk data in risk analysis:** Based on the basic idea behind the investigation of Paper V and findings of the focus group meeting it is found that the historical data is very important for future risk analyses. The historical risk data or information can be used as an input to risk analysis, which can help to correctly estimate the likelihood and consequences of potential risks. Almost all existing risk analysis methods require detailed system information with an overview of already happened unwanted events. To improve the risk analysis process it is believed that this can be done by having an overview of already encountered risks.

   In the focus group it was discussed that historical risk information is very important for carrying out future risk analyses. Furthermore, it also helps to eliminate the negative aspects of subjectivity in risk analysis. An overview of already encountered risks can also help in increasing risk education and awareness in an organization. Therefore, in the focus group meeting it was recommended that risk practitioners must, if such information is available, try to use historical information about risks and system that is being analyzed to correctly estimate risk values.

7 Validity Assessment

The validity of a study represents the trustworthiness of its results, which means for example that the results are not biased by the researcher's own opinion or point of view [63]. The validity of the studies included in this thesis can be assessed regarding construct validity, internal validity, external validity, and reliability [63, 76].

Construct validity considers the studied artifacts and concerns if they represent what the researcher have in mind and also if the studied artifacts are investigated according to the research questions of the study. *Internal validity* is important and mostly applicable in studies of causal relationships. When the researcher investigates that one factor is affected by investigated factor there is a risk that the investigated factor is also affected by a third factor. *External validity* is concerned with to what extent it is possible to generalize the findings, and to what extent the findings are of interest to people outside the investigated case. *Reliability* is concerned with to what extent the data and the analysis are dependent on the specific researchers. In general to improve reliability both the data collection and analysis should be done by a group of researchers instead of one single researcher. Below is a brief description of the countermeasures that were taken during the execution of research work presented in the included papers. For more details of validity assessment see validity sections in the included papers.

The main validity issue for the research work carried out in Paper I concerns missing possible relevant studies due to practical issues. First, there might exist few lesser known journals and conferences that might not be available in the searched databases. Secondly, the full text of a few identified studies were not available, mostly of old studies. Thirdly, it is likely that some possible relevant studies were not identified by the used search query because it is not possible to have a search query that identifies all relevant studies. Finally, there was a chance of incorrectly rejecting possible relevant studies by the authors during the selection process. In order to reduce the afore-mentioned validity threats the following measures were taken. First, different synonyms for IT systems were used in the search query to reduce the chance of missing possible relevant studies. Then, the reference lists of the most relevant identified studies were also examined for missing possible relevant studies. Finally, to reduce the threat of incorrect rejection of relevant study during the selection process, the co-authors cross-checked all the selection steps carried out for the selection of relevant studies.

The main validity issue for Paper II has probably to do with the sampling of interviewed subjects. The interviewed subjects can be seen as a good representation of risk professionals because that are collaborating and supporting many risk professionals working in the same organizations. Moreover, all investigated organizations use international risk analysis and management methods and standards. That is, the findings seem general enough that they should also hold for other large-scale organizations.

The main validity issue for Paper III is the threat to external validity. There could be a chance of this threat because the sample for the experiment consists of students of a project course and are therefore not representative for the entire population. To reduce the effect of this threat, a pilot study was carried out by using experts from industry and academia. There was not a big difference in the number of identified relevant risks found by the experts and students. To reduce the chance of random heterogeneity of subjects, which can affect the results, the participants for both treatments were selected from the same level of education with almost similar knowledge and background.

In Paper IV, the second author analyzed the system using FMEA and all the documentation and information were available, which were used during the hazard analysis performed by the first author. There could be a risk of not understanding the analyzed system and its description by the authors. To decrease this threat a simple system was selected and also its detailed description was acquired and made available to all the authors of this study. Furthermore, to decrease the risk of history threat the following measure were taken. The second author of this study was selected to apply FMEA on the collision avoidance system. The first author already knows the existing hazards in the selected system because he has applied STPA on the selected system in the previous study [VII]. Therefore, to improve research validity it was decided that the first author would not apply the FMEA method on the selected system. Instead another author did that. The second author of this study did not have access or review the previous study results [VII].

The main validity issue for Paper V is the scalability of the proposed solution. In the presented work we proposed a prototype solution using an example dataset. There is a need for further research to implement this system that can be executed in runtime while reading text from online sources. Another validity issue is that it is not clear how to select a sufficient and representative set of information sources to be used by the system. Solutions to these issues require more investigation and for that further research is needed.

The main validity issue for Paper VI is that whether the yielded results are relevant and meaningful to the practitioners or not. To address this threat, results from applying the model were presented to the investigated organization. This was carried out in an informal setting where representatives from the organization participated and were able to give feedback on the usefulness of the approach. The researchers actively asked for information about the usefulness of the results at the meeting.

One potential validity threat concerns carrying out research in the lab environment instead of live industry environment and using embedded systems for analysis instead of IT systems. In this thesis the methods are evaluated and compared by using embedded systems, e.g., train door system and forward collision system instead of IT systems. There are a few reasons behind the selection of lab environment and these embedded systems. The first reason in selection of these systems was that we wanted to have as simple as possible systems for the analysis. The other reason was that it is cost effective as it was easy to invite students to participate in the research than practitioners from industry. As suggested by Gorschek et al. [21], carrying out research in a lab environment can provide fast, valuable feedback, identifying obvious flaws so that researcher can fix them before industry piloting. Moreover, it was a challenge to find a real IT system in live industry environment to evaluate and compare different risk analysis methods for the research presented in this thesis.

8 Conclusions and Future Work

The main objective of the research work presented in this thesis was to improve the analysis process of risks pertaining to IT systems in large-scale organizations, which is addressed in three different ways as mentioned in Section 3. That is, firstly by understanding current literature and practices related to risk analysis of IT systems. Secondly, by evaluating and comparing the existing risk analysis methods. Finally, by suggesting improvements in the risk analysis process and by developing new effective and efficient risk analysis methods to analyze IT systems.

A systematic mapping study is presented in Paper I to answer RQ1 partly, i.e., to understand existing methods and approaches used for analyzing IT systems. 63 primary studies were identified in the mapping study that present different risk analysis methods. A majority of the identified studies focus on presenting new risk analysis methods and non-empirical research. Only five studies were identified that describe, analyze and compare existing, well-known, risk analysis methods. Based on this we conclude that there is a need for empirical investigation of risk analysis methods for analyzing IT systems by conducting case studies and controlled experiments. A semi-structured interview study is presented in Paper II to answer RQ1. In this study, several factors were investigated regarding the current practices of risk analysis and management, e.g., its importance, identification of critical resources, definitions of roles, involvement of different stakeholders, used methods, and follow-up analysis. Furthermore, this study also investigates existing challenges in the current practices of risk analysis. A number of important challenges are unfolded by this study, e.g., that risk analysis requires competence both about the risk analysis procedures and the analyzed system, which is challenging to identify, and that it is challenging to follow-up and repeat a risk-analysis that is conducted, also it is important to eliminate the negative effects of subjectivity in the risk analysis process.

A controlled experiment is presented in Paper III and a case study is presented in Paper IV to answer RQ2, i.e., to evaluate the effectiveness and efficiency of risk analysis methods. Based on the results of paper III and IV, we conclude that the effectiveness and efficiency of risk analysis methods can be evaluated and compared by counting the number of relevant and non-relevant risks identified by the participants or risk analysts. Moreover, the ease of use is another suitable attribute to evaluate effectiveness and efficiency of risk analysis methods. The time efficiency is also a suitable attribute for evaluation and comparison of different risk analysis methods. Furthermore, the evaluation presented in Paper IV investigates the effectiveness of the methods by performing a comparison of how a hazard analysis is conducted for the same system. It is also possible to evaluate the analysis process of risk analysis methods by using the presented set of qualitative criteria, derived from the Technology Acceptance Model [10, 11]. This way we can evaluate and compare different risk analysis methods to conclude which method is effective among others. Another case study is presented in Paper VI to answer RQ2 that assesses resilience of critical IT systems and networks by applying a simulation method. A hybrid modeling approach is used which considers the technical network, represented using graph theory, as well as the repair system, represented by a queuing model. The findings of this study show that the method is applicable for the studied IT network and that we can obtain the resilience metrics sought for. Results show that disturbance scenarios for which resilience is low can be identified. Results of this type can be useful in the process of increasing system resilience. The findings further show how system robustness and rapidity change with level of strain. This makes it possible for system owners/operators to see if and for what levels of strain they are presently reaching their desired targets concerning system resilience.

A controlled experiment is presented in Paper III to answer RQ3 partly, i.e., to improve the risk analysis process. For this, a new risk analysis method is presented, Perspective Based Risk Analysis (PBRA), that suggests the use of different perspectives. In this study the use of different perspectives in risk analysis is suggested and empirically assessed. We found a statistically significant result that more relevant risks were found by using perspectives than by not using perspectives. Based on these findings, we can conclude that the use of different perspectives improves effectiveness of the risk analysis process and also eliminates the negative aspects of subjectivity. Another case study is presented in Paper V to answer RQ3. Based on the findings of Paper V, it can be concluded that it is possible to identify interesting texts from a large number of potential texts but it requires a substantial effort to set up. We found that it is possible to support the work of identifying texts about IT incidents with automated methods like one presented in Paper V. This means it could be an important aid in the process of building a database of occurred IT incidents that later can be used as an input to improve the risk analysis process. Furthermore, based on the findings of the studies included in this thesis and the the investigation carried out in Paper II a list of potential suggestions was created and evaluated in a focus group meeting. For example, risk analysis awareness and education, defining clear roles and responsibilities, easy-to-use and adapt risk analysis methods, dealing with subjectivity, carry out risk analysis as early as possible and finally using historical risk data to improve the risk analysis process. Based on the findings of the focus group meeting it can be concluded that these suggestions are important and useful for risk practitioners to improve the risk analysis process.

Conducting research answers some questions and raises many more. Based on the results of Paper I we found that there is a need for empirical investigation of different risk analysis methods. Therefore, the ambition is to further investigate the different risk analysis methods for their adaptation to IT systems or to develop new risk analysis methods and techniques specific for IT systems. A first important continuation of the work in future is to replicate the study presented in Paper III with practitioners. We plan to apply Perspective Based Risk Analysis (PBRA) method on more complex systems by involving practitioners having extensive experience. Another continuation of the work in future is to develop a risk management framework for the large-scale organizations based on the findings of Paper II. These findings will help in the development of the risk management framework. Then, there is a need to evaluate the planed risk management framework by conducting a case study in a large-scale organization by analyzing and managing IT risks.

BIBLIOGRAPHY

[1] A. Abdulkhaleq and S. Wagner, “A controlled experiment for the empirical evaluation of safety analysis techniques for safety-critical software,” in *Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering, EASE’15*. ACM, 2015, pp. 16:1–16:10.

[2] C. Alberts, A. Dorofee, J. Stevens, and C. Woody, “Introduction to the OCTAVE approach,” Networked Systems Survivability Program, Carnegie Mellon Software Engineering Institute, Pittsburgh, PA 15213-3890, Tech. Rep., 2003.

[3] C. Alberts and A. Dorofee, *Managing Information Security Risks: The Octave Approach*, ser. SEI Series in Software Engineering. Addison-Wesley, 2003.

[4] V. R. Basili, S. Green, O. Laitenberger, F. Shull, S. Sørumgård, and M. V. Zelkowitz, “The empirical investigation of perspective-based reading,” *Empirical Software Engineering*, vol. 1, no. 2, pp. 133–164, 1996.

[5] H.-P. Berg, “Risk management: procedures, methods and experiences,” *Reliability: Theory & Applications*, vol. 1, no. 2, pp. 79–95, 2010.

[6] N. Boudriga, M. Hamdi, and J. Krichene, “Netram: A framework for information security risk management,” Techno-parc El Ghazala, Route de Raoued, Ariana, 2083, Tunisia, Tech. Rep., 2007.

[7] S. Brinkmann and S. Kvale, *InterViews: Learning the Craft of Qualitative Research Interviewing*. SAGE Publications, 2014.

[8] S. Campbell, “Risk and the Subjectivity of Preference,” *International Journal of Risk Research*, vol. 9, no. 3, pp. 225–242, 2006.

[9] Central Computer and Telecommunications Agency, Great Britain, Treasury, *Prince User’s Guide to CRAMM*, ser. Programme and Project Management Library, 1993.

[10] F. D. Davis, “Perceived usefulness, perceived ease of use, and user acceptance of information technology,” *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989.

[11] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, “User acceptance of computer technology: A comparison of two theoretical models,” *Management Science*, vol. 35, no. 8, pp. 982–1003, 1989.

[12] S. Easterbrook, “Empirical research methods for software engineering,” in *proceedings of the 22:nd International Conference on Automated Software Engineering*, ser. ASE ’07. ACM, 2007, pp. 574–574.

[13] S. Easterbrook, J. Singer, M.-A. Storey, and D. Damian, “Selecting empirical methods for software engineering research,” in *Guide to Advanced Empirical Software Engineering*, F. Shull, J. Singer, and D. I. Sjøberg, Eds. Springer, 2008, pp. 285–311.

[14] ENISA Ad Hoc Working Group on Risk Assessment and Risk Management, “Inventory of risk assessment and risk management methods,” European Union Agency for Network and Information Security, 2006.

[15] C. A. Ericson, “Fault Tree Analysis - A History,” in *proceedings of The 17:th International System Safety Conference*, 1999.

[16] C. F. Eubanks, S. Kmenta, and K. Ishii, “Advanced failure modes and effects analysis using behavior modeling,” in *Proceedings of Design Engineering Technical Conferences and Design Theory and Methodology Conference, Sacramento, California, USA*, 1997.

[17] F. Ewald, *The Foucault Effect: Studies in Governmentality*. Harvester Wheatsheaf, 1991, ch. Insurance and Risk, pp. 197–211.

[18] C. H. Fleming, M. Spencer, N. G. Leveson, and C. Wilkinson, “Safety assurance in NextGen,” NASA/CR-2012-217553, Tech. Rep., 2012.

[19] C. H. Fleming, M. Spencer, J. Thomas, N. Leveson, and C. Wilkinson, “Safety assurance in NextGen and complex transportation systems,” *Safety Science*, vol. 55, pp. 173–187, 2013.

[20] M. Gleirscher, “Hazard Analysis for Technical Systems,” in *Software Quality, Winkler D., Biffl S., Bergsmann J. (eds). Increasing Value in Software and Systems Development (SWQD). Lecture Notes in Business Information Processing, Vol 133*. Springer, 2013, pp. 104–124.

[21] T. Gorschek, P. Garre, S. Larsson, and C. Wohlin, “A model for technology transfer in practice,” *IEEE Software*, vol. 23, no. 6, pp. 88–95, Nov. 2006.

[22] L. Grunske, R. Colvin, and K. Winter, “Probabilistic model-checking support for FMEA,” in *Proceedings of the 4:th International Conference on the Quantitative Evaluation of Systems, QEST 2007*, Sept 2007, pp. 119–128.

[23] G. Gurvitch, *The Spectrum of Social Time*. D. Reidel Publishing Company, 1964.

[24] C. M. Haissig and R. Brandao, “Using TCAS Surveillance to Enable Legacy ADS-B Transponder Use for In-trail Procedures,” in *Proceedings of the 31st Digital Avionics Systems Conference (DASC), Williamsburg, Virginia, USA*, 2012, pp. 5D5:1–5D5:114.

[25] T. Henschel, “Typology of risk management practices: an empirical investigation into German SMEs,” *International Journal of Entrepreneurship and Small Business*, vol. 9, no. 3, pp. 264–294, 2010.

[26] M. Höst, A. Oručević-Alagić, and P. Runeson, “Usage of open source in commercial software product development – findings from a focus group meeting,” in *Proceedings 12th International Conference on Product-Focused Software Process Improvement (PROFES)*, D. Caivano, M. Oivo, M. T. Baldassarre, and G. Visaggio, Eds. Springer, 2011, pp. 143–155.

[27] N. Hurst, *Risk Assessment: The Human Dimension*. The Royal Society of Chemistry, 1998.

[28] T. Ishimatsu, N. G. Leveson, J. P. Thomas, C. H. Fleming, M. Katahira, Y. Miyamoto, R. Ujiie, H. Nakao, and N. Hoshino, “Hazard analysis of complex spacecraft using Systems-Theoretic Process Analysis,” *Journal of Spacecraft and Rockets*, vol. 51, no. 2, pp. 509–522, 2014.

[29] ISO 27002:2005, “Information technology - Security techniques - Code of practice for information security management,” 2005.

[30] ISO 27005:2011, “Information technology - Security techniques - Information security risk management,” 2011.

[31] ISO 31000:2009, “Risk management - principles and guidelines,” 2009.

[32] H. Jansen, “The logic of qualitative survey research and its position in the field of social research methods,” *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, vol. 11, no. 2, p. 21, 2010.

[33] P. Johannessen, F. Torner, and J. Torin, “Actuator based hazard analysis for safety critical systems,” in *Computer Safety, Reliability, and Security*, ser. Lecture Notes in Computer Science, vol. 3219. Springer, 2004, pp. 130–141.

[34] J. Johansson, “Risk and vulnerability analysis of interdependent technical infrastructure,” Ph.D. dissertation, Lund University, Sweden, 2010.

[35] S. Kaplan and B. J. Garrick, “On the quantitative definition of Risk,” *Risk Analysis*, vol. 1, no. 1, pp. 11–27, 1981.

[36] B. Kitchenham, D. Budgen, and O. P. Brereton, “Using Mapping Studies as the Basis for Further Research - A Participant-Observer Case Study,” *Information and Software Technology*, vol. 53, pp. 638–651, June 2011.

[37] B. Kitchenham and S. Charters, “Guidelines for performing systematic literature reviews in software engineering,” *Technical Report Keele University and University of Durham*, vol. Version 2.3, 2007.

[38] B. A. Kitchenham, D. Budgen, and P. Brereton, *Evidence-Based Software Engineering and Systematic Reviews*. Chapman & Hall/CRC, 2015.

[39] J. Kontio, L. Lehtola, and J. Bragge, “Using the focus group method in software engineering: obtaining practitioner and user experiences,” in *Proceedings of the International Symposium on Empirical Software Engineering (ISESE)*, Aug 2004, pp. 271–280.

[40] Krisberedskaps Myndigheten (KBM), “Hot och riskrapport (in swedish),” Västerås Sweden, Tech. Rep., 2005.

[41] O. Laitenberger, K. E. Emam, and T. G. Harbich, “An internally replicated quasi-experimental comparison of checklist and perspective based reading of code documents,” *IEEE Trans. Software Engineering*, vol. 27, no. 5, pp. 387–421, 2001.

[42] F. Landegren, “Technical infrastructure networks as socio-technical systems: Addressing infrastructure resilience and societal outage consequences,” PhD Thesis, Lund University, Sweden, 2017.

[43] F. Landegren, J. Johansson, and O. Samuelsson, “Comparing societal consequence measures of outages in electrical distribution systems,” in *Safety and Reliability: Methodology and Applications*, 2014, pp. 189–196.

[44] J. Laudon and K. Laudon, *Management Information Systems: Managing the Digital Firm (10th ed)*. Prentice-Hall, 2006.

[45] T. C. Lethbridge, S. E. Sim, and J. Singer, “Studying software engineers: Data collection techniques for software field studies,” *Empirical Software Engineering*, vol. 10, no. 3, pp. 311–341, 2005.

[46] N. G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*. The MIT Press, 2012.

[47] N. G. Leveson, C. H. Fleming, M. Spencer, J. Thomas, and C. Wilkinson, "Safety assessment of complex, software-intensive systems," *SAE International Journal of Aerospace*, vol. 5, no. 1, 2012.

[48] C. Lindholm, J. P. Notander, and M. Höst, "A case study on software risk analysis and planning in medical device development," *Software Quality Journal*, vol. 22, no. 3, pp. 469–497, 2014.

[49] J. A. McDermid, M. Nicholson, D. J. Pumfrey, and P. Fenelon, "Experience with the application of HAZOP to computer-based systems," in *proceedings of the 10:th Annual Conference on Computer Assurance, 1995. COMPASS '95, Systems Integrity, Software Safety and Process Security*, 1995, pp. 37–48.

[50] "Mehari 2010 – evaluation guide for security services," Methods Working Group, Club De La Securite De L'Information Francais (CLUSIF), Paris, Tech. Rep., 2010.

[51] M. W. Meyer and K. A. Solomon, "Risk-management practices in local communities," *Policy Sciences*, vol. 16, pp. 245–265, 1984.

[52] R. J. Mikulak, R. McDermott, and M. Beauregard, *The Basics of FMEA*. Productivity Press, paper back, 2008.

[53] G. Motta, G. Pignatelli, T. Barroero, and A. Longo, "Service Level Analysis method - SLAM," in *3rd International Conference on Computer Science and Information Technology*, vol. 5, July 2010, pp. 460–466.

[54] C. A. Murdock, M. Squeri, C. Jones, and B. S. Smith, "Risk Management in Non-DoD U.S. Government Agencies and the International Community: best practices and lessons learned," Center for Strategic and International Studies (CSIS), Tech. Rep., 2011.

[55] H. Nakao, M. Katahira, Y. Miyamoto, and N. G. Leveson, "Safety guided design of crew return vehicle in concept design phase using STAMP/STPA," in *proceedings of the 5:th International Association for the Advancement of Space Safety (IAASS) Conference*, 2011, pp. 497–501.

[56] P. G. Neumann, "Risks of Untrustworthiness," in *proceedings of the 22:nd Annual Computer Security Applications Conference*, 2006, pp. 321–328.

[57] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Information and Software Technology*, vol. 64, pp. 1 – 18, 2015.

[58] F. Redmill, M. Chudleigh, and J. Catmur, *System Safety : HAZOP and Software HAZOP*. John Wiley & Sons, 1999.

[59] B. Regnell, P. Runeson, and T. Thelin, “Are the perspectives really different? further experimentation on scenario-based reading of requirements,” *Empirical Software Engineering*, vol. 5, no. 4, pp. 331–356, Dec. 2000.

[60] G. Reith, “Uncertain times,” *Time & Society*, vol. 13, no. 2-3, pp. 383–402, 2004.

[61] C. Robson, *Real world research*, 2nd ed. Blackwell, 2002.

[62] RTCA/DO-312, “Safety, Performance, and Interoperability Requirements Document for the In-Trail Procedure in Oceanic Airspace (ATSA-ITP) Application,” RTCA Incorporate, Washington DC, Tech. Rep., 2008.

[63] P. Runeson, M. Höst, A. Rainer, and B. Regnell, *Case Study Research in Software Engineering: Guidelines and Examples*. Wiley, 2012.

[64] G. Sabaliauskaite, F. Matsukawa, S. Kusumoto, and K. Inoue, “An experimental comparison of checklist-based reading and perspective-based reading for UML design document inspection,” in *proceedings of the International Symposium on Empirical Software Engineering*, 2002, pp. 148 – 57.

[65] C. B. Seaman, “Qualitative methods in empirical studies of software engineering,” *IEEE Transactions on Software Engineering*, vol. 25, no. 4, pp. 557–572, Jul 1999.

[66] F. Sebastiani and C. N. Ricerche, “Machine learning in automated text categorization,” *ACM Computing Surveys*, vol. 34, pp. 1–47, 2002.

[67] SFS (1997:857), “Ellag, swedish code of statutes, stockholm,” (In Swedish).

[68] SFS (2006:544), “Lagen om kommuners och landstings åtgärder inför och vid extraordinära händelser i fredstid och höjd beredskap, swedish code of statutes, stockholm,” (In Swedish).

[69] G. Sindre and A. L. Opdahl, “Eliciting security requirements with misuse cases,” *Requirements Engineering*, vol. 10, no. 1, pp. 34–44, 2005.

[70] G. Stoneburner, A. Goguen, and A. Feringa, *Risk Management Guide for Information Technology Systems*, ser. National Institute of Standards and Technology, Special Publication 800-30. U.S. Government Printing Office, 2002.

[71] T. Stålhane and G. Sindre, “A comparison of two approaches to safety analysis based on use cases,” in *Conceptual Modeling - ER 2007*, ser. Lecture Notes in Computer Science, vol. 4801. Springer, 2007, pp. 423–437.

[72] A. Syalim, Y. Hori, and K. Sakurai, “Comparison of Risk Analysis Methods: Mehari, Magerit, NIST800-30 and Microsoft’s Security Management Guide,” in *proceedings of the International Conference on Availability, Reliability and Security, ARES ’09*, 2009, pp. 726–731.

[73] U.S. Dept. of Homeland Security, “The national strategy for the physical protection of critical infrastructures and key assets,” Tech. Rep., 2003.

[74] K. Weyns, “IT Dependability Management in Governmental Organisations,” Ph.D. dissertation, Lund University, Sweden, 2011.

[75] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in software engineering*. Springer Publishing Company, Incorporated, 2012.

[76] R. K. Yin, *Case Study Research: Design and Methods*. SAGE Publications, 2003.

[77] A. Yoran and L. J. Hoffman, “Role-based risk analysis,” in *proceedings of the 20:th National Information Systems Security Conference*, 1997, pp. 37–51.

[78] S. Yu, Q. Yang, J. Liu, and M. Pan, “A comparison of FMEA, AFMEA and FTA,” in *Proceedings of the 9th International Conference on Reliability, Maintainability and Safety (ICRMS)*, June 2011, pp. 954–960.

[79] Z. Zhang, V. Basili, and B. Shneideman, “Perspective-based usability inspection: an empirical validation of efficacy,” *Empirical Software Engineering*, vol. 4, no. 1, pp. 43 – 69, 1999.

INCLUDED PAPERS

PAPER I

A REVIEW OF RESEARCH ON RISK ANALYSIS METHODS FOR IT SYSTEMS

Abstract

Context: At the same time as our dependence on IT systems increases, the number of reports of problems caused by failures of critical IT systems has also increased. This means that there is a need for risk analysis in the development of this kind of systems. Risk analysis of technical systems has a long history in mechanical and electrical engineering. **Objective:** Even if a number of methods for risk analysis of technical systems exist, the failure behavior of information systems is typically very different from mechanical systems. Therefore, risk analysis of IT systems requires different risk analysis techniques, or at least adaptations of traditional approaches. This means that there is a need to understand what types of methods are available for IT systems and what research that has been conducted on these methods. **Method:** In this paper we present a systematic mapping study on risk analysis for IT systems. 1086 unique papers were identified in a database search and 57 papers were identified as relevant for this study. These papers were classified based on 5 different criteria. **Results:** This classification, for example, shows that most of the discussed risk analysis methods are qualitative and not quantitative and that most of the risk analysis methods that are presented in these papers are developed for IT systems in general and not for specific types of IT system. **Conclusions:** The results show that many new risk analysis methods have been proposed in the last decade but even more that there is a need for more empirical evaluations of the different risk analysis methods. Many papers were identified that propose new risk analysis methods, but few papers discuss a systematic evaluation of these methods or a comparison of different methods based on empirical data.

Sardar Muhammad Sulaman, Kim Weyns and Martin Höst,
In Proceedings of the 17:th International Conference on Evaluation and Assess-

ment in Software Engineering (EASE '13), pages 86–96. Association for computing machinery (ACM) 2013.

1 Introduction

IT systems have become an essential part of our modern society. This evolution has not only created new opportunities, but also new threats to our society. The presence of IT systems everywhere has made us dependent on IT systems for our daily life. This is the case both for individuals and organizations, both private as well as public organizations. However, at the same time as the usage of, and dependence on, IT systems increases, the number of reports of problems caused by failures of critical IT systems has also increased [18].

One of the common aspects of these failures is the faith in systems that are not sufficiently dependable. The core of the problem is not that these systems suddenly become unreliable, but that we have become critically dependent on a wide variety of systems without analyzing whether they are dependable enough and what the consequences could be of a possible failure [18]. To prevent critical systems from causing problems for the organizations dependent on them, risk analysis is a necessary activity.

Risk analysis of technical systems has a long history in mechanical and electrical engineering where many well-established methods exist. The failure behavior of IT systems is typically different from mechanical systems and, at the same time, the complexity can be significantly higher. The high rate at which new IT systems are being developed and updated for many critical applications usually means there is not enough historical data available for a strictly statistical analysis of the reliability of each system and its components, as is sometimes the case in risk analysis of mechanical systems.

For all these reasons, risk analysis of IT systems requires different risk analysis techniques or at least adaptations of these traditional risk analysis approaches. In this article we present a systematic overview of previously published research on risk analysis for IT systems.

Risk analysis can be performed during the development of the system, at deployment of the system or at any time afterwards. In the ideal situation, the risk analysis should be re-evaluated each time major changes occur in the system or in the environment in which the system is used.

In this article we present an overview of operational risk analysis methods for IT systems. This includes many different types of systems and methods, but does not include project risk analysis methods, used to analyze the project management risks in software development projects.

Section 2 presents related work in the field of risk analysis and systematic literature reviews. Section 3 discusses the methodology used in this study in detail.

Section 7 contains the special measures that were taken to improve the validity of this research. Next, Section 5 contains the results of this mapping study and presents the categorization of the identified articles based on different attributes of the research and the risk analysis methods presented in each article. Finally, Section 8 summarizes and analyses the results of this classification.

2 Related Work

Many different national and international high-level frameworks exist for information technology risk management and assessment. Such frameworks have for example been published by the International Organization for Standardization (ISO), such as ISO/IEC 27005 [7] and ISO/IEC 27002 [6], by national governmental organizations, such as the National Institute of Standards and Technology (NIST) [21] or the British Central Communication and Telecommunication Agency (CCTA) [2], by non-governmental organizations such as Club de la Sécurité de l'Information Français (CLUSIF) [15] or by research institutes such as the Carnegie Mellon Software Engineering Institute (SEI) [1]. A detailed comparison of some of these frameworks can for example be found in [4] and [22].

There also exist a number of low-level risk analysis methods for technical systems in general or for IT-systems in particular. Some of the most well-known methods are Fault Tree Analysis (FTA) [5], Failure Mode and Effect Analysis (FMEA) [16] and Hazard and operability study (HAZOP) [19]. Some of the frameworks mentioned above specifically recommend one or more of these risk analysis methods.

The goal of the study presented in this article is to identify research articles that describe or evaluate new or established risk analysis methods for IT systems, which includes both high- and low-level methods. To identify and categorize these research articles this study uses the methodology of mapping studies [11], which is a variation of systematic literature reviews [12].

Systematic literature reviews and mapping studies have been conducted in different studies [10] in widely different areas such as cost estimation (e.g. [8]), open source software (e.g. [20]), and testing (e.g. [3]). Two systematic reviews, [14] and [9], have focused on project risk assessment in software development projects. However, to the best of our knowledge, no reviews have looked specifically at operational risk analysis methods for IT systems.

3 Methodology

This article presents a study of available risk analysis, assessment, and management methods for IT systems. The review presented here is a systematic mapping study, conducted based on the guidelines presented in [12]. This article presents, in addition to the overview of the identified risk analysis methods, a categorization of the identified methods.

A review protocol was developed in the initial phase of the review. It contains research background, research questions, search strategy, study selection criteria and procedures, validity assessment, data extraction instructions, and data synthesis strategies.

This research is conducted as a planned study and was carried out in the following steps:

1. Defining the research questions.
2. Selection of sources to be searched for relevant articles.
3. Defining the search query and performing the search on the selected sources, resulting in 1203 articles.
4. Removing 117 duplicate articles by using EndNote reference manager and by manual search.
5. Defining the inclusion and exclusion criteria and initial selection based on titles and keywords according to the defined criteria, leaving 320 articles for the next steps of the study.
6. Second round of selection by reading abstracts according to the same criteria and first classification of the articles, leaving 200 articles for the next steps of the study.
7. Final selection of articles based on careful reading of the full text of each article, resulting in a final list of 57 relevant articles for this study.
8. Analysis of the results of the classification of the final list of articles.

During each step special measures were taken to improve the validity of the research. Each step is described in more detail in the following subsections.

The steps involved in the identification and selection of articles are summarized in Figure 1.

3.1 Research Questions

The objective of this article is, as described above, to present an overview of risk analysis methods for IT systems, by summarizing and synthesizing the results from research that has already been carried out on available risk analysis methods for IT systems. This general goal has been broken down to the following main research questions:

1. What risk analysis methods and approaches are reported in the research literature for IT-systems?

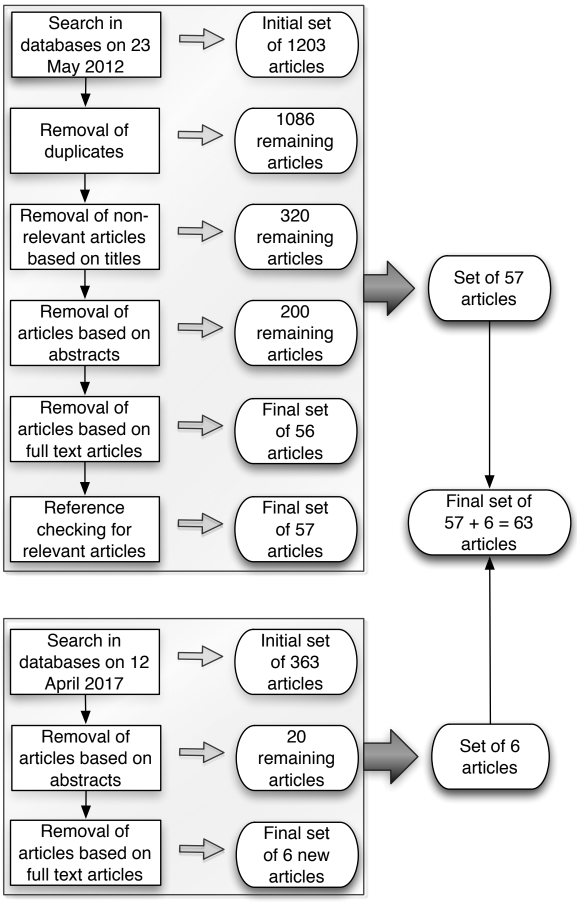

Search in databases on 23 May 2012
Initial set of 1203 articles
Removal of duplicates
1086 remaining articles
Removal of non-relevant articles based on titles
320 remaining articles
Removal of articles based on abstracts
200 remaining articles
Removal of articles based on full text articles
Final set of 56 articles
Reference checking for relevant articles
Final set of 57 articles
Set of 57 articles
Final set of 57 + 6 = 63 articles
Search in databases on 12 April 2017
Initial set of 363 articles
Removal of articles based on abstracts
20 remaining articles
Removal of articles based on full text articles
Final set of 6 new articles
Set of 6 articles

Figure 1: Identification and selection of articles.

2. To what extent are the identified methods used in practice?

3. Is there empirical research published where the identified methods are evaluated/compared/etc.? If there is, which research methodologies are used?

4. Which phases of the risk management process have been the focus of the identified research articles?

5. What type of risk analysis methods are presented in the published research, qualitative or quantitative?

This research can be categorized as a systematic mapping study that is carried out in the same way as a systematic review. It focuses on the main research that has been conducted in the area of risk analysis for IT systems, and it is done by adopting a systematic approach to identify relevant research and classify the identified research articles according to predefined categories.

3.2 Search Strategy

Searched Resources

The following databases were searched (through Engineering Village[1]) for relevant research:

- INSPEC: This database is provided by Elsevier Engineering Information Inc. and the Institute of Electrical Engineers (IEE). It includes articles from 1969 to present.

- COMPENDEX: This database is provided by Elsevier Engineering Information Inc. It includes papers from 1969 to present.

The above mentioned databases provide a broad coverage of the area of interest, i.e. “Risk analysis methods for IT systems”, and they include articles from the main conferences, journals, and publishers (IEEE, ACM, Springer, etc.).

Search Query

After a number of iterations, the following search query was considered a good compromise between finding as many of the relevant articles as possible, and returning a manageable number of results:

```
({risk analysis} OR
 {risk analyses} OR
 {risk identification} OR
 {RA})
```

[1] http://www.engineeringvillage2.org

```
AND (
 method* OR
 technique* OR
 approach*)
AND (
 {computer system} OR
 {information system} OR
 {IT system} OR
 {network system} OR
 {web?based system} OR
 {computer systems} OR
 {information systems} OR
 {IT systems} OR
 {network systems} OR
 {web?based systems})
NOT
( oil OR gas OR flood OR
 agricultur* OR chemi*))
```

The search string has four main parts separated by AND and NOT clauses.

The first part of the search string excludes articles that are not about ‘risk analysis’ or ‘risk identification’.

The second part of the search string excludes articles that do not discuss one or more specific methods for risk analysis, or a synonym to ’method’. The ’*’-character is a wildcard representing any string of characters, which allows different grammatical numbers of the term to be identified, e.g. both ’method’ and ’methods’.

The third part of the search string excludes articles that are not in the field of information technology or computer science. The ’?’-character is a wildcard representing one character, included because we want to identify both ’-’ and ’ ’.

The last part of the search string explicitly excludes articles about oil, gas, agriculture or chemistry. These research fields traditionally have a strong safety focus and contain many papers about risk analysis. They are, however, not domains in which IT systems are considered as the most critical components, and this part of the search string was included to prevent irrelevant papers from these domains from dominating the returned results.

3.3 Inclusion and Exclusion Criteria

When articles were identified with the search string from the databases, it was necessary to manually remove non-relevant articles from the selection. This was done first based on the title and keywords, then based on the abstract, and finally based on the full text. The inclusion and exclusion criteria were defined during the design of the review protocol. The manual selection of articles was carried out based on the following criteria:

- Articles not about methods for risk analysis or risk management of computer system were excluded from the selection.
- Articles about the risk analysis of system development projects were excluded from the selection. That is, articles about risk management of *project risks* were excluded. The focus in this article is on risks for the organization depending on the operation of IT systems, i.e. operational risk, not about the project risks associated with developing the systems.
- Articles specifically about the risk analysis of computer networks were excluded from the selection because the focus in this study is on risk analysis for complete IT systems not just the network component of the system. The excluded articles present risk analysis of network components such as, firewalls, intrusion detection systems, routers and implementation of security policies to cope with unauthorized access of data or resources, e.g., [17].
- Articles about the risk analysis of space systems, nuclear power plants, embedded medical devices, and military systems were also excluded from the selection. These domains have a long history of risk analysis methods, but these methods are often very time-consuming and mostly suited for embedded systems that are analyzed in great detail. This study, however, focuses on risk analysis for large IT systems that are applicable to a wide range of systems in many types of organizations. An example of excluded article is [13].

Each of these criteria was necessary to limit the scope of this study. It would be impossible to cover risk analysis for all types of risk associated with all categories of IT systems in one review like this, because of the large number of relevant articles.

3.4 Selection of Relevant Articles

The above mentioned search query was carried out two times, first on 23 May 2012 and second on 12 April 2017 to update the mapping study. On 23 May 2012 it retrieved 1203 articles, and it has been decided to continue systematic review with these records. After this, the title, keywords, abstract and author names were downloaded for the initial selection of relevant articles. Then the EndNote (Reference manager) was used for the removal of duplicate articles. It found (automatically) 91 duplicate articles that have been removed from the initial list. After this, 26 duplicate articles were found by manual search and removed from the initial list as well.

In each step of the selection process (based first on the title and keywords, then on the abstract and then on the full text) these criteria were used by the first author of this article to manually remove non-relevant articles from the initial selection. This resulted, in each step, in three groups of articles:

- **Relevant:** Articles that clearly fulfill the criteria established above.
- **Not relevant:** Articles that are out of the scope of this study.
- **Possibly relevant:** Articles for which there was not enough information to establish whether they are relevant for this study. This list was rechecked by the co-authors for the selection. The remaining Articles (from selection based on title and keywords, and abstracts) were then added to the relevant articles for further selection in the next step.

After removing irrelevant articles based on the title and keywords, a first effort to remove non-relevant articles was carried out by the first author of this article. This selection resulted in a list containing 229 relevant and 48 possibly relevant articles. To check the reliability of this first step, the second author of this article cross checked 100 randomly chosen articles from the initial list and found disagreement on 3 relevant articles not added and 6 non-relevant articles added. To increase the reliability of the selection, it was therefore decided to repeat the initial selection process based on this information and to only exclude those articles that were not relevant in light of this. The selection process was by this conducted once again and resulted in 70 more articles from the initial list to the main selected list. After this, the possibly relevant articles list was checked and 21 out of 48 articles were selected and added in the main list for the next step of review. After doing the initial selection process again the resulted selection list came up with a total of 320 relevant articles.

The second selection was conducted based on the abstracts, the first author read the abstracts and found 183 relevant articles out of a total of 320. The second author again rechecked this selection and he found 17 more relevant articles. After adding these 17 articles the second selection list came up with a total of 200 relevant articles for the next step of review.

In the third step of the selection process, the full text of the relevant articles needed to be downloaded. The full text for all articles was not always available for all articles and 57 articles were removed from the selection because the articles were not written in English (most often in Chinese) or because the full text could not be downloaded (mostly older articles).

After this, the first author carefully read the full text of all downloaded articles and selected 77 relevant articles. The second author of this article cross-checked the excluded articles from the final list suggested adding two more relevant articles in the final list, which resulted in 79 relevant articles. Then, he cross checked the finally selected articles by reading the full text and removed 23 irrelevant articles.

After removing the irrelevant articles the list contained 56 relevant articles. There was a disagreement for the selection of [article 24], the third author carefully read it, and after discussion all authors agreed to select it for the review.

Finally, the reference lists of the most relevant articles were inspected for further relevant articles that were not included in the selection. Initially 5 articles were selected from reference inspection, after reading the full text of selected articles only one article identified as relevant for this study. This article was from a source that was not included in the searched resources. After adding this article the final list contains the 57 articles listed in the appendix of this article.

The same search query was used again on 12 April 2017 to update the mapping study by finding the newly published relevant literature between 2012-2017. It retrieved 363 articles. All the same steps and selection criteria mentioned above were used and carried out for this selection. After reading titles, keywords, and abstracts of the articles the first author found 20 relevant articles out of a total of 363. In the next step of the selection process, the full texts of the relevant articles were downloaded. The full texts of two articles were not found and one article was in Chinese, this left 17 relevant articles with full text. After this, the first author carefully read the full text of all downloaded articles and 6 were selected as relevant articles for this study. It should be noted that the new 6 selected articles will be referred as 1n, 2n, 3n, 4n, 5n, and 6n in this study.

3.5 Data Extraction and Synthesis

In the final steps of the selection, i.e. the selection based on the full text of the articles, the articles were classified based in the following classes:

Class A Articles describing or evaluating existing risk analysis methodologies.

Class B Articles presenting improvements or changes to existing risk analysis methodologies.

Class C Articles presenting new methods for risk analysis of IT systems.

Further, a number of relevant attributes were also extracted from each of the articles with respect to the research questions discussed in Section 3.1. The results of this data extraction and classification are discussed in Section 5.

4 Validity Assessment

The main objective of this research is to summarize the available research in the field of risk analysis for IT systems. An important threat to the validity of this study is that it cannot be guaranteed that all possible relevant articles in this field have been included in the study. First of all, only research published in English was included for practical reasons. Secondly, some lesser known journals or conferences are not available in the searched databases, and were therefore not searched in this study. Also, articles for which the full text was not available were excluded from this study. This mostly affects older articles. Thirdly, it is likely that some relevant articles were rejected by the search string, since it is impossible to define a search string that finds absolutely all relevant articles without returning an unmanageable number of false positives. Finally, it is of course also possible that relevant articles were incorrectly rejected during the manual selection process from over one thousand articles to the final selection of 57 articles.

To increase the validity of this study, the reference list of the most relevant articles from the final selected list were examined for missing important articles. This validity check resulted in only one new article being added to the selection of articles. This article had not been found in the automatic search because it was from a source not included in the searched databases.

In order to reduce the risk of incorrect rejection of an article during the selection process, the co-authors of this article cross-checked the selection in each step. Whenever there was doubt about whether to include an article or not, the article was retained for the next step of the selection process. After initial selection process based on the title and keywords, the second author of this article cross checked 100 randomly selected articles from the initial list, and suggested a few additions and removals of articles. Instead of just adding and removing these articles, it was decided to repeat the selection process and to keep any articles selected in either case.

After the second selection process based on abstracts, the second author of this article re-checked the complete selection and found 17 more relevant articles that had possibly been rejected incorrectly, and in this way made sure that also articles where we were in doubt were included.

After the third selection process that was conducted after reading the full text of articles, the second author of this article cross checked the excluded articles from the final list and suggested the adding of two more relevant articles to the final list. Then he cross checked the finally selected articles by reading their full text and found 23 non-relevant articles according to the defined research questions.

That is, whenever there was a doubt in selection of an article it was retained for the next step, where more information was available to decide the relevance of an article with more accuracy. Whenever one author was not sure about the classification of an article, the co-authors reviewed the article and decision about the classification was based on the agreement by all authors.

By taking the above mentioned measures for the validity of this study we are more confident that most of the relevant articles for this study have been identified and included in the final list of articles.

Table 1: Classification of articles

Classification	articles	#
Class A	2, 4, 5, 7, 8, 10, 12, 14, 18, 21, 22, 26, 32, 38, 41, 45, 52, 56	18
Class B	34, 42, 43, 44, 47	05
Class C	1, 3, 6, 9, 11, 13, 15, 16, 17, 19, 20, 23, 24, 25, 27, 28, 29, 30, 31, 33, 35, 36, 37, 39, 40, 46, 48, 49, 50, 51, 53, 54, 55, 57, 1n, 2n, 3n, 4n, 5n, 6n	40

5 Results

This section presents an analysis of the data extracted from the selected articles.

5.1 Year of Publication

In Figure 2, the publication year for the selected articles is displayed. It can be observed that the oldest selected article is from the year 1980, and the most recent from 2012. About half of the articles were published in the last 5 years before the publication of this study. That is, this indicates that the number of publications in the area has increased the later years, at least if we were able to find as many of the older articles as the newer articles.

5.2 Risk Analysis Method Classification

Table 1 shows the classification of the selected articles into classes A, B, and C, see Section 3.5. Class A, about existing risk analysis methods, includes 18 articles. Articles in this class describe general risk analysis concepts and its importance for dependable IT systems. This class also contains some articles about the comparison of different risk analysis methods. Class B includes 5 articles that present improvements in existing risk analysis methods.

The majority of the articles are in class C. It includes 34 from previous search and 6 articles from the new search that are about presenting new frameworks, methods and models for risk analysis. They are in total 40 articles.

5.3 Types of Systems

Table 2 shows the types of system that the selected articles focus on. The majority of the selected articles, 49 articles out of 57 from the previous search, are about risk analysis of IT systems in general. This means that the paper does not specify

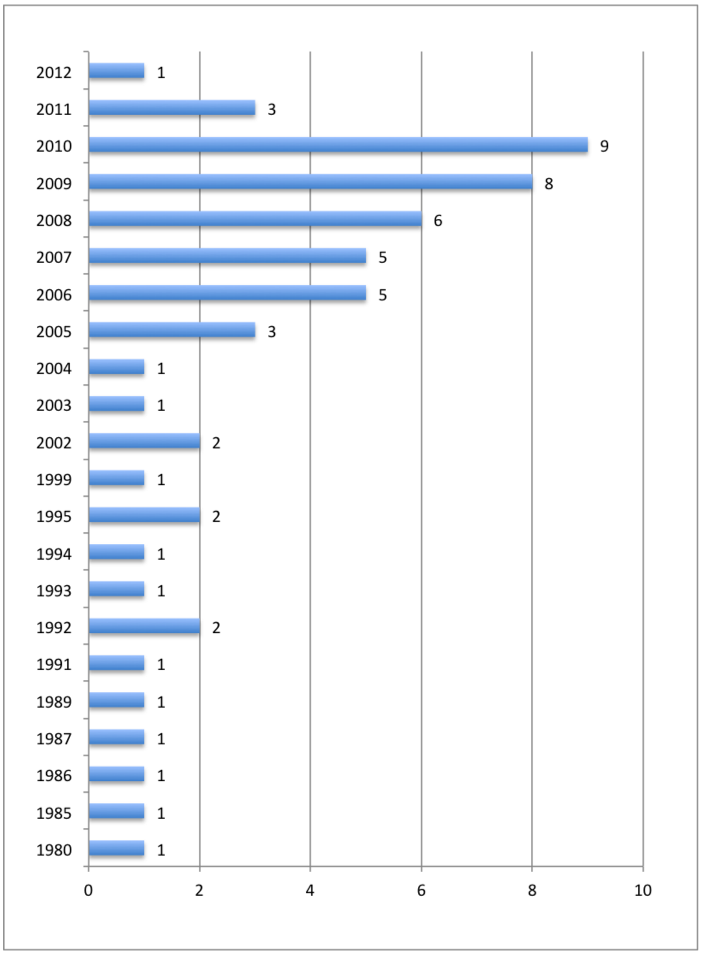


Figure 2: Histogram of publication year for the identified articles

Table 2: Focused systems in selected articles

Type of System	articles	#
IT systems in general	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 22, 23, 24, 26, 27, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 52, 53, 54, 55, 56, 1n, 2n, 3n, 4n, 5n, 6n	55
Hospital systems	2, 32, 57	3
E-Commerce	25, 50	2
Cloud computing	16	1
E-government	51	1
Web-service systems	21	1

which type of systems the research is about, and thereby it can be assumed that the intention is that the research results should be generally valid. However, 2 articles are specifically about risk analysis for e-commerce systems, 3 are about hospital systems, 1 is specifically about web service systems, 1 is about cloud computing and 1 is about e-government systems. It should be noted that articles about space technology and military systems were specifically excluded before the classification.

All 6 selected articles from the new search are about risk analysis of IT systems in general.

5.4 Analytical or Empirical Research

In Table 3, the research methodologies that were used in the selected articles are categorized as either completely *analytical* (not containing any research based on the application of a risk analysis method on an actual system) or *empirical* (containing an explicit description of an application of at least one risk analysis method, either in a real-life setting or in a controlled experiment). 36 articles were identified as analytical and 21 as empirical research. These 21 articles all presented case studies on risk analysis methods, no surveys or experiments were identified.

From the new search only one empirical article (4n) as a case study was found. With this article total number of case studies found is 22.

5.5 Area of Risk Management

Risk management is a process that consists of several activities: risk identification, risk analysis, risk assessment, risk prioritization, and risk mitigation. It is a process that tries to find a balance between loss prevention and cost associated with

Table 3: Type of research presented in selected articles

Research type	Selected articles	#
Analytical	1, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 18, 19, 20, 22, 23, 24, 25, 27, 29, 30, 31, 33, 37, 38, 40, 42, 43, 44, 45, 47, 48, 49, 50, 53, 1n, 2n, 3n, 5n, 6n	41
Empirical		
-Case study	2, 7, 15, 16, 17, 21, 26, 28, 32, 34, 35, 36, 39, 41, 46, 51, 52, 54, 55, 56, 57, 4n	22

countermeasures. It usually starts with the *risk identification* activity to determine a list of possible risks. Next, *risk analysis* is applied to combine the probability and the expected consequences associated with each risk. Sometimes the term 'risk analysis' is also used to include the risk identification step. Then, in *risk prioritization*, all the identified risks are prioritized based on the results of the risk analysis. Finally, *risk mitigation*, deals with implementing appropriate measures and controls to reduce the probability or the consequences of the identified risks, based on the results of the prioritization. *Risk assessment*, on the other hand, usually deals with the analysis of a system with existing security measures and anticipates the weaknesses present in assessed system. However, these definitions are not generally accepted and sometimes each of these terms is used to describe a process that includes several of the other activities.

Although our search for articles specifically searched for articles about risk analysis or risk identification, the final list of selected articles contain some articles that mainly focus on risk management as a whole and some articles that focus only on one or more of the different sub-activities. Table 4 shows the focus of the selected articles within the field of risk management. It can be noticed that the majority of selected articles, 28 articles out of 57, are in fact about risk analysis. Further, it can be seen that 1 article is specifically about risk identification, 15 are about risk assessment, 1 is about risk prioritization, 2 are about risk mitigation and 20 are about risk management as a whole.

From the new search one article (5n) is about risk prioritization, and all other articles are about risk analysis in general.

5.6 Qualitative and Quantitative Risk Analysis

Table 5 classifies the risk analysis methods in the selected articles as quantitative or qualitative. Quantitative methods express the probability and consequences of the identified risk as a numerical result. This makes it possible to calculate the relationship between loss prevention and cost associated with proposed counter-

Table 4: Focused risk management part in the selected articles

Risk management part	Selected articles	#
Risk analysis	1, 2, 3, 4, 5, 8, 11, 13, 14, 15, 19, 20, 21, 23, 24, 25, 26, 27, 29, 32, 34, 35, 40, 42, 47, 50, 56, 57, 1n, 2n, 3n, 4n, 6n	33
Risk identification	32	1
Risk assessment	5, 7, 9, 16, 31, 33, 36, 37, 39, 44, 45, 46, 52, 53, 55	15
Risk prioritization	16, 5n	2
Risk mitigation	12, 52	2
Risk management	4, 9, 10, 16, 17, 18, 19, 20, 21, 22, 27, 29, 30, 32, 35, 38, 43, 47, 48, 49	20

Table 5: Type of risk analysis method (quantitative or qualitative)

Risk analysis type	Selected articles	#
Qualitative	2, 12, 13, 14, 21, 34, 57, 3n	8
Quantitative	4, 5, 15, 19, 23, 24, 25, 26, 28, 31, 33, 36, 37, 39, 40, 41, 44, 49, 52, 53, 54, 55, 56, 1n, 2n, 4n, 5n, 6n	28
Combined approach	9, 10, 18, 42, 45, 46	6
Semi-Quantitative	3, 16	2

measures. Often it is difficult to use quantitative risk analysis because it is hard to estimate the exact probability and loss associated with each risk. Qualitative methods, on the other hand, use descriptive values such as 'high', 'medium' or 'very low' to express the probability and consequences of each risk. Both types of risk analysis methods are widely used for different types of systems, and in some cases they can be used together. Except for qualitative, quantitative and combined risk analysis methods, this study also identified semi-quantitative methods. This is an intermediary risk analysis technique that classifies the probability and consequences by using quantitative categories such as 'financial loss between 10.000 USD and 100.000 USD' or 'less than once per 100 years'. It does not require the exact estimates needed for a quantitative risk analysis, but offers a more consistent approach than qualitative risk analysis. Not all of the selected articles contain enough information to determine whether a qualitative or quantitative approach was used, and for some articles the question is not applicable. Of the 38 articles that could be classified according to this criterion, 23 articles use a quantitative approach, 7 a qualitative approach, 6 contain a combined (quantitative and qualitative) risk analysis approach, and 2 are about semi-quantitative risk analysis methods.

From the new search one article (3n) presents a qualitative risk analysis however all other articles present quantitative type risk analysis.

6 Discussion

First of all it can be observed that, from the previous search, many of the identified articles have been published during the last few years before this study (2006-2011). This may mean that the amount of research has increased. As also discussed above, there may be other reasons, such as that the databases are more complete for later years. However, an increased dependence on information in the society, e.g. when critical processes to an increased extent are supported by IT-systems, may also mean that there is an increased interest in risk management of IT-systems.

In order to investigate the relationship between different investigated factors different pairs of variables were investigated.

It was found that risk management papers are to a larger extent non-empirical than papers in the other categories, see Table 6. This may be because this topic requires more research effort to be studied empirically since it is a process covering a rather long time-span.

Risk analysis methods for a specific type of systems are all found in empirical papers, except for the papers about e-commerce systems. This probably indicates that most risk analysis methods are developed with general IT systems in mind. Only when they are applied in practice they are adapted for specific classes of systems.

Table 6: Paper area vs. empirical or not

Paper area	No	Yes
BCP	0	1
General	1	1
Risk Analysis	15	6
Risk Analysis and Assessment	1	0
Risk Analysis and Management	7	2
Risk Assessment	7	8
Risk Assessment and management	1	0
Risk Assessment and mitigation	0	1
Risk Assessment, prioritization,	1	
and management	0	1
Risk Identification, analysis, and management	0	1
Risk management	7	1
Risk Mitigation	1	0
SUM	41	22

Table 7: Risk analysis approach vs. empirical or not

Approach	No	Yes
Combined approach	5	1
General	14	5
Qualitative	4	4
Quantitative	17	11
Semi-quantitative	1	1
SUM	41	22

It was also found that qualitative risk analysis methods are more likely to be investigated in empirical papers than quantitative analysis methods, see Table 7. This may be because quantitative methods are not as easy in practice as it might seem, because a lot of specific data is needed. When a risk analysis method is used in practice, it is often easier to classify a risk's probability and consequence into some categories than to assign an exact numerical value. This however limits the analysis that can be done later. A lot of information is lost when categories are used instead of a quantitative best estimate, possible combined with an explicit uncertainty range on the estimate.

It can be noticed from the previous section that the majority of the identified articles present either qualitative or quantitative risk analysis and only two articles (3, 16) use a semi-quantitative risk analysis method. Based on this, it could be argued that there is a need for more research on techniques and methods that combine the advantages of both quantitative and qualitative methods.

Two identified articles (9, 10) present research on the well-known risk analysis method CORAS, performing model-based risk analysis by using UML, and one article (54) that proposes a new risk analysis method using fault-tree analysis. This review also has identified some other specific risk analysis methods named in a few articles such as LAVA, LRAM , CRAMM, OCTAVE, Mehari and Magerit (20, 34, 45, 47).

This review has identified five articles (2, 42, 43, 44, 47) that describe, analyze and compare existing well-known risk analysis methods. But from these articles it is not possible to decide that a particular method is better than other.

7 Conclusions

Based on this mapping study of risk analysis methods for IT-systems discussed in the research literature, it can be concluded that most articles focus on new methods, and new frameworks and models for risk analysis. Only few papers focus on already available, and thereby maybe already known, methods. Further, it can be concluded that most research concerns general risk analysis methods, and not methods specific to certain types of IT systems.

The fact that only few articles focused on already available methods also means that it is not possible to say from the identified articles to what extent different methods are used in practice. For the same reason, it has not been possible to find many articles comparing available risk analysis methods, even if we argue that there is a need for this kind of research.

It can also be concluded that a majority of the identified articles present research that is non-empirical (41 articles), and fewer articles (22 articles) present case studies. None of the identified articles present research conducted as surveys or controlled experiments. Concerning what type of risk analysis methods that are presented in the published research, it can be concluded that most identified research concerns quantitative risk analysis methods.

Based on these findings a number of areas for further research can be identified. First of all it can be concluded that there is a need to conduct research where already available methods are investigated. This can for example be carried out as studies where different types of methods are compared in controlled experiments. We believe that methods for risk analysis are quite possible to investigate in controlled experiments [23], since they are possible to isolate from the whole management process to investigate them in a ‘laboratory’ setting. Having said that, we also believe that there is a need to further investigate the whole risk management process in longer case studies, where actual cases of risk management are investigated in practice.

ACKNOWLEDGEMENT

This work was funded by the Swedish Civil Contingencies Agency under a grant for PRIVAD, Programme for Risk and Vulnerability Analysis Development.

LIST OF SELECTED ARTICLES

(1) Beachboard, J., Cole, A., Mellor, M., Hernandez, S., Aytes, K., Massad, N., Improving information security risk analysis practices for small-and medium-sized enterprises: a research agenda, Journal of Issues in Informing Science and Information Technology Journal, vol. 5, pp. 73-85, 2008.

(2) Bennett, S.P., An application of qualitative risk analysis to computer security for the commercial sector, In proceedings of Eighth Annual Computer Security Applications Conference (Cat. No.92TH0470-5), pp. 64-73, 1992.

(3) Birch, D.G.W., McEvoy, N.A., Risk analysis for information systems, Journal of Information Technology, vol. 7, issue 1, pp. 44-53, March 1992.

(4) Bojanc, R., Jerman-Blazic, B., An economic modeling approach to information security risk management, International Journal of Information Management, vol. 28, issue 5, pp. 413-22, 2008.

(5) Breier, J., Risk analysis supported by information security metrics, In proceedings of 12:th International Conference Computer Systems and Technologies, pp. 393-398, 2011.

(6) Chivers, H., Information modeling for automated risk analysis, In proceedings of 10:th International Conference Communications and Multimedia Security (CMS), pp. 228-239, 2006.

(7) Coles-Kemp, L., Triangulating the views of human and non-human stakeholders in information system security risk assessment, In proceedings of the 2007 International Conference on Security & Management, SAM 2007, pp. 172-178, 2007.

(8) De Koning, W.F., A methodology for the design of security plans, Computers & Security, vol. 14, issue 7, pp. 633-643, 1995.

(9) Djordjevic, I., Suitability of risk analysis methods for security assessment of large scale distributed computer systems, In proceedings of the 6:th Conference of International Association of Probabilistic Safety Assessment and Management, 23-28 June, San Juan, Puerto Rico, USA, 2002.

(10) Djordjevic, I., Model based risk management of security critical systems, In proceedings of the 3:rd International Conference on Computer Simulation in Risk Analysis and Hazard Mitigation, pp. 253-264, 2002.

(11) Eloff, J.H.P., Labuschagne, L., Badenhorst, K.P., Comparative framework for risk analysis methods, Computers & Security, vol. 12, issue 6, pp. 597-603, 1993.

(12) Eom, Jung-Ho, Qualitative method-based the effective risk mitigation method in the risk management, In proceedings of the International Conference on Computational Science and its Applications (ICCSA), pp. 239-248, 2006.

(13) Eom, Jung-Ho, Risk assessment method based on business process-oriented asset evaluation for information system security, In proceedings of the International Conference on Computational Science (ICCS), pp. 1024-1031, 2007.

(14) Eom, Jung-Ho, Qualitative initial risk analysis for selecting risk analysis approach suitable for IT security policy, In proceedings of the International Conference on Information Theory and Information Security, pp. 669-673, 2010.

(15) Feng, Nan, A probabilistic estimation model for information systems security risk analysis, In proceedings of the International Conference on Management and Service Science (MASS), p. 4, 2009.

(16) Fito, J.O., Macias, M., Guitart, J., Toward business-driven risk management for Cloud computing, In proceedings of the 6:th International Conference on Network and Service Management (CNSM 2010), pp. 238-241, 2010.

(17) Ghernouti-Helie, S., Reasonable security by effective risk management practices: From theory to practice, In proceedings of the 12:th International Conference on Proceedings of the 2009 12th International Conference on Network-Based Information Systems (NBiS 2009), p 226-33, 2009.

(18) Grob, H. L., Conceptual modeling of information systems for integrated IT-risk and security management, In proceedings of the 2008 International Conference on Security and Management (SAM), pp. 178-184, 2008.

(19) Guarro, S.B., Principles and procedures of the LRAM approach to information systems risk analysis and management, Computers & Security, vol. 6, issue 6, pp. 493-504, 1987.

(20) Guarro, S.B., Risk analysis and risk management models for information systems security applications, Reliability Engineering & System Safety, vol. 25, issue 2, pp. 109-130, 1989.

(21) GutiŐrrez, C., Rosado, G. D., FernǦndez-Medina, E., The practical application of a process for eliciting and designing security in web service systems, Information and Software Technology, vol. 51, issue 12, pp. 1712-1738, 2009.

(22) Hamdi, M., Boudriga, N., Computer and network security risk management: Theory, challenges, and countermeasures, International Journal of Communication Systems, vol. 18, issue 8, pp. 763-793, 2005.

(23) Hu, Zhi-Hua, Knowledge-based framework for real-time risk assessment of information security inspired by danger model, In proceedings of the International Symposium on Intelligent Information Technology, pp. 1053-1056, 2008.

(24) In, H.P., A security risk analysis model for information systems, Systems Modeling and Simulation: Theory and Applications, In proceedings of the Third Asian Simulation Conference, AsiaSim 2004, Revised Selected Papers (Lecture Notes in Computer Science Vol.3398), pp. 505-513, 2005.

(25) Jung, C., Han, I., Suh, B., Risk analysis for electronic commerce using case-based reasoning, International Journal of Intelligent Systems in Accounting, Finance and Management, vol. 8, issue 1, pp. 61-73, 1999.

(26) Kaegi, M., Information systems' risk analysis by agent-based modeling of business processes, In proceedings of the European Safety and Reliability Conference (ESREL) - Safety and Reliability for Managing Risk, 2006.

(27) Kailay, M. P.; Jarratt, P., RAMeX: a prototype expert system for computer security risk analysis and management, Computers & Security, vol. 14, issue 5, pp. 449-463, 1995.

(28) Kim, Young-Gab, Quantitative risk analysis and evaluation in information systems: A case study, In proceedings of the 7:th International Conference on Computational Science (ICCS), pp. 1040-1047, 2007.

(29) La Corte, A., A Process Approach to Manage the Security of the Communication Systems with Risk Analysis Based on Epidemiological Model, In proceedings of the 5:th International Conference on Systems and Networks Communications (ICSNC), pp. 166-171, 2010.

(30) Li Helgesson, Y.Y., Managing risks on critical IT systems in public service organizations, In proceedings of the 2009 International Conference on Computational Science and Engineering (CSE), pp. 470-475, 2009.

(31) Li, He-Tian, Security risk evaluation for it systems based on the Markov chain, Journal of the China Railway Society, vol. 29, issue 2, pp. 50-53, 2007.

(32) Lindholm, C, Pedersen Notander, J., Höst, M. Software Risk Analysis in Medical Device Development, In proceedings of the 37:th EUROMICRO Conference on Software Engineering and Advanced Applications, pp. 362-365, 2011.

(33) Lu, Simei, Security risk assessment model based on AHP/D-S evidence theory, In proceedings of 2009 International Forum on Information Technology and Applications (IFITA), pp. 530-534, 2009.

(34) Maglogiannis, I., Zafiropoulos, E., Platis, A., Lambrinoudakis, C., Risk analysis of a patient monitoring system using Bayesian Network modeling, Journal of Biomedical Informatics, vol. 39, issue 6, pp. 637-647, 2006.

(35) McGaughey Jr. R.E., Snyder, C.A., Carr, H.H., Implementing information technology for competitive advantage: risk management issues, Information & management, vol. 26, issue 5, pp. 273-280, 1994.

(36) Mock, R., Risk analysis of information systems by event process chains, International Journal of Critical Infrastructures, vol. 1, issue 2-3, pp. 247-257, 2005.

(37) Mosleh, A., Bayesian probabilistic risk analysis for computer systems, Performance Evaluation Review, vol. 13, issue 1, pp. 5-12, 1985.

(38) Nassar, P.B, Risk management and security in service-based architectures, In proceedings of the International Conference on Advances in Computational Tools for Engineering Applications (ACTEA), pp. 214-218, 2009.

(39) Patel S.C., Graham J.H., Ralston P.A.S., Quantitatively assessing the vulnerability of critical information systems: A new method for evaluating security enhancements, International Journal of Information Management, vol. 28, issue 6, pp. 483-491, 2008.

(40) Pirzadeh, L., A Cause and Effect Approach towards Risk Analysis, In proceedings of the 3:rd International Workshop on Security Measurements and Metrics (Metrisec), pp. 80-83, 2012.

(41) Post, G. V., Diltz, J. D., A stochastic dominance approach to risk analysis of computer systems, Management Information Systems Quarterly, vol. 10, issue 4, pp. 363-374, 1986.

(42) Rainer, R. K., Snyder, C. A., Carr, H. H., Risk analysis for information technology, Journal of Management Information Systems, vol. 8, issue 1, pp. 129-147, 1991.

(43) Sarkheyli, A., Improving the current risk analysis techniques by study of their process and using the human body's immune system, In proceedings of the 5:th International Symposium on Telecommunications (IST), pp. 651-656, 2010.

(44) Satoh, N., Kumamoto, H., Kino, Y., Norihisa, K., Viewpoint of ISO GMITS and PRA in information assessment, In proceedings of the 8:th conference on Applied Computer Science, pp. 253-258, 2008.

(45) Smith, S.T., LAVA, Proceeding of 12:th National Computer Security Conference, Baltimore, MD, USA, 1989.

(46) Sun, L., Srivastava, R. P., Mock, T. J., An information systems security risk assessment model under the Dempster-Shafer theory of belief functions, Journal of Management Information Systems, vol. 22, issue 4, pp. 109-142, 2006.

(47) Syalim, A., Comparison of risk analysis methods: Mehari, Magerit, NIST800-30 and Microsoft's security management guide, In proceedings of the International Conference on Availability, Reliability and Security (ARES), pp. 726-731, 2009.

(48) Trcek, D., Security metrics foundations for computer security, Computer Journal, vol. 53, issue 7, pp 1106-1112, 2010.

(49) Trcek, D., System dynamics based risk management for distributed information systems, In proceedings of the 4:th International Conference on Systems (ICONS), pp. 74-79, 2009.

(50) Warren, M., Hutchinson, W., A security risk management approach for e-commerce, Information Management & Computer Security, vol. 11, issue 5, pp. 238-242, 2003.

(51) Wei, G., Research on E-government Information Security Risk Assessment - Based on Fuzzy AHP and Artificial Neural Network Model, In proceedings of the 1:st International Conference on Networking and Distributed Computing (ICNDC 2010), pp 218-221, 2010.

(52) Wijnia, Y., Assessing business continuity risks in IT, In proceedings of the IEEE International Conference on Systems, Man and Cybernetics, pp. 3547-3553, 2008.

(53) Winkelvos, Timo, A property based security risk analysis through weighted simulation, In proceedings of the Information Security for South Africa (ISSA), 2011.

(54) Xiao, H., The research of information security risk assessment method based on fault tree, In proceeding of the 6:th International Conference on Networked Computing and Advanced Information Management (NCM), pp. 370-375, 2010.

(55) Xinlan, Z., Information security risk assessment methodology research: Group decision making and analytic hierarchy process, In proceedings of the 2:nd WRI World Congress on Software Engineering, pp. 157-160, 2010.

(56) Yan, H., Power information systems security: Modeling and quantitative evaluation, In proceedings of the IEEE Power Engineering Society General Meeting, pp. 905-910, 2004.

(57) Zain, N. M., Samy, F. N., Ahmad, R., Ismail, Z., Manaf, A. A., Fuzzy based threat analysis in total hospital information system, In Proceedings of the Advances in Computer Science and Information Technology, Lecture Notes in Computer Science, vol. 6059, pp. 1-14, 2010, Springer Berlin, Heidelberg.

(1n) Pirzadeh, L., Jonsson, E., A Cause and Effect Approach towards Risk Analysis, In Proceedings of the Third International Workshop on Security Measurements and Metrics (Metrisec), pp. 80-83, 2011, Los Alamitos, CA, USA.

(2n) Vicente, E., Jimenez, A., Mateos, A., A fuzzy extension of MAGERIT methodology for risk analysis in information systems, In Proceedings of the International Conference of Information Systems, pp. 39-46, 2013, New York, NY, USA.

(3n) El Fray, I., Kurkowski, M., Pejas, J., Mackow, W., A new mathematical model for analytical risk assessment and prediction in IT systems, In Journal of Control and Cybernetics, vol. 41, no. 1, pp. 241-268, 2012.

(4n) Bamakan, S.M.H., Dehghanimohammadabadi, M., A Weighted Monte Carlo Simulation Approach to Risk Assessment of Information Security Management System, In International Journal of Enterprise Information Systems, vol. 11, no. 4, pp. 63-78, 2015.

(5n) Eren-Dogu, Z.F., Celikoglu, C.C., Information security risk assessment: Bayesian prioritization for AHP group decision making, In International Journal of Innovative Computing, Information & Control, vol. 8, no. 11, pp. 8019-8032, 2012.

(6n) Niescieruk, A., Ksiezopolski, B., Motivation-based risk analysis process for IT systems, In Proceedings of the Second IFIP TC5/8 International Conference of Information and Communication Technology (ICT-EurAsia): LNCS 8407, pp. 446-455, 2014, Berlin, Germany.

Bibliography

[1] C.J. Alberts and A.J. Dorofee. *Managing Information Security Risks: The Octave Approach.* SEI Series in Software Engineering. Addison-Wesley, 2003.

[2] Central Computer and Telecommunications Agency, Great Britain, Treasury. *Prince User's Guide to CRAMM.* Programme and Project Management Library. 1993.

[3] E. Engström and P. Runeson. Software Product Line Testing - A Systematic Mapping Study. *Information and Software Technology*, 53:2–13, 2011.

[4] ENISA Ad Hoc Working Group on Risk Assessment and Risk Management. Inventory of risk assessment and risk management methods, 2006.

[5] Clifton A. Ericson. Fault Tree Analysis - A History. In *proceedings of The 17:th International System Safety Conference*, 1999.

[6] ISO 27002:2005. Information technology - Security techniques - Code of practice for information security management, 2005.

[7] ISO 27005:2011. Information technology - Security techniques - Information security risk management, 2011.

[8] M Jørgensen. A Review of Studies on Expert Estimation of Software Development Effort. *Journal of Systems and Software*, 70(1-2):37–60, 2004.

[9] M. A. Khan, S. Khan, and M. Sadiq. Systematic review of software risk assessment and estimation models. *International Journal of Engineering and Advanced Technology*, 1:298–305, 2012.

[10] 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*, 51(1):7–15, 2009.

[11] B. Kitchenham, D. Budgen, and O. P. Brereton. Using Mapping Studies as the Basis for Further Research – A Participant-Observer Case Study. *Information and Software Technology*, 53:638–651, June 2011.

[12] B. Kitchenham and S. Charters. Guidelines for performing systematic literature reviews in software engineering. *Technical Report Keele University and University of Durham*, Version 2.3, 2007.

[13] B. Li, M. Li, K. Chen, and C. Smidts. Integrating Software into PRA: A Software-Related Failure Mode Taxonomy. *Risk Analysis*, 26(4):997–1012, 2006.

[14] D. Liu, Q. Wang, and J. Xiao. The role of software process simulation modeling in software risk management: A systematic review. In *3rd International Symposium on Empirical Software Engineering and Measurement (ESEM)*, pages 302–311. IEEE, 2009.

[15] Mehari 2010 – evaluation guide for security services. Technical report, Methods Working Group, Club De La Securite De L'Information Francais (CLUSIF), Paris, 2010.

[16] Raymond J. Mikulak, Robin McDermott, and Michael Beauregard. *The Basics of FMEA, 2nd Edition*. Taylor & Francis, 2008.

[17] L. Mixia, Y. Dongmei, Z. Qiuyu, and Z. Honglei. Network Security Risk Assessment and Situation Analysis. In *proceedings of the 2007 IEEE International Workshop onAnti-counterfeiting, Security, Identification*, pages 448 –452, april 2007.

[18] P. G. Neumann. Risks of Untrustworthiness. In *proceedings of the 22:nd Annual Computer Security Applications Conference*, pages 321–328, 2006.

[19] F. Redmill, M. Chudleigh, and J. Catmur. *System Safety : HAZOP and Software HAZOP*. John Wiley & Sons, 1999.

[20] K. J. Stol and M. A. Babar. Reporting Empirical Research in Open Source Software: The State of Practice. In *proceedings of the International Conference on Open Source Systems, OSS 2009*, pages 156–169, 2009.

[21] G. Stoneburner, A. Goguen, and A. Feringa. *Risk Management Guide for Information Technology Systems*. National Institute of Standards and Technology, Special Publication 800-30. U.S. Government Printing Office, 2002.

[22] A. Syalim, Y. Hori, and K. Sakurai. Comparison of Risk Analysis Methods: Mehari, Magerit, NIST800-30 and Microsoft's Security Management Guide. In *proceedings of the International Conference on Availability, Reliability and Security, ARES '09*, pages 726–731, 2009.

[23] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén. *Experimentation in software engineering: an introduction*. Kluwer Academic Publishers, Norwell, MA, USA, 2000.

PAPER II

RISK ANALYSIS AND MANAGEMENT OF IT SYSTEMS: PRACTICE AND CHALLENGES

Abstract

Risk analysis is important for safety-critical IT systems and services, both in public and private organizations. However, the actual practices and the challenges of risk analysis in these contexts have not been fully explored. This paper investigates the current practices of risk analysis by an interview-based investigation. This study investigates several factors of the risk analysis process, e.g., its importance, identification of critical resources, definitions of roles, involvement of different stakeholders, used methods, and follow-up analysis. Furthermore, this study also investigates existing challenges in the current practices of risk analysis. A number of challenges are identified, e.g., that risk analysis requires competence both about the risk analysis procedures and the analyzed system, which is challenging to identify, and that it is challenging to follow-up and repeat a risk-analysis that is conducted. The identified challenges can be useful when new risk analysis methods are defined.

Sardar Muhammad Sulaman and Martin Höst,
International Conference on Information Systems for Crisis Response and Management (ISCRAM'18).

1 Introduction

Safety-critical systems are managed by both public and private organizations. Failures in these systems can have catastrophic consequences in society. These systems are typically subject to risk analysis in order to prevent that they unduly harm people, property, or the environment.

Today, both public and private organizations are facing different types of risks, e.g., policy, operational, project, financial, technological, health, safety, and human resources [2]. Risk analysis and management are often conducted in the organizations that are responsible for these safety-critical systems and provide critical services to society, which means that risk analysis is carried out by several different organizations when circumstances are changed. Therefore, clearly described methods and guidelines for conducting these analyses are needed. That is, an effective risk management approach is needed that requires a continuous assessment of potential risks in an organization at relevant lower levels and then it aggregates analysis results at the higher level to improve decision making.

Today, almost every organization has procedures and rules regarding risk analysis and management and knows how one should carry out risk analysis and management, based on available methods, frameworks and guidance documents. That is, organizations often have formal processes defined for how risk analysis should be conducted. Furthermore, there exist a large number of reported normative research studies that investigate how risk analysis and management can be carried out [24]. However, there are only few studies available that investigate how risk analysis and management are actually carried out in practice. This study is carried out to investigate and understand the current state of the practices of risk analysis and management in large-scale organizations for their IT systems. It is intended to provide insights into how practitioners deal with risk analysis and management in different large-scale organizations and to understand what they see as the main challenges.

2 Background and Related Work

There are several frameworks developed by different national and international organizations, e.g., the ISO 31000 standard [11] for risk management, the ISO/IEC 27005 for Information security risk management, CRAMM by the British Central Communication and Telecommunication Agency (CCTA) [5], OCTAVE by the SEI Software Engineering Institute [1], and the Risk Management guide for information technology systems by the National Institute of Standards and Technology (NIST) [23]. A detailed comparison of some of these frameworks is presented by the European Union Agency for Network and Information Security (ENISA) working group on risk assessment and risk management [6]. Furthermore, for a more detailed risk analyses of technical systems, there also exist a number of risk analysis methods. Some of the most well-known methods of this type are Fault Tree Analysis (FTA) [7], Failure Mode and Effect Analysis (FMEA) [16], and Hazard and operability study (HAZOP) [19]. Some of the more high-level frameworks described above specifically recommend one or more of these risk analysis methods.

There are a number of studies that discuss and present the current practices of risk analysis and management in different domains [8, 15, 18] but not for IT systems. For example, [18] present lessons learned and best practices for risk management for developing an enterprise-wide risk management framework are identified from 14 case studies. [8] describes the current state of risk management practices in German small and medium size enterprises (SME) in a study based on questionnaires and interviews. The research on current practices in risk analysis in IT systems is by no means complete, and more knowledge would be needed. Most of the presented research on available methods and frameworks for risk analysis and management is of normative nature that guide how one should carry out risk analysis and management. It is not sufficiently investigated how different organizations are actually carrying out risk analysis and management.

3 Research methodology

The main objective of this study is to investigate the current practices of risk analysis and management for IT systems in the large-scale organizations. Moreover, this study also investigates the existing challenges in the current practices of risk analysis. The research method used for this study is a qualitative interview research method [3] in the form of a qualitative survey, based on the discussions by [12]. According to [12] the qualitative survey analyses the diversity of member characteristics within a population as opposed to the statistical survey, which analyses frequencies in member characteristics in a population. The research process in this study includes goal definition, design and planning, data collection, data analysis, and reporting, as described by [3]. The first author of this study was responsible for the thematizing, design and planning, collection of data through interviews, transcription, and analysis of the collected data. However, there is one interview that was carried out and transcribed by the second author because it was carried out in Swedish. The methodology is illustrated in figure 1. The circular arrow indicates the iterations that took place in the research. The steps were not carried out in strict order without going back. Instead there were several iterations where new findings affected the detailed goals and the interview questions.

3.1 Goal definition

A set of goals and research questions were defined based on the experience, knowledge of the authors, and literature reviews of the area. The main research questions of the presented research are

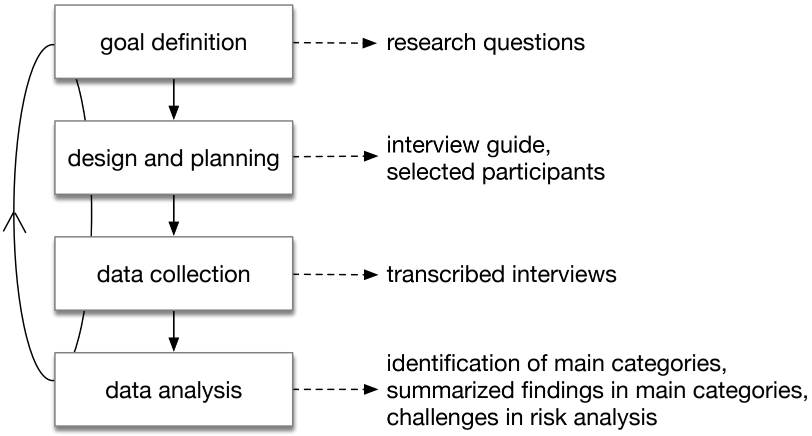


Figure 1: Methodology

- RQ1: What are the main current practices of risk analysis and management for IT systems in the studied governmental organizations?
- RQ2: What are the main challenges in the current practices of risk analysis and management for IT systems in the studied governmental organizations?

RQ1 is answered by interviewing people responsible for risk analysis and management in large-scale organizations about the current practices of IT systems risk analysis. The interviewees were asked how they carry out risk analysis and then the collected qualitative data was analyzed and the results were presented. RQ2 is answered by asking interviewees about the challenges and problems that they face while carrying out risk analysis in their organizations.

3.2 Design and planning

This section presents the research procedures carried out in designing the study and the preparations made for the data collection. This include designing the interview guide with interview questions, and the selection of interviewees. The general design of the study is qualitative. That is, the analysis is based on understanding of the data, and not quantitative findings based on statistical significance, see e.g. [**?**].

The interview guide for the semi-structured interviews was designed based on the research objective and questions of the study. The interview guide was updated several times, e.g., after review by the second author and after carrying out the first interview. In the last update of the interview guide (after the first interview) the general content of the guide remained the same, however the structure and order of the interview questions were modified and improved.

Table 1: Overview of interviewees' roles and their organizations

Org.	Type of org.	Role of interviewee	# empl.
A	Education	Chief security officer	7-8 k
B	Municipality	Coordinator of info. security	21-23 k
C	Health care	Head of IT security	32-34 k
D	Municipality	Responsible of info. security	9-10 k
E	Municipality	Security and safety officer	9-10 k

The interview guide was designed in three themes with questions about current practices of risk analysis, existing challenges in carrying out risk analysis and questions about improving current practices of risk analysis. Before these three main themes there were a number of introductory questions about the interviewees' roles in organizations and their experience with risk analysis, etc. The questions of theme one, current practices of risk analysis, were divided into three groups (general questions about the risk analysis process, questions about carrying out risk analysis and after this questions about risk management). After this in theme two, challenges in current practices, the questions were mainly about identifying the existing challenges in carrying out risk analysis. Finally, in theme three, improvements for current practices of risk analysis, the questions were mainly intended to get suggestions for improvement. However, theme three is not included in this study.

The selection of participants was carried out with the objective to represent the relevant range of viewpoints from security, safety, risk analysis and management domains for IT systems. The following selection criteria were used to select the participants for this study. Participant should be involved with risk analysis and management of IT systems, and must be working or having experience from working in a large-scale organization. Participants should also be located in southern Sweden for ease of accessibility. Initially, 32 persons from 30 different Swedish municipalities, and 3 persons from the regional health care, were contacted through emails. After this, very few responses were received and then all persons who did not reply were contacted again through email after 3 weeks. In total 5 persons responded that they were willing to participate in the study and they all fulfilled the selection criteria, see Table 1.

The interviewees are all responsible for the risk analysis in central units of their organizations. The central unit, which is responsible for the whole organization manages all sub-departments and sub-organizations regarding risk analysis and management e.g., hospitals, schools, emergency services, etc. Furthermore, the persons responsible for risk analysis in the sub-departments and sub-organizations working under a municipality get support from the central risk unit for risk analysis and then after analysis they report analysis results back to the central risk unit for aggregation. This way it can be said that a person responsible for the central risk unit has knowledge about how others are carrying out risk analysis and managing their risks.

3.3 Data Collection and Analysis

A semi-structured interview strategy [21] was, as described above, used for the interviews. The interviews were conducted over a period of about four months. After having consent by the interviewee to record the interview, audio recordings were made of each interview. On average an interview lasted for about 50 minutes. The interviews were transcribed and the transcripts were validated through member checking by sending them to the interviewees for review. After this, the transcripts were divided into chunks of text consisting of a couple of sentences each to map with the themes or categories of the interview questionnaire. Three main themes or categories we defined: current practices, exiting challenges, and suggestions for improvements were defined in the interview guide.

A set of codes, i.e. keywords, based on the research and interview questions was first produced and then iteratively updated during the analysis. The final set of codes consists of 16 codes related to current practices, and 8 codes related to existing challenges. Then, the text relevant to each category was analyzed to find the diversity between the experiences of the interviewees about the risk analysis practices.

3.4 Validity Evaluation

According to [3], the overall validity of an interview-based research study is described by the research reliability, research validity, and the possibility to generalize the research results. The reliability refers to the consistency and trustworthiness of research findings [3]. Here, consistency and trustworthiness denote to what extent a finding is reproducible at other times and by other researchers. This can be interpreted as if another researcher subsequently conducts the same study, the results should be the same. In this study, the interview guide was designed and developed, which helped to carry out interviews in a consistent way. Moreover, all findings were also reviewed by the second author, which means that this threat has been addressed at least to some extent.

The research validity determines if a statement is true, correct, and strong. It helps to solve the issue of whether a method investigates what it claimed to investigate [3]. In this study, it can be assumed that the topic being analyzed (risk analysis and management in large-scale organizations for IT systems) was well known by the all interviewees. Moreover, during the interviews, care was taken not to ask questions with over-complicated terms in order to make sure there was good understanding by the both interviewees and interviewer.

The generalizability is concerned with generalization of results from the chosen population and topics that have been investigated. The main threat to the generalizability of this study has probably to do with the sampling of interviewed subjects. The interviewed subjects can be seen as a representation of risk professionals, not the least since they are collaborating and supporting many risk professionals in their work in the same organizations. The findings are more in the form of understandings than in the form of distributions, which means that the findings are not that difficult to transfer to other situations when the context can be seen as similar. However, the analysis is based on a rather small number of interviews, which needs to be taken into account when conclusions are drawn.

4 Analysis and Results

4.1 Current Practices of Risk Analysis

This section presents the analysis and results of the main categories, that is, the current practices of risk analysis, and the existing challenges that were defined in the interview guide of this study.

Importance of Risk Analysis

During the interviews, it was investigated to what extent risk analysis is seen as important for the interviewees in their work-related activities. According to the majority of the interviewees, it is crucial to carry out risk analysis because most of their security and safety countermeasures and mitigations in different subdepartments and projects are based on it. So, if one finds a severe risk then there is a need to take counter-measures to mitigate the possibility for someone to exploit or miss-use of the system. Furthermore, the interviewees (organizations B and E) mentioned that, in most situations, IT systems that are maintaining other services are not of primary concern in risk analysis. Instead, the services themselves are more critical. For example, if a service is to provide health care, then one needs a system, which consists of staff with expertise, infrastructure, medical facilities etc. However, an important part that is needed is an IT system for the management of different tasks and patient information, which means that these IT systems are more critical because of their importance. Here, the question arises how it is possible to know about the importance and dependency of these services and the answer to this is by carrying out risk analysis. Finally, one of the interviewees (organization D) mentioned that the public organizations are funded by the tax, and therefore need to be really effective. For this effectiveness they should be able to see what parts or services are important. That is, risk analysis is important for the organizations to carry out their tasks and risk analysis helps them manage their critical systems and services.

Identifying Critical Assets and Services

Identification of critical assets and services and their prioritization are important activities in the risk analysis process. Most of the interviewed organizations are aware of this vital activity and in their first step of risk analyses they determine which services are important for society. For example, for delivery of these services one needs several elements, e.g., software, hardware, and staff with knowledge. Therefore, one has to know how important the services are, as they are setting the requirements to the system. For example, organization B has many services that it is providing to the society and for these services they have performed a prioritization for the both normal operations and crisis situations. The IT department at Organization B knows which services are most important for the functioning of the municipality and its inhabitants. If something goes wrong in the infrastructure, then the IT department has a prioritized list of services.

> "The IT department knows which services are most important to serve the municipalities and inhabitants. If something goes wrong in the infrastructure then IT department prioritizes what is more important. Then, they know we have to prioritize for example health care system instead of library service work." (Organization B)

Another interviewee mentioned that:

> "From an IT perspective there is a 'crisis list' that shows what IT systems are important and they are prioritized in a certain order. Moreover, we have developed a critical assets list based on many risk and crisis analyses conducted with in different departments." (Organization C)

However, there are a few organizations (A, D and E) that are aware of this fact but they are trying to improve the overall risk analysis process, including the identification of critical assets. Few organizations are very systematic in doing risk analysis of their IT systems and critical services although they are improving this particular activity and in general the whole risk analysis process.

That is, identification of critical assets is seen as important, but some organizations are still working on defining complete records of important assets.

Defining Roles and Responsibilities

Definition of roles and responsibilities for normal and crisis situations is important. If there is a critical system then it is important to have an owner of it from the risk analysis and management perspectives. Regarding this the interviewed organizations, municipalities and health services providers are at least following a systematic approach and they use some models. For example, one organization B is using a model named PM3[1] for definition of roles and responsibilities.

[1] http://pm3.se/en/

The model demands that all assets need to have an owner or someone responsible.Then, there is an independent audit unit for these organizations. It serves as a bridge between these public organizations and the government to monitor and regulate different activities.

Carrying out Risk Analysis

Regarding the responsibility of risk analysis it was noticed that mostly head of department, project manager or anyone who is responsible for an organization or sub area is responsible for carrying out risk analysis. In the public sector there are normally a few persons responsible for risk analysis and management who carry out risk analysis for central systems. Moreover, they also support other employees, i.e., persons responsible for risk analysis in different sub-departments or organizations, to carry out their risk analysis. The central responsible person or unit for risk analysis requires risk information from all sub-departments or sub-organizations to aggregate risk information to take countermeasures from the central point.

Regarding how these organizations carry out risk analysis, it can be explained in an abstract level because it is not exactly the same in all organizations. The following example presents analysis of acquiring a new information system for health record system. In step 1, they analyze what kind of information is important and what are the potential risks to this information along with legal restrictions, e.g., the personal information act. Based on this they have requirements of the system in hand. This step does not involve technical details at this stage of analysis. In step 2, they have requirements in hand with a list of identified potential risks. In this step they involve technical parts and identify more potential risks by analyzing the requirements to secure the information. Then, they consider potential identified risks with their likelihood and consequences. In carrying out both steps they follow the ISO 27000 standard. In the final step most, but not all, of the investigated organizations re-analyze the acquired system for potential risks after a fixed time interval.

Almost all interviewees carry out risk analysis in their organizations by themselves internally. However, one interviewee mentioned that sometimes they hire a third party to carry out their risk analysis.

Involving People With Different Knowledge

Almost all the investigated organizations involve more than one person in order to get people with different knowledge in the risk analysis process. It is important from the completeness perspective of the risk analysis process i.e., yielding complete risk analysis results. Generally, all the interviewed organizations involve 5 to 7 people in risk analysis consisting of IT function, law function, end-user, system developer representative and one person from higher level management working with the strategies and policies. Here, the IT function means IT security specialist along with system administrator, system architect and system owner. In some cases all representatives from an IT function take part in risk analysis. Then, the law function represents legal team who checks that all legislation is being followed in the risk analysis process. However, there exist some challenges that are presented elsewhere in this paper. For example, one of the interviewees (Organization E) mentioned a risk analysis case where they tried to include people with different knowledge and they ended up with 30 people, which posed several problems in risk analysis.

Used Methods For Risk Analysis

The majority of the interviewees mentioned that generally their main focus is to have a simple method or technique to work with for risk analysis. Almost all of the organizations are using qualitative risk analysis methods with qualitative descriptive values to anticipate probabilities and consequences of unwanted events. They are not using quantitative methods and estimates because they do not want to make the risk analysis process more complicated and difficult.

Almost all organizations are following international standards but they have adapted these standards and methods to their needs. Organization A is using their own risk analysis method that is based on international standards and methods. Organization B is also using their own risk analysis method based on ISO 27000 and ISO 31000, and Organization C is using a risk analysis method that is based on ISO 31000. The risk analysis method used by Organization D and E is based on ISO 27000, adapted by a national organization[2].

Follow-Up Analysis

Follow-up risk analysis means that if a critical risk is found and is being treated then it requires to carry out risk analysis again for whole system or a specific part of system depending on the treatment of found risk. Most of the investigated organizations carry out follow-up risk analysis. For example, Organization B carries out follow-up after every two years and Organization C and Organization D carry it out every time after introducing significant changes in the system. However, Organization C is dealing with more safety critical systems as it is providing health services to a large number of inhabitants. Moreover, because of its severity it does not have any defined time period for follow-up analysis instead it has instructions that follow-up analysis should be carried out after every major change.

Education and training About the Risk Analysis Process

Not very many concrete activities regarding education and training about risk analysis in the investigated organizations were described. It was found that these organizations are not investing enough to educate and train their employees to carry

[2] Swedish Association of Local Authorities and Regions, https://www.skl.se

out or participate in the risk analysis process. For example, persons responsible for risk analysis in these organizations has duty to educate his/her sub-departments or organizations to carry out risk analysis by their own. The majority of the interviewees (organizations A, B and E) mentioned that they educate others about the risk analysis methods and models. However, except this education there is no other courses or trainings for risk analysis in these organizations. Only one interviewee (Organization C) mentioned that they have both internal and external training courses about risk analysis.

4.2 Existing Challenges in Carrying Out Risk Analysis

This section presents the identified challenges, which are encountered by interviewees while carrying out risk analysis of their IT systems.

Required Competence and Skills

The required competence and skills are considered to be significant to carry out risk analysis. In general, all of the interviewees mentioned that the required competence and skills are important to carry out risk analysis. The required competence and skills are not the same in all sub-departments or organizations. For example, one of the interviewees mentioned that:

> “The competence and skills change dramatically between different sub-departments and sub-organizations. This variation in competence and skills of different sub-departments makes the risk analysis process more difficult and complex.” (Organization B)

The important challenge is knowledge of doing risk analysis because most of the times there is a need of someone who leads the risk analysis process with the understanding of what should be covered in the risk analysis process. Therefore, lack of people with required competences and skills could be a risk in itself for risk analysis because it can yield incomplete analysis results.

Different Opinions About Risks and The Risk Analysis Process

The next unfolded challenge by this study is that people involved in risk analysis have different opinions about risks and the risk analysis process itself. Here, the difference in opinions regarding risks and the risk analysis process exist because of how people perceive and trust them. Trust in the risk analysis and management process is very crucial and helps in yielding better results, which consequently improves critical societal services. The view of distrust comes from considering something not important. [22] argues that this distrust among people is not because of ignorance. On the other hand, after these many years, today we can notice that it is somehow because of public ignorance. Therefore, there exist challenges that risk practitioners are facing because people think differently and sometimes do not have enough information. All the interviewees mentioned that more or less they encounter this challenge, e.g.,

> "Mostly people involved with risk analysis are divided into three categories i.e. people who say 'yes', people who say 'why?', and people who say nothing about carrying out risk analysis." (Organization A)

People who say 'yes' represents the group that knows the importance of risk analysis and this group trusts in it. People who say 'why?' represent the group that knows the importance of risk analysis but only partially. Therefore, after carrying out risk analysis once they are reluctant to carry out a follow-up analysis. Then, the third group who says nothing seems ignorant and this group does not believe and trust in the risk analysis and management process. Usually the third group tries to move to next step without carrying out risk analysis. One of the interviewees mentioned that it is really important to communicate that risk analysis is an important activity for a project or mission and it can not be skipped at any cost.

After this, comes the issue of different priorities in a department, which is also relevant to the different thinking about the risk analysis process. For example, one of the interviewees mentioned that everything has a priority and it might be that carrying out risk analysis for critical IT systems and services in a particular department is not given the highest priority. Sometimes, when the central risk unit question about it then the department's risk people answer yes we know that risk analysis is important but we do not have time to carry out. Here, it seems like these kinds of problems are more exposed to the large-scale organizations however it is interesting to investigate this in small-scale organizations as well.

Pre-Understanding of The Risk Analysis Process and Its Context

To carry out an effective and complete risk analysis, people involved in it should have good pre-understanding of the potential risks, the risk analysis process, and the system being analyzed within its context. Almost all of the interviewees mentioned that having a clear pre-understanding about these elements is very important, although they do not always see this in their practices. One of the interviewees mentioned that if an analysis unit does not have this pre-understanding then there are chances that they will also identify some risks that are ongoing or already been eliminated.

Subjectivity in Risk Analysis

There is a discussion going on whether risk analysis is subjective, objective or some combination of both, e.g. [4, 9]. Since the risk analysis process involves subjective judgments or estimates in its all activities, it does not produce exact estimates that means one cannot call it a complete objective process. Here, it worth to mention that the estimates for potential risks made by either experts or a normal user of an analyzed system can not eliminate involved subjectivity in the risk analysis process [20].

In this study, all the interviewees mentioned issues related to subjectivity in the risk analysis process. Some of the interviewees mentioned that this subjectivity is a severe challenge that makes risk analysis more complex and difficult to carry out. People think differently and have different opinions about things similarly this goes for potential risks and the risk analysis process as well. Interestingly, one of the interviewees mentioned that this subjectivity, of course, brings some challenges, but it is also a significant strength of the risk analysis process. All the interviewees also mentioned that because of subjectivity they carry out risk analysis in groups and try to reach consensus regarding risks and their values. Furthermore, they mentioned by involving different kinds of roles in risk analysis could eliminate this challenge at some extent.

Follow-Up Risk Analysis

All investigated organizations have well defined and designed normative documents saying that they should carry out follow-up analysis after some specified time interval. However, because of priorities in the organizations, they are not always following these instructions strictly. One of the interviewees mentioned that:

> "We are reworking with the instructions for follow-up analysis that it should be carried out more on the regular basis. We have work in progress there to improve this area." (Organization C)

Interestingly, Organization D has this follow-up analysis as a major activity in their risk analysis and management process, and it has also defined the person responsible for it. However, here the challenge is if one has performed risk analysis a year or more ago then he/she question why this should be done again.

Moreover, the interviewee from Organization A, mentioned that he carried out more than 20 risk analyses of different departments. He tried to educate all employees about risk analysis and also motivated them to carry out risk analysis on regular basis to improve decision-making. He offered to all departments of the organization that whenever they need help or suggestions to eliminate risks or regarding anything for risk analysis and follow-up they can contact him. After this, no one contacted him regarding anything just because they think risk analysis is over and now they can do other things by saying that we are done with risk analysis.

5 Discussions

Several factors were investigated regarding the current practices and existing challenges of risk analysis as presented in detail in the results section. For example, regarding the importance of risk analysis in large-scale organizations the findings show that it is an important activity for organizations that are dealing with safety-critical systems and services corroborating the findings of [18]. It is an important activity because most of an organization's security and safety countermeasures or mitigations in different sub-departments and projects are based on risk analysis and assessment. However, the associated challenge with risk analysis is that different employees have different opinions about risks and the risk analysis process. This difference in opinions is probably because of different level of perception, trust, and priorities about the risk analysis process. Therefore, this difference in opinions about importance of risk analysis makes difficult to carry out risk analysis.

Then, regarding carrying out risk analysis practices (RQ1) more or less all investigated organizations are analyzing their critical systems or services in the same way, at least on an abstract level. The reason for this could be that all organizations are of the same nature, i.e., governmental organizations dealing with critical services to society. It was found that these organizations mainly focus on the information or services in risk analysis instead on an IT system. Here, it can be said that they are using the system level analysis method [17]. The associated challenges with carrying out risk analysis practice are the following. Firstly, the required competences and skills are a great challenge in carrying out risk analysis in these critical organizations. The findings of this study suggest that lack of knowledge and expertise about doing risk analysis is itself a risk. Moreover, the knowledge about the system context that is being analyzed and its boundary definitions are very crucial as discussed in [14]. Secondly, pre-understanding of the risk analysis process is also a challenge while performing risk analysis of safety-critical services. This challenge is very similar to the required skills and competences challenge. However, it is about having good pre-understanding of potential risks, the risk analysis process, and the system being analyzed with its context. On the other hand, required skills and expertise deal with the knowledge of different risk analysis methods or tools and then the knowledge used for defining the system boundaries. The best practices of risk analysis and management identified in study [18] also suggest that the risk management process should start with context establishment that includes organizational objectives, stakeholders, constraints, risk criteria, and other factors.

Next, involvement of different people (RQ1) in risk analysis is investigated and it is found that these organizations involve more than one person with different knowledge in the risk analysis process. Several authors, e.g., [13], [10], and [25] advocate to involve various roles or perspectives in risk analysis. Regarding involvement of different people in risk analysis the investigated organizations are seem to be mature. The associated challenge (RQ2) with this practice is subjectivity involved in risk analysis.

Furthermore, the analysis of this study reveals that the simpler methods, models and tools for risk analysis are better and being used in the investigated organizations as this corroborates with the findings of the study [18].

There is, of course, an evolution of risk analysis methods for IT systems. There are methods available today, and there will be new methods developed. Based on this study it is possible to point at some of the challenges which could be possible to take into account in the development of new methods. Especially the fact that risk analysis on the one hand requires involvement of many experts in the organization but on the other hand that it is challenge to overcome the need for training about risk analysis and that it is not always given the highest priority in all parts of the organization as described above. This means that risk analysis approaches must overcome these challenges.

6 Conclusions

This study investigates and presents the current practices (RQ1) and existing challenges (RQ2) of the risk analysis and management process for IT systems in different large-scale organizations. Based on the results of this study regarding RQ1, it can be concluded that risk analysis is an important activity in large-scale organizations because most of their security and safety countermeasures are dependent on it. Furthermore, based on the findings of this study it can be concluded that the investigated organizations are not sophisticated regarding the identification of critical assets and services and its documentation, which is an important activity for risk analysis. Regarding methods for risk analysis it can be concluded that the simpler methods, and models for risk analysis are better than the complex ones because they are being used in the investigated organizations.

Based on the results of this study, regarding RQ2, it can be concluded that the required competences and skills are a great challenge in carrying out risk analysis in large-scale critical organizations. Another identified challenge is lack of knowledge and expertise about doing risk analysis, which is itself a risk. Next, the pre-understanding of potential risks, domain, and the risk analysis process is also a challenge while performing risk analysis of safety critical services. The findings of this study also found that involvement of different roles in the risk analysis process eliminates at least the negative effects of subjectivity in risk analysis. After this, the next challenge identified is relevant to follow-up analysis that is how people perceive risk analysis and its priority among other things, which is different in different organizations and their sub-departments. There is a need for further research in order to provide general suggestions to improve the risk analysis process in large-scale organizations dealing with safety-critical IT systems and services. New risk analysis methods that are defined should meet the identified challenges.

Bibliography

[1] C. Alberts and A. Dorofee, *Managing Information Security Risks: The Octave Approach*, ser. SEI Series in Software Engineering. Addison-Wesley, 2003.

[2] H.-P. Berg, "Risk management: procedures, methods and experiences," *Reliability: Theory & Applications*, vol. 1, no. 2, pp. 79–95, 2010.

[3] S. Brinkmann and S. Kvale, *InterViews: Learning the Craft of Qualitative Research Interviewing*. SAGE Publications, 2014.

[4] S. Campbell, "Risk and the Subjectivity of Preference," *International Journal of Risk Research*, vol. 9, no. 3, pp. 225–242, 2006.

[5] Central Computer and Telecommunications Agency, Great Britain, Treasury, *Prince User's Guide to CRAMM*, ser. Programme and Project Management Library, 1993.

[6] ENISA Ad Hoc Working Group on Risk Assessment and Risk Management, "Inventory of risk assessment and risk management methods," European Union Agency for Network and Information Security, 2006.

[7] C. A. Ericson, "Fault Tree Analysis - A History," in *proceedings of The 17:th International System Safety Conference*, 1999.

[8] T. Henschel, "Typology of risk management practices: an empirical investigation into German SMEs," *International Journal of Entrepreneurship and Small Business*, vol. 9, no. 3, pp. 264–294, 2010.

[9] N. Hurst, *Risk Assessment: The Human Dimension*. The Royal Society of Chemistry, 1998.

[10] S. Ierace, "The basics of FMEA, by robin e. mcdermott, raymond j. mikulak and michael r. beauregard," *Journal of Production Planning and Control*, vol. 21, no. 1, pp. 99–99, 2010.

[11] ISO 31000:2009, "Risk management - principles and guidelines," 2009.

[12] H. Jansen, "The logic of qualitative survey research and its position in the field of social research methods," *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, vol. 11, no. 2, p. 21, 2010.

[13] N. G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*. The MIT Press, 2012.

[14] C. Lindholm, J. P. Notander, and M. Höst, "A case study on software risk analysis and planning in medical device development," *Software Quality Journal*, vol. 22, no. 3, pp. 469–497, 2014.

[15] M. W. Meyer and K. A. Solomon, "Risk-management practices in local communities," *Policy Sciences*, vol. 16, pp. 245–265, 1984.

[16] R. J. Mikulak, R. McDermott, and M. Beauregard, *The Basics of FMEA, 2nd Edition*. Taylor & Francis, 2008.

[17] G. Motta, G. Pignatelli, T. Barroero, and A. Longo, "Service Level Analysis method - SLAM," in *3rd International Conference on Computer Science and Information Technology*, vol. 5, July 2010, pp. 460–466.

[18] C. A. Murdock, M. Squeri, C. Jones, and B. S. Smith, "Risk Management in Non-DoD U.S. Government Agencies and the International Community: best practices and lessons learned," Center for Strategic and International Studies (CSIS), Tech. Rep., 2011.

[19] F. Redmill, M. Chudleigh, and J. Catmur, *System Safety : HAZOP and Software HAZOP*. John Wiley & Sons, 1999.

[20] F. Redmill, "Risk analysis - a subjective process," *Engineering Management Journal*, vol. 12, pp. 91–96(5), April 2002.

[21] C. Robson, *Real world research*, 2nd ed. Blackwell, 2002.

[22] P. Slovic, "Perceived risk, trust, and democracy," *Risk Analysis*, vol. 13, no. 6, pp. 675–682, 1993.

[23] G. Stoneburner, A. Goguen, and A. Feringa, *Risk Management Guide for Information Technology Systems*, ser. National Institute of Standards and Technology, Special Publication 800-30. U.S. Government Printing Office, 2002.

[24] S. M. Sulaman, K. Weyns, and M. Höst, "A review of research on risk analysis methods for it systems," in *proceedings of the 17:th International Conference on Evaluation and Assessment in Software Engineering (EASE '13)*. ACM, 2013, pp. 86–96.

[25] S. M. Sulaman, K. Wnuk, and M. Höst, "Perspective Based Risk Analysis - a Controlled Experiment," in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, ser. EASE '14, London, UK. ACM, 2014, pp. 47:1–47:10.

Paper III

Perspective Based Risk Analysis - A Controlled Experiment

Abstract

Context: The increasing dependence on critical IT systems makes them more and more complex, which results in increased complexity and size. Risk analysis is an important activity for the development and operation of critical IT systems, but the increased complexity and size put additional requirements on the effectiveness of risk analysis methods. There complexity means that there is a need to involve different perspectives into risk analysis. **Objective:** The objective of the research carried out in this study is to investigate the effectiveness of perspective-based risk analysis (PBRA) methods compared to traditional risk analysis (TRA) methods. **Method:** A controlled experiment was designed and carried out. 43 subjects performed risk analysis of a software-controlled train door system using either TRA or PBRA. **Results:** The results suggest that PBRA helps to identify more relevant risks than TRA. On the other hand, our experiment failed to provide supporting evidence that PBRA helps to identify fewer non-relevant risks. This study also found that PBRA is more difficult to use than TRA. **Conclusions:** Some potential benefits of using perspective-based risk analysis are uncovered and experimentally confirmed. In particular, it was discovered that PBRA is more effective than the traditional method and identifies more relevant risks.

Sardar Muhammad Sulaman, Krzysztof Wnuk and Martin Höst,
In Proceedings of the 18:th International Conference on Evaluation and Assessment in Software Engineering (EASE '14). Association for computing machinery (ACM) 2014.

1 Introduction

The increasing complexity of socio-technical IT systems and our dependence on them put additional pressure on the effectiveness of risk analysis methods. More complex IT systems contain more interacting components and sub-systems, which in turn increase the probability of serious failures [13]. Moreover, failures in these complex safety-critical systems are often results of multiple interacting decisions and errors [11].

The complexity, size, and heterogeneity of today's IT systems call for involving different perspectives into risk and hazard analyses. Several authors, e.g., Leveson [11] and Ierace [5] recognized the benefits from multiple views analysis and encouraged adding internal and external organizational perspectives into the hazard analysis teams. Yoran and Hoffman proposed defining roles and identifying actors before performing risk analysis in order to improve the process [24]. Morevoer, involving different perspectives is also recommended by several risk analysis standards and methods [1, 3, 7, 18].

Perspective-based reading was successfully used for reviews and inspections during software projects, e.g. [14]. However, the potential benefits of involving perspectives into risks analysis have, to our knowledge, not been explored in an experimental way. It can also be observed that only one study listed in a survey about controlled experiments in software engineering was classified as software and system safety [16], which also indicates the need for experimentation in the area.

In this paper, we report the results from an experiment designed to investigate if Perspective-Based Risk Analysis (PBRA) that involves different views and perspectives is more effective and offers higher confidence than Traditional Risks Analysis (TRA). 43 subjects performed risks analysis of a software-controlled train door system using either TRA or PBRA. The effectiveness of the methods is measured by counting the number of relevant and non-relevant risks. A questionnaire was used to assess the difficulty of the methods and the confidence of the subjects concerning the correctness of the identified risks.

This paper is structured as follows: Section 2 provides related work while Section 4 outlines the experimental design. Section 4 describes the execution of the experiment and Section 5 provides the experimental results. Section 6 analyzes the results and Section 7 discusses the validity threats. Section 8 presents the discussion. Finally, the paper is concluded in Section 6.

2 Related Work

There exist a number of risk analysis methods for technical systems in general or for IT-systems in particular, e.g. [1, 3, 7, 18] just to name a few. The Risk Management guidelines for Information Technology Systems [18] highlight that management, CIOs, system owners, business managers and security program managers should be involved into the risk management process. The OCTAVE method for risk-based information security assessment also advocates involving business and IT perspectives into the risk analysis processes [1]. The NetRAM method for network security analysis is also adapted for different enterprise structures on different levels and therefore can also involve the business perspective [3].

Some of the most well-known low-level risk analysis methods are Fault Tree Analysis (FTA) [4], Failure Mode and Effect Analysis (FMEA) [13] and Hazard and operability study (HAZOP) [5]. These methods have successfully been used for decades for technical and IT systems. However, these traditional methods do not consider the use of perspectives.

The recent advances in risk analysis methods or techniques include an actuator-based approach that identifies failures in four different severities [9] and the System Theoretic Process Analysis (STPA) method proposed by Leveson that considers safety as a control problem rather than a component failure problem [11]. STPA was applied to various systems with positive outcomes [6, 11, 21].

The idea of using perspectives is not new. Perspectives were utilized for reading software engineering artifacts with the purpose of improved defect identification [2, 14]. Perspective-based reading was also applied for object oriented design inspections [15], code reviews [10] and usability inspections [25]. Different perspectives, e.g. developers, testers and domain experts are often involved in requirements elicitation. This results in increased quality of elicited requirements and often uncovers new requirements based on various views and perspectives.

Yoran and Hoffman proposed the Role-Based Risk Analysis (RBRA) method that defines roles and identifies actors before performing risks analysis activities in order to reduce the set of vulnerabilities and controls to those appropriate to a given role [24]. RBRA was presented on an illustrative example from the computer software engineering domain but not experimentally investigated. Leveson [11] and Ierace [5] advocated to involve various perspectives during risk analysis, also from external organizations. It is always recommended, in almost all risk analysis methods, to have experts with domain knowledge while performing risk analysis but to our knowledge no one has proposed the use of specific perspectives for risk analysis. In this study we have used specific perspectives for the performed risk analysis. To summarize, the potential of perspectives in risk analysis was not yet experimentally assessed.

3 Experimental Design

In this study, the research is carried out through a controlled experiment based on the guidelines presented by Wohlin et al. [22] and reported based on the reporting

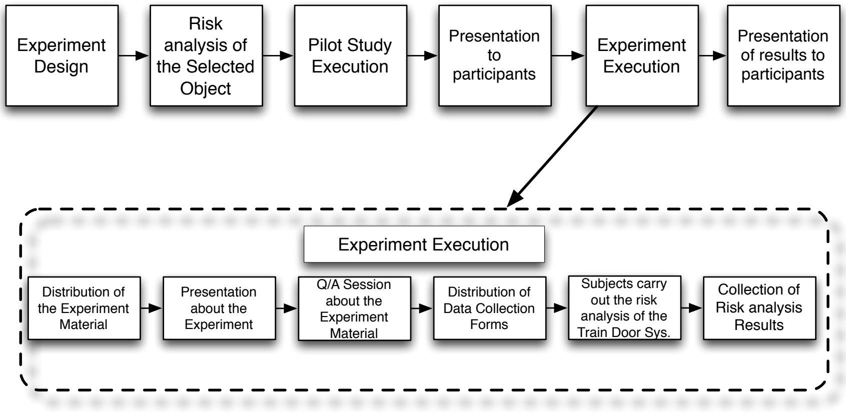

Experiment Design
Risk analysis of the Selected Object
Pilot Study Execution
Presentation to participants
Experiment Execution
Presentation of results to participants
Experiment Execution
Distribution of the Experiment Material
Presentation about the Experiment
Q/A Session about the Experiment Material
Distribution of Data Collection Forms
Subjects carry out the risk analysis of the Train Door Sys.
Collection of Risk analysis Results

Figure 1: Carried out steps for the experiment

guidelines presented by Jedlitschka et al. [8][1]. This research is carried out in the following steps as shown in Figure 1.

1. Experiment design
2. Risk analysis of the selected object for this experiment. This resulted in a first set of “correct risks”.
3. Pilot study
4. Presentation about risk analysis to the subjects at a lecture
5. Experiment execution
6. Presentation of results to the subjects

3.1 Research Questions

The objective of the research carried out in this study is to investigate the effectiveness of the PBRA method in comparison with the TRA method. Here, effectiveness means a large number of relevant risks and a small number of non-relevant risks. This general objective is broken down to the following research questions:

- RQ1: Which risk analysis method is more effective?
- RQ2: Which risk analysis method is more difficult to use?

[1] The experimental package including all the guidelines and results is available at `http://serg.cs.lth.se/index.php?id=87041`

- RQ3: How confident are the participants about the risks they find using the studied methods?

RQ1 is important to investigate since a general goal of any risk analysis method is to find as complete set of risks as possible [19] and to minimize the number of non-relevant risks. RQ2 is relevant to investigate since the successful introduction of any method is dependent on that it is not seen as too hard to use by the users. Moreover, if the users do not feel confident (RQ3) with the results of the proposed method, they will be reluctant to apply the method in the real safety-critical systems.

3.2 Variables and Hypothesis

The following independent and dependent variables are used in this experiment. The independent variable is the used risk analysis (RA) method. Two methods are compared in this experiment:

- Traditional risk analysis
- Perspective-based risk analysis

The dependent variables for this experiment are:

- N_r: Number of relevant risks found
- N_{nr}: Number of non-relevant risks found
- D: Difficulty level while using risk analysis method. The difficulty is measured on a Likert scale with five possible values, from *very easy* (1) to *very difficult* (5).
- C: Confidence level of the participants about found risks. The confidence level is measured on a Likert scale with five possible values, from *Very Confident* (1) to *Strongly not confident* (5).

The values of the dependent variables, N_r and N_{nr}, are calculated based on the identified relevant and non-relevant risks. The confidence and difficulty levels are determined using a questionnaire. The statistical analysis was performed to accept or reject the hypotheses H_0^1, H_1^1 and H_2^1.

RQ1 is broken down into two null hypotheses, detailed below. The first null hypothesis is that both risk analysis methods, PBRA and TRA, find the same numbers of relevant risks.

- H_0^1: The mean of PBRA and TRA is equal that both found same number of relevant risks (N_r).

The alternative hypothesis is:

- H_1^1: The mean of PBRA and TRA is not equal that both found different number of relevant risks (N_r).

The second null hypothesis for the RQ1 is that both risk analysis methods, PBRA and TRA, find the same numbers of non-relevant risks.

- H_0^2: The mean of PBRA and TRA is equal that both found same number of non-relevant risks (N_{nr}).

The alternative hypothesis is:

- H_1^2: The mean of PBRA and TRA is not equal that both found different number of non-relevant risks (N_{nr}).

The null hypothesis for the RQ2 is that both risk analysis methods, PBRA and TRA, are equally difficult to use.

- H_0^3: Both PBRA and TRA methods have same median that is same difficulty level to use (D).

The alternative hypothesis is:

- H_1^3: TRA method has lower median that it is less difficult to use (D).

The null hypothesis for the RQ3 is that the participants of both methods are equally confident about the identified.

- H_0^4: The median for both methods, PBRA and TRA, is same. i.e. the participants of both treatments are equally confident (C).

The alternative hypothesis is:

- H_1^4: TRA method has small value of median that means the participants that used TRA are less confident (C).

3.3 Subjects

The sample included participants of a project course in software development at Lund University, offered in autumn 2013[2]. The course is an optional advanced-level Masters' course for students from several engineering programs, e.g., Computer Science, Electrical Engineering, Civil Engineering, and Information and Communication Technology. The course gives 7.5 ETCS points that corresponds to five weeks full-time study. This experiment was a non-mandatory part of the course. 43 out of the total 70 students took part in the experiment. The participants were instructed clearly that the results of this experiment were completely anonymous and do not have any effect on the final grade of the course. It was also explained that results of the experiment will be used for research, and if they do not want to participate in the research then they are not required to submit their results.

[2] http://cs.lth.se/kurs/etsn05-programvaruutveckling-foer-stora-system/

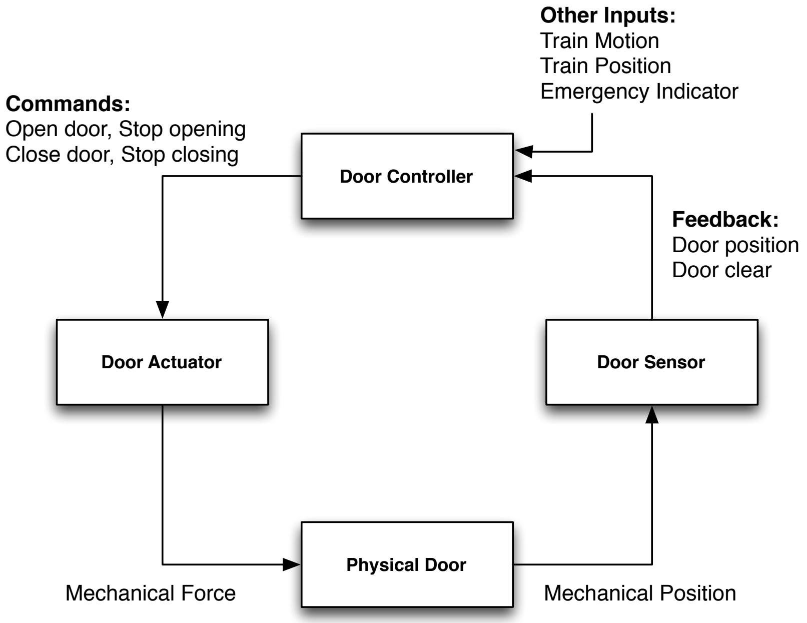


Figure 2: Functional diagram of a Train Door System [21]

3.4 Objects

The objects used a software-controlled insulin pump as an example in the guidelines, and a software-controlled train door system during the experiment. Both systems represent embedded socio-technical safety-critical systems.

Train Door System

The train door system (see Figure 2) was selected for the experiment because of the following reasons: (1) it is a simple system and it has fewer components than the insulin pump, (2) it is highly possible that almost every participant has used this kind of system, (3) the system is rather simple and should be easy and quick to understand and (4) the participants should be able to find many risks for this system. The automated train door system has four main components, shown in Figure 2, the door sensor, door controller, door actuator and the physical door.

The door sensor sends a signal about the door position and the status of the doorway (if the doorways is clear or not) to the door controller. Then, the door controller receives input from the door sensor with some other inputs from the external sensors about the motion and the position of the train. It also gets an

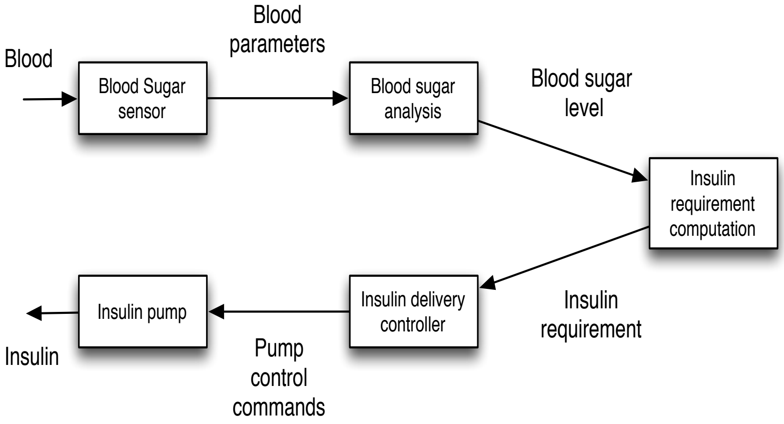


Figure 3: Functional diagram of an Insulin Pump [17]

indication about possible emergencies from an external sensor. After receiving inputs, the controller performs some computation and then it issues door open and close commands as shown in Figure 2. After this, the door actuator receives commands from the controller and it applies mechanical force on the physical door. Finally, there is a physical door in the system that is closed and opened by the door actuator.

Insulin Pump

The software-controlled insulin delivery system (see Figure 3) provides automated insulin delivery by monitoring blood sugar levels. The insulin pump is a portable device that delivers insulin via a needle attached to the body. It was selected to be an example system in the experiment guidelines because it is a representative example of a small and simple safety-critical system. Moreover, it has already been used for risk analysis [17].

3.5 Treatments

Traditional Risk Analysis (TRA)

The TRA method is an iterative activity and it consists of the four following steps [20]:

1. **Planning**: In this step, after forming groups all group members carefully read the system description individually and then decide who will be the moderator and who will be the scribe for the group.

2. **Risk identification**: This step determines a list of possible risks. It is an iterative activity that is normally carried out by brainstorming. In this step every risk analyst in the group attempts to find an individual list of possible risks by answering the following question: What could happen or what can go wrong?

   After performing individual analysis, all analysts in the group compare and merge their individually identified risks with others and make a common risk list. During this process new risks can also be identified.

3. **Determine likelihood**: Step 3 determines the likelihood of occurrence of all identified risks from step 2 by using the qualitative descriptors, i.e., highly unlikely, unlikely, possible, likely, very likely.

4. **Determine consequence**: Step 4 determines the consequences (severity level) of all identified risks from step 2 by using the qualitative descriptors, i.e., insignificant, minor, moderate, major, catastrophic.

Perspective-Based Risk Analysis

The PBRA method is also an iterative activity that supports the risk analysts to view and analyze the system from different perspectives. For example, one analyst may analyze the system from the point of view of the designer, another from the point of view of the developer, and another from the point of view of the user/client of the system. We believe that by using different perspectives risk analysts can perform a better and more in-depth analysis by thinking about different safety and security requirements. The used guidelines for the both treatments were the same; there was no extra information for the PBRA participants except the used perspectives.

PBRA consists of four steps just like TRA, but in step 1, the planning step, every member of the risk analysis team (group) is assigned one perspective for the identification step (this is similar to the approach suggested by Yoran and Hoffman [24]). The other steps of PBRA are the same as in TRA.

The selection of perspectives can be done by the participants in the groups, or can be assigned to the group before they start, in this case by the experimenter. In this experiment, during the pilot study the experimenter assigned the specific perspectives to the participants according to their experience. In the experiment execution with subjects, the perspectives were selected by the participants themselves according to their own choice.

In this experiment, PBRA was performed from the following three perspectives for the train door system.

- System Engineer (SE)
- Tester (T)

- Train Staff Member (TS)

The participants were informed that it is possible and even likely that several of the identified risks from the different perspectives are the same.

3.6 Instrumentation

The detailed guidelines were written in an understandable language and reviewed by the authors to execute the experiment effectively. Minor changes were introduced in the guidelines for the PBRA method about the use of different perspectives, in PBRA the participants have to use different perspectives unlike TRA[3].

The first section of the guidelines is about motivation to perform the experiment and the risk analysis. The main motivation for the subjects to participate in the experiment was to use the gained knowledge and experience from the experiment in their own course projects since risks analysis was a mandatory part of the course.

Then, the guidelines present the risk analysis method in detail with step-by-step instructions to perform it. The guidelines also present one example system (insulin pump) with some of the identified risks to give a solid idea about the risk analysis process to the participants. The guidelines also present qualitative descriptors for the likelihood of occurrence and the consequence levels with their definitions. The example presented in the guidelines shows all steps of risk analysis for the example system (insulin pump) with likelihood and consequences and example risks.

The description of the system (train door system), selected for the experiment, was appended in the appendix of the guidelines. The system description contains the technical details of the system and shows the boundaries of the system and the system context. For risk analysis, defining the boundaries (scoping) of the system being analyzed including all dependencies between components is very important otherwise risk analysts could easily become confused or could find many non-relevant risks. The system context, i.e., where the system is used, how and by whom, is also very crucial for the risk analysis. To perform an effective and efficient risk analysis the risk analysts should have clear understanding of the system context [12].

A post-experiment questionnaire was designed to measure the understanding of the guidelines, system description, and prior experience of risk analysis process. It contains 8 questions in total, where 6 of them are quantitative and 2 are qualitative.[4]

Two different data collection forms were designed to be used by the participants, one for each risk analysis method. The participants were asked to write

[3] The guidelines can be accessed at `http://serg.cs.lth.se/index.php?id=87041`

[4] The questionnaire can be accessed at `http://serg.cs.lth.se/fileadmin/serg/Questionnaire.pdf`

identified risks on the provided data collection forms. The example presented in the guidelines used the same data collection forms. The motivation behind this was to give good understanding of the risk analysis process to the participants.

For data collection a complete set of risks was needed to decide which risks, identified by the participants, are relevant and non-relevant. The set of risks was incrementally developed in several phases. The first author performed the initial analysis and identified 28 risks. Then, one independent researcher working in the software safety domain evaluated the list. After evaluation and discussion 13 more risks were added. During the pilot study, 23 additional new risks were identified and added to the list. As a result, the final risk list contains 64 risks. The participants of the experiment found 5 new risks during the experiment execution that were not identified by the experimenters before. After adding these 5 new risks in the risk list the total identified risks became 69.

3.7 Pilot Study/Experiment

After preparing the instrumentation a pilot experiment was carried out on 13:th September 2013. The pilot study was carried out to evaluate the instrumentation of the experiment. Therefore, the results of pilot study are not used in the analysis of the experiment.

The sample contained 9 participants, where 5 participants were from the IT industry with 1.5–5 years of experience in software testing and development. The four other participants were researchers; one was PhD in biology and the other three were PhD students in computer science and electrical engineering.

Since there were 9 participants, they formed three groups each with three members. Each group was in the separate room when they performed the risk analysis for the pilot experiment. Two groups performed risk analysis by using PBRA and one group by using the TRA method.

The pilot experiment was carried out by following same steps for the main experiment mentioned in Section 4. The participants of the pilot study were asked to give feedback verbally after the experiment. After this, the feedback was noted down by the experimenter for the later analysis of instrumentation.

The participants of the pilot study mentioned the following problems or ambiguities in the guidelines, and system description.

- The example system and the experiment system are not clearly distinguished in the guidelines.

- Some information was missing in the given presentation for the experiment, e.g., example about one risk having multiple causes and other way around.

- In the functional diagram of the train door system there is one ambiguous input (*control command*) and one ambiguous output (*status*).

Table 1: Results from the Pilot Study

Group label	Applied treatment	# of relevant risks found	# of non-relevant risks found
G1	PBRA	26	1
G2	PBRA	14	0
G3	TRA	11	0

- *Other inputs* mentioned in the functional diagram of the train door system are un-clear.
- The detail of mentioned *emergency indicator* in the functional diagram of the train door system is missing.

Based on the identified problems and suggestions from the participants in the pilot study, changes were made to the instrumentation for the main experiment. The guidelines were improved by explaining the differences between the both example (insulin) and experiment (train door) systems. For the missing information in the presentation, it was decided that the example risks should be explained in the main experiment presentation clearly. The functional diagram was improved by removing the mentioned ambiguous input and output (*control command* and *status*). Here, the problem was unclear system boundaries because the mentioned ambiguous input and output were connected with some external systems. There were three *other inputs* (train motion, position and emergency indicator) to the train door systems that were not clearly explained in the system description. This problem was fixed by adding explanation for each input with headings (clearly visible).

Results From the Pilot Study

Table 1 shows the results of the pilot study. There were three groups in the pilot study. Two groups (G1 and G2) used PBRA and one group (G3) used TRA. It can be noticed that the number of identified risks of G1 are significantly higher than for the other two groups. This may be because of differences in experience of participants. Group G1 had one member with 5 years of experience working as a system tester and a second member was a PhD in biology. The experience of the participants in G2 and G3 was almost same (1.5-2 years) and there is not a significant difference in the number of found risks between them. However, more risks were found with PBRA than TRA.

3.8 Data Collection Procedure

The data collection procedure was kept same for both the pilot experiment and main experiment. The subjects were given a presentation including the motivation for the experiment, explanation of the risk analysis method to be used (TRA or PBRA) with an example. The system that they should work with (train door) was also described.

Then, there was a short answer/question session about the guidelines, system description etc. Then, each group was asked to perform risk analysis. All members of each group performed an individual risk analysis as mentioned in the instrumentation. After this, each group was asked to compare and merge the individual risk lists to come up with a common group risk list.

Data collection forms, designed by the experimenter, were distributed among the participants for writing the identified risks during the risk analysis. After completion of the risk analysis, data collection forms were collected group by group. Then, all the participants were given a post-experiment questionnaire. The results from the experiment were collected by analyzing the information written in the data collection forms and then these results were also checked against the post-experiment questionnaire.

Each group was assigned a label and asked to write that on the data collection forms. Group labels were used to know that both data collection forms and post-experiment questionnaires are from one specific group, which was required for the analysis. The participants of the experiment were completely anonymous.

4 Experiment Execution

There were total 43 participants of the experiment, see Section 3.3. These participants were divided into 14 groups (7 groups for each treatment) with 3 members in each as shown in Table 2.

Three course seminars were assigned for this experiment. The first seminar was on 9:th September 2013, at 15-17, the second on 10:the September at 8-10, and the third was also on 10:th September at 10-12. It was decided to perform the experiment with the only treatment PBRA in the first seminar, and with the treatment TRA in the second seminar. The third seminar was allocated to balance the number of groups for both treatments.

21 students attended the first seminar, forming 7 groups, and they all participated in the experiment with the PBRA treatment. 4 students attended the second seminar and used the TRA treatment by forming 1 group of three members. The remaining one student was not part of the experiment. In the third seminar, 18 subjects participated. All attendees of the third seminar used the TRA method and formed 6 groups, which balanced the experiment so that there were equally many groups for both treatments.

This experiment was carried out by following steps.

1. The experiment guidelines were distributed among the participants.

Table 2: Summary of groups at seminars

Seminar	# of Subjects		# of Groups
	PBRA	TRA	
I	21	-	7
II	-	$4-1=3$	1
III	-	18	6
Sum	21	21	14

2. The participants were given a brief (10 minutes) presentation about the experiment task. This included an explanation of the risk analysis method, the example presented in the guidelines and the system description. Some examples of relevant and non-relevant risks about the example system were also presented in order to show concrete examples of what type of risks that can be identified, and on what level of abstraction the risks can be formulated on.

3. There was a short (5 minutes) session with questions and answers about the guidelines, system description, etc. The participants were given a chance to ask immediate questions that they had after reading the guidelines, but they were also allowed to ask questions during the later sessions.

4. The data collection forms were distributed for writing the risks found in the system during the risk analysis.

5. Each group was asked to perform the risk analyses.

   (a) 10 minutes were given for the planning step of the risk analysis. It was possible to have as short time as this since the participants had already read the system description.

   (b) 35 minutes were given to perform the remaining steps (risk identification, determine the likelihood level and the consequence level) of the risk analysis. During this time, every member of a group performed individual risk analysis.

6. Each group was given 20 minutes to compare and merge the individual risk lists to come up with a common group risk list.

7. After the collection of data forms the post-experiment questionnaire was given to all the participants.

5 Results

Table 3 shows the experiment results carried out in seminars I, II and III. In seminar I, the PBRA treatment was used by 7 groups (21 participants). It can be seen that

Table 3: The results from the main Experiment

Seminar	Group label	Applied treatment	# of relevant risks found	# of non-relevant risks found
I	M1	PBRA	14	3
I	M2	PBRA	19	0
I	M3	PBRA	13	1
I	M4	PBRA	14	0
I	M5	PBRA	8	1
I	M6	PBRA	10	2
I	M7	PBRA	14	1
II	T1	TRA	7	0
III	T2	TRA	9	0
III	T3	TRA	10	1
III	T4	TRA	11	0
III	T5	TRA	11	0
III	T6	TRA	10	0
III	T7	TRA	9	5

group M2 found the highest number of relevant risks, 19, and group M5 found the lowest number of relevant risks, 8. Group T7 found the highest number of non-relevant risks, 5, and M1 found 3. Groups M3, M5 and M7 found 1 non-relevant risk each. The remaining two groups (M2 and M4) found only relevant links.

In seminar II and III, TRA was carried out by the 7 groups. It can be seen that group T4 and T5 found most relevant risks, 11, and group T1 found least relevant risks, 7. Group T7 found most non-relevant risks, 5, and T3 found 1. All other groups did not find any non-relevant risk. The remaining five groups found only relevant risks.

As described in section 3.6, the experiment participants identified 5 new risks that were not present in the risk list identified by the experimenters. These new identified risks were also added in the risk list.

6 Analysis

The data collected from the experiment (the number of found relevant risks) was analyzed for normality. Figure 4 shows the normal distribution plot for both datasets (results of PBRA and TRA). The line on the left is from the TRA dataset, the data points are quite clearly forming a straight line. However, the line of PBRA dataset, on the right, does not look clearly straight. Since the datasets are rather small, it was decided to use the Shapiro-Wilk normal distribution test. It is used

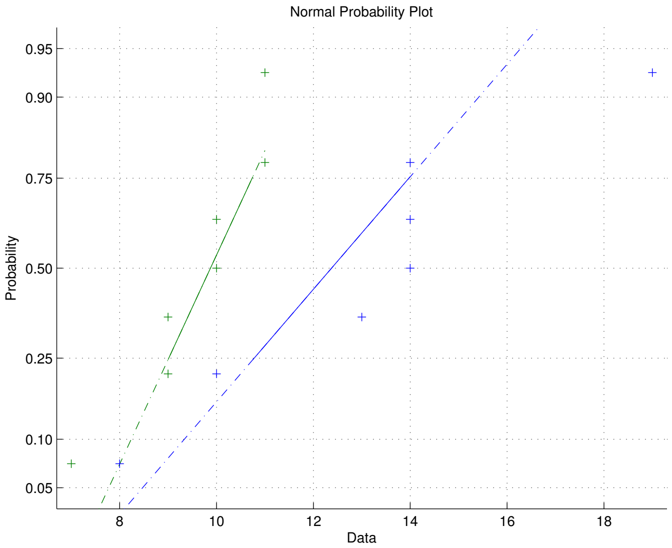


Figure 4: Normal distribution plot for the data

to test the null hypothesis that data comes from a normally distributed population. The null hypothesis was not rejected with the p-values 0.306 for TRA and 0.505 for PBRA. Both datasets proved to be normally distributed by using the Shapiro-Wilk normality test, which is one of the most powerful normality tests [23].

After testing the datasets for the normality, the T-test was performed to check for statistically significant difference between the efficiency of the TRA and PBRA methods measured by the number of identified relevant risks (research question RQ1). The T-test was applied to investigate the null hypothesis that the data from the two methods are normally distributed with equal means and equal but unknown variance, against the alternative that they are not. It revealed a statistical significant difference between TRA and PBRA methods by rejecting the null hypothesis H_0^1 with the p-value 0.027. As a result, we could accept the alternative hypothesis H_1^1 that the subjects found more relevant risks using the PBRA method.

The box plots for the number of found risks are shown in Figure 5. It can be seen that there is a difference in the number of found risks by TRA and PBRA methods. The participants that used TRA method found on average 9.57 relevant risks and the participants that used PBRA found 13.14 relevant risks.

To answer the second hypothesis regarding research question RQ1, the number of identified non-relevant risks with both treatments was first analyzed for normality. The *Shapiro-Wilk* normal distribution test was used to test the null hypothesis that data comes from a normally distributed population. The null hypothesis was

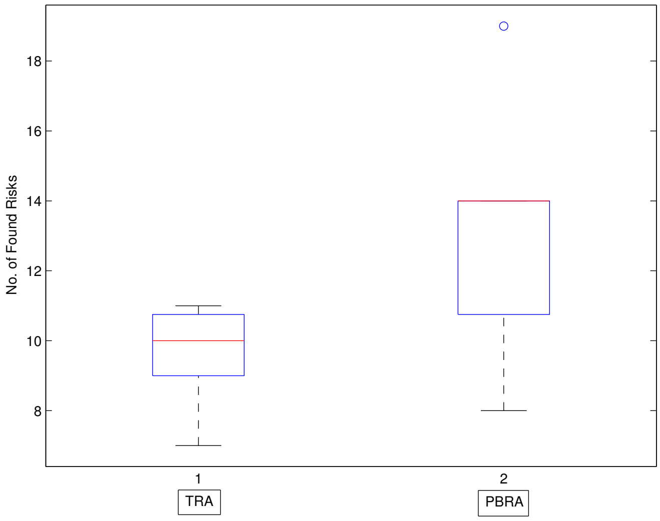


Figure 5: Box Plot of Found Risks

rejected for TRA dataset with the p-value 0.0014 and it was not rejected for the PBRA dataset with the p-value 0.587. Thus, it was decided to use non-parametric tests.

The results were tested for the statistical difference using the *Mann-Whitney U*-test. No statistically significant difference was revealed by the test resulting in the p-value of 0.249. Therefore, we cannot state that the PBRA method helped to identify fewer non-relevant risks.

For the research question RQ2 and RQ3, the following two questions were asked using an ordinal scale in the post-experiment questionnaire[5] respectively.

1. How difficult was the risk analysis method to use? (RQ2)

2. How confident are you that you have found all the relevant risks? (RQ3)

The data for RQ2 and RQ3 is collected by using an ordinal scale (Likert). Therefore, the collected data has been tested by using a non-parametric test (*Mann-Whitney U*-test).

[5] Due to space limitations we do not present complete survey results in this paper. We present the frequencies of the answers in Table 4. The questionnaire and the complete set of answers are available at
`http://serg.cs.lth.se/index.php?id=87041`.

Table 4: Summary of result regarding RQ2 and RQ3 given in frequencies of answers given for each option

Question Method	Frequencies of answers				
Difficulty (RQ2)	Very Easy	Easy	Fair	Difficult	Very difficult
TRA	4	4	11	2	0
PBRA	0	2	12	7	0
Confidence (RQ3)	Very confident	Confident	Fair	Not confident	Strongly not confident
TRA	0	1	5	9	6
PBRA	0	2	4	11	4

The collected data regarding RQ2 was saved in two vectors, $x1$ and $y1$, and *Mann-Whitney* U-test with the left tail was used to test the statistical difference in difficulty level while using both treatments. It tests the null hypothesis that data in vectors $x1$ and $y1$ comes from continuous distributions with equal medians, against the alternative that the median of $x1$ (TRA) is less than the median of $y1$ (PBRA). *Mann-Whitney* U-test rejected the null hypothesis H_0^2 with the p-value 0.004 meaning that the TRA method is less difficult to use than the PBRA method. The descriptive statistics, see Table 4, provides additional explanations for the test result. No subject considered PBRA *very easy* while four subjects considered TRA *very easy*. Moreover, seven subjects considered PBRA *difficult* while only two subjects considered TRA *difficult*.

The data regarding RQ3 was also saved in two vectors, $x2$ and $y2$, and *Mann-Whitney* U-test with the left tail was used to test the statistical difference in confidence level between the two samples. *Mann-Whitney* U-test could not reject the null hypothesis H_0^3 with the p-value 0.691 meaning that there is no statistical difference. Looking at Table 4, there could be several indications of lack of difference. Firstly, no subject was *very confident* of any method results. Secondly, six subjects were *strongly not confident* of the TRA method results and four subjects were *strongly not confident* of the PBRA method results. Thirdly, there are only subtle differences between the number of subjects that were *confident*, *fair*, *not confident* or *strongly not confident* about the results.

7 Validity Evaluation

The validity threats can be divided into four types [22]: conclusion, construct, internal, and external. We discuss the most relevant validity threats below.

7.1 Conclusion Validity

Use of wrong statistical tests: In order to reduce this threat, the collected data was investigated for normality before parametric tests (t-test) were used.

Reliability of treatment implementation: In order to reduce this threat, all subjects received the same standard instructions in all seminars. The illustrating example in the guidelines was also same for both treatments.

Random irrelevancies: Elements outside the experimental setting can disturb the experiment's results i.e. noise, and unplanned interrupt in the experiment. In order to reduce this threat, the subjects were not interrupted during the experiment and there was no significant noise in the experiment room. Subjects were instructed to discuss as quietly as possible while merging the individual risk lists.

Random heterogeneity of subjects: We believe that there is a very little chance of this threat because the students were selected from the same level of education (master students of engineering programs) and also had almost similar knowledge and background. That is, the students come from a rather homogeneous group.

7.2 Internal Validity

Maturation: In order to reduce this threat the subjects were asked to perform risk analyses in 35 minutes. It was assumed that 35 minutes would be enough to perform individual risk analysis and also subjects will not get bored.

Instrumentation: In order to reduce this threat the instrumentation of the experiment was carefully written and then evaluated by one of the co-authors. After that, an independent researcher evaluated the instrumentation. Finally, a pilot study was carried out to evaluate and improve the instrumentation.

Compensatory rivalry: This threat to internal validity is minimized since the subjects did not know that there is two different treatments.

7.3 Construct Validity

Construct validity generalizes the experiment's results to the theory of the experiment. Here, the theory is that PBRA method performs better and finds more relevant risks as compared to TRA. Previous work advocated using perspectives during risk analysis [5, 11, 24] as well as provided supporting evidence that perspectives support reviews and inspections [14]. This theory is based on the assumption that the use of different perspectives can support a better and more in-depth analysis by encouraging the participants to think of different safety and security requirements.

There could be a threat to the construct validity that the participants do not interpret relevant and non-relevant risks as the experimenter intended. There could be difference of risks interpretation between the participants and experimenters. Similarly, the likelihood and consequence levels can also be misinterpreted. In order to reduce the threats to construct validity, the guidelines were written to be as clear and understandable as possible, and help was provided by clarifying any ambiguity to the participants when they asked. An example was also mentioned in the guidelines to make them unambiguous and clear. A very simple and common system was selected for the experiment and we believed that almost all the participants have already used it many time. This means that the selected system was easy to understand without domain knowledge. Finally, a pilot study was also carried out to mitigate any potential ambiguities.

The fear to be evaluated (also known as evaluation apprehension) threat to validity was reduced by clearly stating that the results of the experiment do not have any affect on the studentsÕ final grades. It is not possible to track the individual participants of the experiment for evaluation because the participants were anonymous.

7.4 External Validity

Interaction of selection and treatment: There could be a chance of this threat because the subjects for the main experiment were students of a project course and are therefore not representative for the entire population. However, to reduce the affect of this threat, the pilot study was carried out by using experts from industry and academia. There was not a big difference in the number of identified relevant risks found by the industry experts and students.

Another threat to external validity is that the subjects were given 35 minutes for the individual risk analysis and then 20 minutes for the comparison and merger of individual risk lists. They were asked to find as many risks as they can but there was no upper or lower limit for the number of identified risks. The given time was also limited in order to reduce the effect of maturation. The time was chosen as a tradeoff between having the possibility to spend a lot of time and be sure to find “all” risks, and the risks of spending too much time and obtain maturation.

8 Discussion

The experiment confirms that subjects using PBRA found more relevant risks. This result provides supporting evidence about the potential of roles and perspectives in risk analysis, stretching outside a simple scenario of role-based risk analysis given by Yoran and Hoffman [24] and recommendations given by Leveson [11] and Ierace [5]. Moreover, our results suggest that perspectives could increase the efficiency of not only document reviews [14] but also risk analysis and identification. Contrary to expectations, our results do not bring the supporting evidence that PBRA helps to identify fewer non-relevant risks. This indicates that there can be an advantage to assign perspectives to participants in a risk analysis, as a complement to only rely on the more natural differences between different roles in a group.

Our experiment brings statistically significant evidence that PBRA is seen as more difficult than TRA. This does not have to be interpreted as negative for PBRA. It may be that this level of difficulty is necessary, and acceptable, if the results can be more relevant risks. It may also be possible that the higher difficulty level may not be appropriate for the rather inexperienced students participating in the experiment, this calls for a replication of this study with much more experienced practitioners.

Regarding the confidence in the identified risks, no statistical significance may be caused by a lack of experience and domain knowledge in train systems. The results in Table 4 seems to support this interpretation as most subjects were highly not confident about the risks identified using any of the methods. Thus, further studies with more experienced risk managers and engineers are needed to further explore this aspect.

9 Conclusions and Future Work

Involving perspectives into risk analysis brings a potential to increase the efficiency of the risk analysis and confidence in the identified risks. In this paper, we present the results from a study designed to experimentally assess the potential of perspectives in risk management and therefore further experimentally explore the suggestions given in previous work [1, 3, 5, 7, 11, 18, 24]. 43 subjects performed risks analysis of a software-controlled train door system using TRA and PBRA. We measured the efficiency of the methods by counting the number of relevant and non-relevant risks and a questionnaire to measure the difficulty of the methods and the confidence of the subjects in the identified risks.

Revisiting our research questions, we can with a statistical significance claim that PBRA helps to identify more relevant risks than TRA. On the other hand, our experiment failed to provide supporting evidence that PBRA helps to identify fewer non-relevant risks (RQ1). Contrary to expectations, this study did find with a statistical significance that PBRA is more difficult to use than TRA (RQ2). We interpret this result as a consequence of the subjects' limited experience in system engineering and rail domain. It may be that this level of difficulty is necessary, and acceptable, if the results can be more relevant risks. Finally, we cannot say that any of the studied methods generated risks with higher confidence (RQ3). However, most subjects were highly not confident about the risks identified using any of the methods.

In future work, we plan to replicate our study with practitioners experienced in rail domain. We also plan to apply PBRA on more complex systems by involving practitioners that have extensive experience in the system engineering approach and measure their performance. Finally, we plan to explore if different perspectives than used in this experiment (tester, train staff member and system engineer) impact the number of relevant and non-relevant risks identified using the PBRA method.

Acknowledgement

This work was funded by the Swedish Civil Contingencies Agency under a grant for PRIVAD, Program for Risk and Vulnerability Analysis Development.

Bibliography

[1] C. Alberts and A. Dorofee, *Managing Information Security Risks: The Octave Approach*, ser. SEI Series in Software Engineering. Addison-Wesley, 2003.

[2] V. R. Basili, S. Green, O. Laitenberger, F. Shull, S. Sørumgård, and M. V. Zelkowitz, "The empirical investigation of perspective-based reading," *Empirical Software Engineering*, vol. 1, no. 2, pp. 133–164, 1996.

[3] N. Boudriga, M. Hamdi, and J. Krichene, "Netram: A framework for information security risk management," Techno-parc El Ghazala, Route de Raoued, Ariana, 2083, Tunisia, Tech. Rep., 2007.

[4] C. A. Ericson, "Fault Tree Analysis - A History," in *proceedings of The 17:th International System Safety Conference*, 1999.

[5] S. Ierace, "The basics of FMEA, by robin e. mcdermott, raymond j. mikulak and michael r. beauregard," *Journal of Production Planning and Control*, vol. 21, no. 1, pp. 99–99, 2010.

[6] T. Ishimatsu, N. G. Leveson, J. Thomas, M. Katahira, Y. Miyamoto, and H. Nakao, "Modeling and hazard analysis using STPA," in *proceedings of the International Conference on Association for the Advancement of Space Safety*. NASA, Sep. 2010.

[7] ISO 27005:2011, "Information technology - Security techniques - Information security risk management," 2011.

[8] A. Jedlitschka, M. Ciolkowski, and D. Pfahl, "Reporting experiments in software engineering," in *Guide to Advanced Empirical Software Engineering*. Springer London, 2008, pp. 201–228.

[9] P. Johannessen, F. Torner, and J. Torin, "Actuator based hazard analysis for safety critical systems," in *Computer Safety, Reliability, and Security*, ser. Lecture Notes in Computer Science, vol. 3219. Springer, 2004, pp. 130–141.

[10] O. Laitenberger, K. E. Emam, and T. G. Harbich, "An internally replicated quasi-experimental comparison of checklist and perspective based reading of code documents," *IEEE Trans. Software Engineering*, vol. 27, no. 5, pp. 387–421, 2001.

[11] N. G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*. The MIT Press, 2012.

[12] C. Lindholm, J. P. Notander, and M. Höst, “A case study on software risk analysis in medical device development,” ser. Lecture Notes in Business Information Processing. Springer Berlin Heidelberg, 2012, vol. 94, pp. 143–158.

[13] R. J. Mikulak, R. McDermott, and M. Beauregard, *The Basics of FMEA*. Productivity Press, paper back, 2008.

[14] B. Regnell, P. Runeson, and T. Thelin, “Are the perspectives really different? further experimentation on scenario-based reading of requirements,” *Empirical Software Engineering*, vol. 5, no. 4, pp. 331–356, Dec. 2000.

[15] G. Sabaliauskaite, F. Matsukawa, S. Kusumoto, and K. Inoue, “An experimental comparison of checklist-based reading and perspective-based reading for UML design document inspection,” in *proceedings of the International Symposium on Empirical Software Engineering*, 2002, pp. 148 – 57.

[16] D. I. K. Sjoeberg, J. E. Hannay, O. Hansen, V. B. Kampenes, A. Karahasanovic, N.-K. Liborg, and A. C. Rekdal, “A survey of controlled experiments in software engineering,” *IEEE Transactions on Software Engineering*, vol. 31, no. 9, pp. 733–753, 2005.

[17] I. Sommerville, *Software Engineering*, 7th ed. Harlow, England: Addison-Wesley, 2010.

[18] G. Stoneburner, A. Goguen, and A. Feringa, *Risk Management Guide for Information Technology Systems*, ser. National Institute of Standards and Technology, Special Publication 800-30. U.S. Government Printing Office, 2002.

[19] S. M. Sulaman, K. Weyns, and M. Höst, “A review of research on risk analysis methods for it systems,” in *proceedings of the 17:th International Conference on Evaluation and Assessment in Software Engineering (EASE ’13)*. ACM, 2013, pp. 86–96.

[20] Swedish Civil Contingencies Agency (MSB), “Guide to risk and vulnerability analyses,” 2012.

[21] J. Thomas and N. G. Leveson, “Performing hazard analysis on complex, software- and human-intensive systems,” in *proceedings of The 29:th International System Safety Conference*, 2011.

[22] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in software engineering*. Springer Publishing Company, Incorporated, 2012.

[23] B. W. Yap and C. H. Sim, “Comparisons of various types of normality tests,” *Journal of Statistical Computation and Simulation*, vol. 81, no. 12, pp. 2141–2155, 2011.

[24] A. Yoran and L. J. Hoffman, "Role-based risk analysis," in *proceedings of the 20:th National Information Systems Security Conference*, 1997, pp. 37–51.

[25] Z. Zhang, V. Basili, and B. Shneideman, "Perspective-based usability inspection: an empirical validation of efficacy," *Empirical Software Engineering*, vol. 4, no. 1, pp. 43 – 69, 1999.

Paper IV

Comparison of the FMEA and STPA safety analysis methods – A case study

Abstract

As our society becomes more and more dependent on IT systems, failures of these systems can harm more and more people and organizations. Diligently performing risk and hazard analysis helps to minimize the potential harm of IT systems failures on the society and increases the probability of their undisturbed operation. Risk and hazard analysis is an important activity for the development and operation of critical software intensive systems, but the increased complexity and size puts additional requirements on the effectiveness of risk and hazard analysis methods.

This paper presents a qualitative comparison of two hazard analysis methods, Failure Mode and Effect Analysis (FMEA) and System Theoretic Process Analysis (STPA) using case study research methodology. Both methods have been applied on the same forward collision avoidance system to compare the effectiveness of the methods and to investigate what are the main differences between them. Furthermore, this study also evaluates the analysis process of both methods by using a qualitative criteria derived from the Technology Acceptance Model (TAM). The results of the FMEA analysis was compared to the results of the STPA analysis, which was which was presented in a previous study. Both analyses were conducted on the same forward collision avoidance system. The comparison shows that FMEA and STPA deliver similar analysis results.

Sardar Muhammad Sulaman, Armin Beer, Michael Felderer and Martin Höst, *Software Quality Journal*, “Online First”, 2017.

1 Introduction

The increasing dependence of our society on IT systems brings not only new development opportunities but also new severe risks and threats. As our daily life is almost completely dependent on IT systems, both for individuals and for organizations (private and public), failures of these IT systems can have serious negative consequences and effects on the society. In-depth and a completely performed risk and hazard analyses help to minimize the potential harm of IT systems failures on the society [18,40]. However, risk/hazard analyses of modern socio-technical systems are far from trivial, mainly due to the dynamic behavior that pervades almost every modern software intensive system and a high number of interacting components. As a result, many traditional low-level risk or hazard analysis methods fail to encompass the dynamic behavior of the systems, as they focus solely on the system component failures [18]. These traditional methods mainly focus on identification of critical components of a system and then either try to prevent the failures of these components or add redundant components. In case of dynamically changing systems, a new risk can emerge from wrong or non-synchronized commands that may lead to severe accidents. Therefore, new methods for performing risk and hazard analysis, optimized for dynamic systems, are required.

There are still a number of uncertainties when it comes to what risk and hazard analysis method to apply in a given situation. The main objective of this study is to empirically compare two existing risk analysis methods, Failure Mode and Effect Analysis (FMEA) and System Theoretic Process Analysis (STPA). The study compares the results of and investigates the effectiveness of the well-established bottom-up FMEA and the rather new top-down STPA hazard analysis methods by performing a comparison of how a hazard analysis is conducted for the same system. Furthermore, this study also evaluates the analysis process of both the methods by using a set of qualitative criteria, derived from the Technology Acceptance Model (TAM) [4, 5]. The results of the FMEA analysis yielded from this study are compared with the results of a previous study [41] that presents an STPA hazard analysis of a system. It should be noted that this study does not aim at comparing both methods quantitatively, but instead to understand the differences through a qualitative analysis. That is, we investigate both methods qualitatively by analyzing hazard analysis results gathered by applying both methods, FMEA and STPA, on a collision avoidance system. Furthermore, this study also evaluates the analysis process of both methods by using the qualitative criteria derived from the Technology Acceptance Model (TAM).

The remainder of this paper is structured as follows. Section 2 provides a background on FMEA, STPA, and other risk and hazard analysis methods, as well as an overview of the forward collision avoidance system which is analyzed. Section 2 presents related work. Section 4 discusses the design of the case study and Section 4.5 presents the data collection procedure. Section 5 presents the results of the conducted analyses, Section 6 provides an analysis of the results, and Section 7 discusses the validity of the study Section 8 discusses the results from the study, and Section 9 concludes the paper.

2 Background

This section presents a brief description of the FMEA and STPA hazard analysis methods that are compared in this study. It also provides an overview of other existing risk and hazard analysis methods. In addition, this section presents the description of the selected system, forward collision avoidance system, on which both methods are applied.

2.1 FMEA

FMEA is a bottom-up analysis method that is used to identify potential failure modes with the causes for all the parts in system to find negative effects [14, 25]. The analysis starts with the lowest level components and proceeds up to the failure effect of the overall system. A failure effect at a lower level becomes a failure mode of the component at the next higher level. FMEA also measures severity, occurrence and detection probability that are used to calculate risk priority numbers for the identified failure modes. The main purpose of FMEA is to identify potential problems in the early design process of a system or product that can affect its safety and performance, and to introduce countermeasures to mitigate or minimize the effects of the identified potential problems (failure modes). Moreover, FMEA can complement FTA and identify many more failure modes and causes [24]. Failure Modes and Effects Criticality Analysis (FMECA) is an extension to FMEA that ranks the identified failure modes based on their severity, which is used for prioritization of countermeasures [1, 25].

Another extension is provided by [12] who introduced an extension to the conventional FMEA, namely “the probabilistic FMEA”. It has the advantage of formally including rates at which component failures can occur. This method helps safety engineers to formally identify if a failure mode occurs with a probability higher than its tolerable hazard rate.

Software FMEA (SFMEA) [29] is an extension to system FMEA to analyze software-intensive system components, such as embedded real-time systems. FMEA was originally aimed at the reliability of hardware. However, its benefits for performing a software FMEA were also shown by [38]. Software FMEA considers specific aspects of software in an FMEA, for instance the fact that software components often do not fail in the traditional way but instead result in incorrect behavior. Software FMEA is a preventive measure for risk management and should therefore be carried out during the development of a system. [36] states that SFMEA is best suited for a qualitative high level analysis of a system in the early design phase. A general limitation of the FMEA analysis is the restriction to analyze only single cause of an effect. By assessing the severity of failure effects, the probability of their occurrence and the detection of the probability of failure causes, a distinction between components of high or low risk is feasible and appropriate actions can be planned.

FMEA is applied to components in the design phase of the software system life cycle. The level of abstraction can take the levels of the V-model into account [23]. In this study both Software FMEA and system FMEA were applied in the following five steps:

1. Partition the system to be examined into subsystems and components, taking the architecture of hardware and software into account.

2. Assign the application function to each component. In this step, functional and non-functional requirements have to be interpreted.

3. Determine and analyze the potential failure mode, cause of failure and failure effect that can lead to a hazardous state. For instance the failure mode 'false break activation' could have the cause of a defect in the SW of pressure determination and the effect of a potential crash situation between two cars. Another example regarding security is, for instance, a failure or threat mode classifying the way in which vulnerabilities are exploited [36]. A threat mode could be 'Attacker is pretending to be a measurement device' violating the integrity of the system. The cause could be an encryption problem or security breach and results in 'System is unreliable and potentially unsafe'.

   Each failure mode represents potential product failures that can occur. Failure mode, cause and effect are entered in the spreadsheet fields related to the appropriate component and function. The causal factors are associated with software defects, interface errors (architectural, protocol), HW/SW interaction (signaling), reliability, security and real time constraints. The potential failure effects could be the following: Risk of collision, the operator is not alerted, a potential crash situation or the authorization of external hackers to manipulate the collision avoidance system.

4. Evaluate risk and calculate the risk priority number (RPN). To calculate the RPN as described by [22] the severity of the failure effect, the probability of their occurrence and the detectability of the failure causes have to be assessed first.

   The abbreviations used below for severity, probability, and detectability i.e. B, A, and E are adapted from the study [20].

   - Severity (B): The severity value is assessed taking the potential failure effect into account. A five-point Likert scale is used, ranking the impact from 1 (no impact) to 5 (catastrophic, i.e., potential crash situation)

- Probability of occurrence (A): To assess the probability of occurrence the complexity, the potential failure mode and cause of a failure have to be taken into account. A five-point Likert scale is used to rank the probability, starting from 1 (very low, 0.01%) to 5 (very high, 50%).
- Detectability (E): The detectability depends on the complexity of the HW/SW component and potential cause of a failure. A five-point Likert scale is used to rank the detectability, starting from 1 (very low probability (0 to 19%) that current controls will detect the cause) to 5 (very high probability (80 to 100%) that current controls will detect the cause).
- Calculation of risk priority number: RPN is calculated by multiplying the values of severity, probability of occurrence and detectability. $RPN = B \times A \times E$, where B, A, and E denote severity, probability, and detectability according to above. RPN ranges from 1 to 125.

5. Specify defect avoidance or risk mitigation measures. This step is not taken into account in the current case study.

Software FMEA (SFMEA) [22] allows the categorization of components taking the degree of their failure risk into account. It fosters the risk oriented development of software intensive systems. The complexity of a software system plays an important role in the development and the maintenance of products. SFMEA relates the complexity of a component to the probability of a failure. The practical experience in large-scale system development of the second author shows that if requirements are adapted iteratively, the complexity of the affected software components increases. In case also the system architecture has to be altered, the complexity will even increase significantly. SFMEA enables the partitioning of components into sets of different complexity. It considers complexity as an important influence factor in a hazard analysis. For example, a developer who focuses on the implementation of specific functions may overlook relations in the architecture of the system and therefore insert software defects. The benefit of FMEA is that complexity is taken into account to assess the risk of a failure and to issue preventive and analytical quality assurance measures like software testing [9].

2.2 STPA

The System Theoretic Process Analysis (STPA) method for hazard analysis focuses on analyzing the dynamic behavior of the systems and is intended to provide advantages over traditional hazard analysis methods [19]. STPA is a top-down method, just like the FTA method presented in section 2.3. However, STPA uses a model of the system that consists of a functional control diagram instead of a physical component diagram used by traditional hazard analysis methods. STPA is based on system theory unlike FMEA, which is based on reliability theory. Moreover, STPA considers safety as a system's control (constraint) problem rather than a component failure problem. Among the most prominent benefits of STPA, [15] listed the efficiency of the later phase of STPA when the broader scenarios are analyzed. According to [15], STPA takes into consideration the interactions of system components, and considers the evaluated system and its components as a collection of interacting control loops (control action and safety constraints on the component behaviors). STPA requires a control structure diagram for hazard analysis consisting of components of a system and their paths of control and feedback, i.e., acknowledgment. STPA is applied in the following two steps:

1. Identify the potential for inadequate control of the system that could lead to a hazardous state. A hazardous state is a state that violates the system's safety requirements or constraints and therefore can cause some loss regarding life, mission, or financial.

2. Determine how each potentially hazardous control action, identified in step 1, could occur (finding causal factors). An inadequate control action can lead a system to a hazardous state, and that could be one of the following:

    - A control action required is not provided
    - An unsafe (incorrect) control action is provided
    - A control action is provided too early or too late (wrong time or sequence)
    - A control action is stopped too early or applied too long.

The aforementioned term 'provided' means the correct delivery of a control action or command from one component to another component of the system. A control action or command can encounter communication errors, e.g., delayed, failure, corrupted, etc. For the application of STPA, a functional control structure diagram of the system is required and all control loops in system are identified from it. After this, in each control loop all components that contribute to unsafe behavior of the studied system are identified.

[41] applied STPA on a socio-technical system that has three controllers. They are critical components of system because they contain a process model [19]. The controller receives input from almost all components of the system, e.g., sensors and actuators and then it performs internal calculations to issue a command.

2.3 Other Methods

A few more risk and hazard analysis methods exist in addition to FMEA and STPA. For example, there exist a number of low-level risk analysis methods that analyze systems and subsystems at lower level considering only systems and their components. Some of the most well-known methods are Fault Tree Analysis (FTA) [7] and Hazard and Operability Study (HAZOP) [32].

FTA is a top-down hazard analysis approach. It is a deductive approach and carried out by repeatedly asking: how can this (a specific undesirable event) happen, and what are the causes of this event? It involves a logical diagram that shows the relation between the system components and their failures. [7] presented a review of the research performed on FTA with its advantages and shortcomings. Because FMEA is restricted to analyze only single cause of an effect, FTA augments the feasibility of FMEA. An analysis using FTA in combination with FMEA may support an assessment considering, for instance all security risks [36].

HAZOP is a qualitative technique commonly used in the planning phase of system development. It identifies hazards by analyzing how a deviation can arise from a design specification of a system. It is used to identify the critical aspects of a system design for further analysis. It can also be used to analyze an operational system. A multi-disciplinary team of 5–6 analysts lead by a leader usually carries out the HAZOP analysis. The HAZOP team identifies different scenarios that may result in a hazard or an operational problem, and then their causes and consequences are identified and analyzed [21].

2.4 Forward Collision Avoidance System

The forward collision avoidance (FCA) system was selected in this study to compare and evaluate hazard analysis methods. Here, it should be noted that the main focus of this study is on the comparison and evaluation of the analysis methods (FMEA and STPA) rather than the FCA system itself. Moreover, the FCA system was selected because it was decided to use an operational and real system to for the analysis.

The FCA system alerts a driver of a vehicle about crash situations and applies automatic brakes after a certain time period if the driver does not respond to a warning alert that provides passive and active safety. The system performs two main functions: (1) object/obstacle detection (by using forward-looking sensors that detect hindrance in front of the vehicle) and (2) generation of warning or applying auto breaks (passive/active response). The forward-looking sensors could have some or all of these components: radar, infrared, motion sensors and cameras [2, 3].

Figure 1 shows the forward collision avoidance system [2] that has been divided into parts A (the collision controller), B (the brake controller), and C (the engine torque controller).

The *collision controller* (part A of the system) is connected with the following system components: The *collision controller* is connected with the *radar* and the *camera* through the *object detection system*. An object detection system could have more sensors or devices to detect an object in front of the vehicle. In this

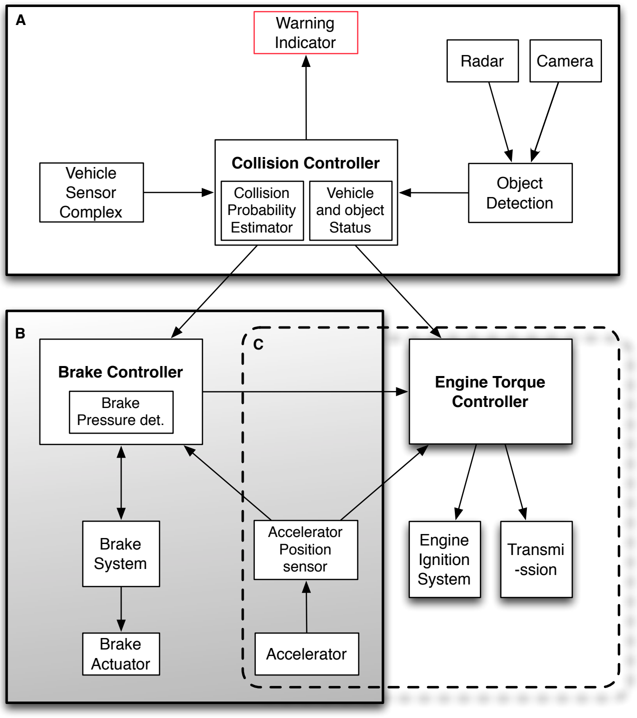


Figure 1: Forward collision avoidance system with autonomous braking [2]

study, we suppose that it uses more than one motion sensors to complement the radar and the camera. The object detection system could be very simple or very complex but in this study we consider the simple version. In the next sections we will only refer to the object detection system instead of referring individually to the radar, camera and sensors.

The *vehicle sensor complex* is also connected with the collision controller that generates a signal, and then sends it to the collision controller. The vehicle sensor complex consists of several vehicle system sensors, such as a brake position sensor, throttle position sensor, steering sensor, suspension sensor, speed sensor, and seat belt sensor. The information from these sensors can either be used individually or together to complement the collision avoidance system.

The *warning indicator* connected with the collision controller generates a collision warning signal in response to the collision-assessment of the collision controller. The collision controller gets input from the object detection system and the vehicle sensor complex when it performs the collision assessment.

The *collision controller* (shown in part A), works as follows: The *vehicle* and *object* status provider in the collision controller calculates and provides the current status of the object in front of the vehicle and the current status of the vehicle to the collision probability estimator. The *collision probability estimator* in the collision controller calculates the vehicle collision probability based on the received information. If there is a risk of collision then the estimator sends a signal to the indicator, which is for the vehicle's operator. This is known as collision detection, which is a passive safety system that just warns the vehicle operator. If the vehicle operator does not respond to the collision warning then the system activates the collision avoidance system also known as the active safety (autonomous brake). The *collision controller* uses an algorithm to estimate the risk of collision and generates a collision-assessment signal. It is a critical component of the collision avoidance system, because both active safety and passive safety depend on the output of this component. It also calculates some other parameters, such as the time to collision that is going to happen, point of collision, object identification, etc. If the vehicle's operator responds to the collision warning on time then the forward collision avoidance system resets all its components and calculated parameters. However, if the operator does not respond to the received warning then the collision controller sends a collision-assessment signal with the object and vehicle status signals to the *brake* and *engine torque* controllers to apply autonomous brake.

The *brake controller* (part B of the system) works as follows: It receives the *vehicle status* signal, *detected-object status* signal and *collision-assessment* signal from the *collision controller*. The brake controller has one *brake pressure measurement or determination* component that determines the required brake pressure for the current situation based on the received information from the *collision controller* and *accelerator position sensor*. After determining the required brake pressure, the brake controller sends an autonomous brake signal to the *brake system* and to the *engine torque controller*. The *brake system* has one *brake pedal* and one *brake actuator* that apply the autonomous brakes. One important action of the brake controller and brake system is that they allow the vehicle's operator intervention during the application of autonomous braking. Operator can increase the brake pressure by intervening the autonomous braking that also deactivate the collision avoidance system in that particular collision situation. The *engine torque controller* (part C of the system) works as follows: It reduces the torque to almost zero after receiving signals from the *collision controller* and *brake controller* during the application of autonomous braking by using different methods like, by limiting air or fuel supply to engine, downshifting the transmission, and switching the engine off. The *accelerator position sensor* is electrically coupled to the brake controller and the engine torque controller that indicates and provides the position of accelerator.

2.5 Hazard

The term "hazard" used in this study generally defined as "anything that has the potential to do harm" or "anything that can lead to an accident". According to [18] if every state of a system is considered then system can always pose a potential danger or itself in danger. Therefore, this definition should preclude states that the system must normally be in to accomplish the mission. However, this study is not trying to define a new definition for "hazard" instead it follows the general definition adapted by [18].

> *"A system state or set of conditions that, together with a particular set of worst-case environmental conditions, will lead to an accident (loss)."*

3 Related Work

The research objective of the current study is to compare the STPA and FMEA hazard analysis methods. There exist some studies that have compared different risk and hazard analysis methods. For example, [39] performed a comparison of two safety analysis methods, MisUse Case (MUC) method and FMEA. The MUC method was originally proposed for eliciting security requirements [37], but it has also been used for safety analysis. The MUC method was developed by the software community as an alternative to FMEA and HAZOP. Both methods were compared in an experiment to investigate which method is better than the other for identifying failure modes and if one of the methods was easier to learn and to use. The authors concluded that when the system's requirements are described as use cases MUC is better than FMEA for analyzing failure modes related to user interactions. Furthermore, FMEA is better than MUC for analyzing failure modes related to the inner working of the system. The authors also concluded that MUC will create less confusion and in general be easier to use than FMEA.

[44] compared and discussed three well known risk analysis methods by applying them on a box fan, FMEA, AFMEA (advanced failure mode and effect analysis) [8], and FTA. The authors presented the advantages and disadvantages of these methods and concluded their study with an attempt of combining both deductive (top down) and inductive (bottom up) risk/safety analysis methods.

[16] compared the STPA hazards analysis results with the FTA analysis results that were used to certify the H-II Transfer Vehicle (HTV). The HTV is an unmanned cargo transfer spacecraft that is launched from the Tanegashima Space Center aboard the H-IIB rocket and delivers supplies to the international space station (ISS). In the development of the HTV the potential HTV hazards were analyzed using FTA and during the analysis the NASA safety requirements were also considered. After comparison of the results, the authors concluded that STPA identified all the traditional causes of losses identified by FTA and FMEA, but it also identifies additional causes. The additional factors include those that cannot be identified using fault tree analysis, including software and system design as well as system integration.

[10, 11] analyzed the NextGen In-Trail Procedures (ITP) application by using the STPA analysis method and compared its results with the official NextGen ITP application analysis [34]. NextGen is the next generation of air traffic management systems that contains In-Trail Procedures application. In-Trail Procedures (ITP) is an application of Automatic Dependent Surveillance-Broadcast (ADS-B) that allows aircraft to change flight levels in areas where current radar separation standards would prevent desirable altitude changes [13]. To summarize, ITP helps to increase operational efficiency and throughput in oceanic airspace [11]. The authors concluded that STPA found more potential causes of the hazards considered (violation of separation requirements) than the traditional hazard analysis performed on ITP [34]. In the comparison, the authors identified 19 safety requirements that were not in either of the two official NextGen analysis documents.

[10, 11] also compared STPA with bottom-up and other top-down analysis techniques. According to the authors, bottom-up analysis techniques, FMEA, start by identifying all possible failures. This list can be very long if there are a lot of components and all the permutations and combinations of component failures are considered. However, STPA only identifies the failures and other causes that can lead to a system hazard and does not start by identifying all possible failures. Moreover, in the top-down STPA analysis approach, the analyst can stop refining causes at the point where an effective mitigation can be identified and does not go down any further in detail. The analyst only has to continue refining causes if an acceptable mitigation cannot be designed. That is the major difference between STPA and FMEA (and any other bottom-up technique), which explains the differences in time and effort required [10, 11].

Furthermore, [26] evaluated STPA in a case study where it was applied on an operational crew-return vehicle design. The authors conclude that with STPA it is possible to recognize safety requirements and constraints of the system before the detailed design.

[31] compare risk identification techniques for safety and security requirements. From the safety field the Functional Hazard Assessment (FHA), the Preliminary Hazard Analysis (PHA), HAZOP, FMEA as well as FTA are considered. Each technique is assessed based on several quality criteria addressing the context, the application area, the application method as well as advantages and disadvantages of utilizing the technique. The assessment is based on evidence reported in the literature. The authors conclude that risk identification techniques for safety are more mature than for security and that they have found a balance between creativity and formalism, which is needed for identification process.

As it can be noticed from the literature mentioned in this section, STPA is a quite new analysis technique as compared to other techniques (FMEA, FTA etc.). [10, 11] mention that traditional analysis methods (FMEA, FTA, etc.) are more than 50 years old and, while analyzing safety critical software-intensive systems, they cannot identify software faults or the errors pertaining to dynamic behavior of the system. SFMEA [29] and especially also STPA [19] have been developed to overcome the existing problems in traditional analysis methods. According to [19] and [10, 11] STPA can find more component interaction, software and human hazards than traditional methods. Therefore, according to the authors, STPA is more effective because it is developed by considering system thinking that considers whole system as a single unit and finds more hazards. Moreover, previously STPA is compared and evaluated with bottom-up methods (e.g. FMEA) by the same authors who presented it or they were involved in its development. Several authors [11,15,18,26,28,42] reported positive outcomes from applying STPA on various systems. However, the traditional methods are still in use in practice even though they are more than 50 years old for the analysis of safety critical systems in early design, development and operational phases. This means that there is a need for further investigation of effectiveness of the STPA method compared to other traditional safety analysis methods that are used in industry. If further investigations find STPA as an effective method then these results can help industry to shift to this new analysis method.

To summarize, it is interesting to investigate what are the main differences in STPA and other traditional methods (in this case FMEA) and also the types of hazards identified by them.

4 Case Study Design

4.1 Research Objective

The main objective of this study is to compare and investigate effectiveness of FMEA and STPA hazard analysis methods in the software intensive safety-critical system domain. In this study, hazard analysis results from both FMEA and STPA are compared to find the main differences in methods by investigating e.g., types of hazards identified by them. Based on the comparison results, this study also investigates which method is more effective. Moreover, this study also evaluates the analysis process of both methods by using a qualitative criteria derived from the Technology Acceptance Model (TAM).

4.2 Research Questions

The aforementioned research objective has been broken down in the following main research questions.

RQ1: What are the main differences between the selected hazard analysis methods regarding types of the identified hazards?

RQ2: What are the main differences in the analysis process of both methods?

RQ3: Which method is more effective, FMEA or STPA?

In our context, effectiveness is high if a large number of relevant hazards but only a small number of non-relevant hazards are identified.

RQ1 is answered by analyzing and investigating the results from both the FMEA and STPA analyses to find the main differences between the two methods. Five error types were defined based on the related studies [10, 11, 19] and then all the identified hazards are classified according to the defined error types. Furthermore, the classification of error types (identified hazards) is investigated to answer which method finds what types of hazards.

RQ2 is answered by developing the qualitative criteria to evaluate the analysis process of both methods. The qualitative criteria were derived from the Technology Acceptance Model (TAM) to evaluate the analysis process considering ease of use and usefulness. Then, the developed qualitative criteria were applied on both methods to analyze and evaluate them.

RQ3 is answered by analyzing and investigating the results from both the FMEA and STPA analyses. It should be noted that, in this research initiative hazard analysis of collision avoidance system is carried out using only the FMEA method. After this, the FMEA results are compared to the STPA results found in a previous study [41].

4.3 Research Methodology

In this study the FMEA hazard analysis method is applied on the forward collision avoidance system in order to compare the results with the results of the previous study [41]. In the previous study, the STPA hazard analysis method was applied on

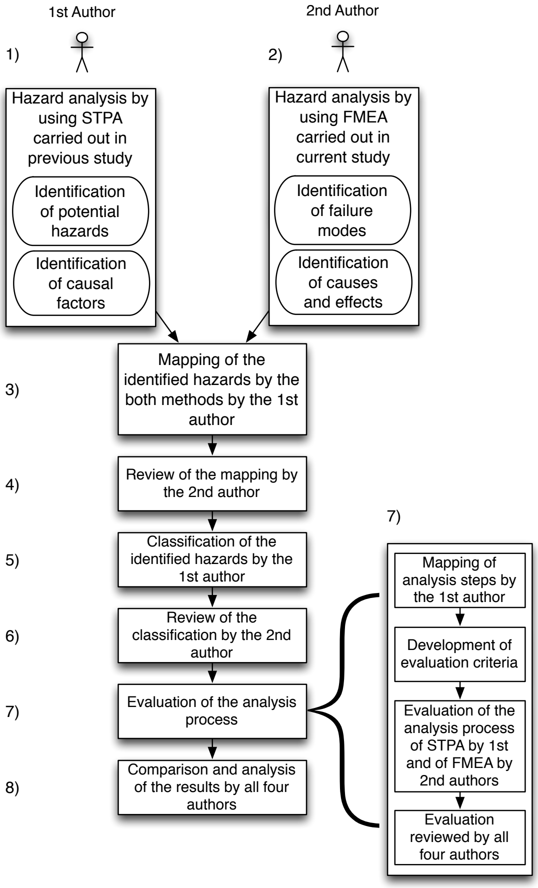

1st Author
2nd Author
1)
2)
Hazard analysis by using STPA carried out in previous study
Identification of potential hazards
Identification of causal factors
Hazard analysis by using FMEA carried out in current study
Identification of failure modes
Identification of causes and effects
3)
Mapping of the identified hazards by the both methods by the 1st author
4)
Review of the mapping by the 2nd author
5)
Classification of the identified hazards by the 1st author
6)
Review of the classification by the 2nd author
7)
Evaluation of the analysis process
8)
Comparison and analysis of the results by all four authors
7)
Mapping of analysis steps by the 1st author
Development of evaluation criteria
Evaluation of the analysis process of STPA by 1st and of FMEA by 2nd authors
Evaluation reviewed by all four authors

Figure 2: Steps taken for the carried out research

the same system in order to understand more about STPA and assess its effectiveness and efficiency. The steps carried out for the presented research in this study are shown in Figure 2.

Step 1 denotes the steps carried out in the previous study [41], where the first author of the current study analyzed the forward collision avoidance system and identified inadequate control commands or events. After this, the identified inadequate control commands or events were analyzed for their causal factors.

In step 2, the second author of this study applied the FMEA method on the same collision avoidance system to analyze operational hazards in it. The first author already knew the existing hazards in the selected system because he had applied STPA on the selected system in the previous study [41]. Therefore, to improve the research validity, it was decided that the first author would not apply the FMEA method; instead the second author would carry out FMEA analysis as he has experience of analyzing safety critical systems. For the FMEA analysis, all the documents about the selected system description that were used during the analysis in the previous study [41] were provided to the second author of this study to apply FMEA. During the FMEA analysis carried out in this study a number of measures were taken in order to increase the research validity and to decrease researcher bias. The same system information and description were available to the second author to analyze the system, as were used in the previous study, and the second author of this study did not review the previous study results.

After this, in step 3, the first author of this study performed an initial comparison (mapping) of the identified hazards yielded from the FMEA analysis and STPA analysis. The first author created a list of the common hazards that were identified by both analysis methods and another list was created for the distinct hazards identified by only FMEA or STPA.

Then, in step 4, the second author reviewed the initial comparison performed by the first author. The second author identified one more hazard (no. 18 in Table 4) as a common hazard.

In step 5, the first author classified all the identified hazards into the following five error categories: component interaction error, software error, human errors, component error, and system error. These categories were selected because the STPA [10, 11, 19] method claims to identify these types of hazards. According to these sources, STPA can find more component interaction, software and human hazards, which could be investigated by classifying the hazards in this way.

After this, in step 6, the second author reviewed the classification performed by the first author, and additionally the third and fourth author reviewed the results of step 1 to 6.

Then, in step 7, all four authors had a discussion regarding the development of qualitative criteria to evaluate the analysis process of both methods. After this discussion, the criteria were developed and then the first author mapped analysis steps of both methods. Then, the first author evaluated STPA and the second author evaluated FMEA according to the developed evaluation criteria. After this, the third and fourth authors reviewed the mapping and evaluations performed by the first and second authors.

Finally in step 8, a final comparison and analysis was performed by the all authors to investigate the differences between the both methods.

4.4 Case and Unit of Analysis

As the objective of this study is to evaluate and compare hazard analysis methods, the case for this study is a composite case that consists of a risk analysis method and a system on which the method is applied to analyze hazards. The selected case for this study is the FMEA and STPA risk analysis methods along with collision avoidance system. A similar case study was carried out in an earlier study [41] that also contains a composite case consisting of a hazard analysis method, STPA, and a system, collision avoidance system, on which method was applied to analyze hazards. However, in this study the case consists of both methods and the analyzed system because the objective is to evaluate and compare both methods based on the results yielded from the analyses.

4.5 Data Collection Procedures

In this study, the system description along with the system control structure diagram that shows how it works is used as study objects for the hazard analysis and evaluation of the methods. The system description is gathered from the existing patents for collision avoidance systems and also from the published literature for collision avoidance system [2, 3]. Moreover, data collected through the hazard analysis from both methods is also used for the analysis and evaluation of the methods. Besides this the expert opinions and knowledge are also used to evaluate and investigate the performed hazard analyses for analysis and results.

5 Results

5.1 Safety Analysis Using FMEA

The risk analysis using FMEA is performed to view the occurrence of failures in the collision avoidance system in a preventive manner. Table 1 shows the failure mode and effect analysis for the collision avoidance system. The FMEA was performed by the second author of this study according the procedure described in Section 2.1.

The first three columns show the identified subsystems, components and functions, for instance FMEA No. 9 the brake actuator, which switches from the auto brake to the manual brake, in case of a failure (corresponding to steps 1 and 2). The fourth column (Potential Failure Mode(s)) shows the failure mode, i.e., ‘Activation of manual brake fails’, corresponding to step 3. Each failure mode is taken

Table 1: Failure Mode and Effects Analysis (Quality Risk Analysis)

No.	System	Component	Application Function	Potential Failure Mode(s)	Potential Cause of Failure	Potential Failure Effect	Risk assesment				Risk mitigation measures
							B	A	E	RPN	
1	Part A: Collision controller system	Collision probability estimator	Calculation of probability of collision	Erroneous probability estimation	SW-defect: Calculating probability or in taking environment constraints into account.	Risk of collision with vehicle in front	4	4	5	**80**	Funct. test - strength=3, code review
2		Vehicle and object status	Calculation of speed and position of object	Erroneous calculation of speed of object	Erroneous specification or implementation of state transitions.	Risk of collision with vehicle in front	4	4	4	**64**	Funct. test - strength=3
3		Warning indicator	Man-machine interface	Display of warning fails	SW defect	Operator is not alerted about potential danger	3	3	2	**18**	Funct. test - strength=2
4		Vehicle sensor complex	Steering of collision controller	Erroneous signaling of sensors	Architecture erroneous or missing services provided by sensors.	Operator is not alerted about potential danger	3	4	4	**48**	Funct. test - strength=3, design review
5		Object Detection	Object detection by radar or camera	Erroneous radar and camera	SW defect or failure: object detection devices.	Operator is not alerted about potential danger	3	3	2	**18**	Funct. test - strength=3
6		Collision controller interface	Switch over to complement device	Object recognition device fails	SW defect or failure of communication stack.	Operator is not alerted about potential danger	3	3	4	**36**	Funct. test - strength=3, code review
7	Part B: Brake controller system	Brake controller	Activation of brake system and engine torque controller	Steering of brakes fails	SW defect or failure of communication stack.	Auto brake is not activated properly, risk of collision	4	3	2	**24**	Funct. test - strength=3
8		Brake pressure determination	Calculation of brake pressure	Steering of brakes fails	SW defect in pressure determination.	Auto brake is not activated properly, risk of collision	4	3	5	**60**	Funct. test - strength=3, code review
9		Brake system	Switch over from auto brake to manual brake	Activation of manual brake fails	SW defect: handling events or queues.	Improper brake activation, potential crash situation	5	4	4	80	Funct. test - strength=3, code review
10		Brake system	Switch over from manual brake to auto brake system	Activation of auto brake fails	SW defect: handling events or queues.	Improper brake activation, potential crash situation	5	4	4	**80**	Funct. test - strength=3, code review
11		Engine torque controller interface	Deactivation of torque controller	Erroneous torque controller is still active	SW defect or missing services provided: engine torque controller	Operator is not alerted about potential crash situation	4	3	3	**36**	Funct. test - strength=3, code review
12		Collision controller interface	Deactivation of collision controller	Automatic collision avoidance system is still running, when it should stop.	SW defect: activation or deactivation of collision avoidance system	Operator is not alerted about potential crash situation	4	2	3	**24**	Funct. test - strength=3, code review
13	Part C: Engine torque controller system	Collision controller interface	Receiption of data from collision controller	Vehicle and object status determination fails.	Architecture erroneous or missing services provided by components, sensors.	Operator is not alerted about potential crash situation	4	3	2	**24**	Funct. test - strength=2, design review
14		Brake controller interface	Receiption of data from braking system	Brake pressure determination fails	SW defect or missing services provided: interface, sensors.	Operator is not alerted about potential crash situation	4	3	2	**24**	Funct. test - strength=2
15		Accelerator position sensor and accelerator	Calculation of acceleration and position	Erroneous interpretation of sensor signals	SW defect: processing of sensor data.	Operator is not alerted about potential danger	3	3	4	**36**	Funct. test - strength=3, code review
16		Engine ignition system	Steering of engine ignition	Engine ignition is not activated	SW defect: steering of engine ignition.	Improper stop of vehicle and potential crash situation.	5	2	4	**40**	Funct. test - strength=3, code review
17		Transmission	Steering of transmission	Transmission downshifting fails.	SW defect or failure of communication stack.	Potential crash situation	5	2	4	**40**	Funct. test - strength=3
18		Engine torque controller	Torque reduction	Torque is not reduced	SW defect: State recognition	Operator is not alerted about potential danger	3	2	4	**24**	Funct. test - strength=3
19	Parts A,B,C	Operating system	Communications	Bad performance	SW defect: handling events or queues.	Potential crash situation	5	2	4	**40**	Performance test - strength=3
20		SW / HW components	Processing	Frequent restarts of the system	Sporadic SW defects	System is no longer reliable	4	3	5	**60**	Stress test - strength=3
21		Trust boundaries	Protecting system assets	Attacker is pretending to be a measurement device	Encryption problem or security breach.	System is unreliable and potentially unsafe.	5	1	3	**15**	Security test with attack pattern

to the potential causes (column 5) and effect of a failure (column 6). For instance, focusing on FMEA No. 9, the cause may be a software defect of handling events or queues, as a consequence improper brake activation and a potential 'crash' with an adjacent vehicle may occur. Each failure mode would be a hazard for a safe usage of the product.

In step 4 the worst-case impact of the effect of a failure, i.e., severity, the probability of occurrence and detectability are assessed. For example for FMEA No. 9, the severity is 5, because of a potential 'crash' situation, the probability of occurrence is 4 (high), because the complexity of the HW/SW component is very high and the likelihood of a failure is 10%. Its detectability is 4 (low), because the probability, that current reviews and testing of artifacts will detect the defect, is 20%–39% only.

The RPN is, as described above, calculated by multiplying the values of severity, probability of occurrence and detectability, $RPN = B \times A \times E$, which means that it ranges from 1 to 125. In the case of FMEA No. 9 the $RPN = 80$, which is the highest value shown in Table 1. To mitigate the risk of a software failure in the first place (step 5), extensive functional tests and code reviews are performed in the development phase. However, the RPN value is in some cases misleading. For instance, in FMEA No. 21, the effect of the failure mode 'attacker is pretending to be a measurement device' is defined as 'system is unreliable and potentially unsafe', the severity is 5, but the probability of occurrence is very low (0.01%), the detectability is 3 (medium probability) and the $RPN = 15$ only. Lower detectability, however, results in more risk and is therefore ranked higher. An intrusive attack of the collision avoidance system has to be blocked, for instance by encryption of signals, to mitigate the risk of manipulation of the braking or engine torque controller system. Security tests based on attack patterns are performed during development to mitigate the risk of a security breach in the system.

To summarize to reduce the occurrence of probability, error preventive measures during development, such as coding guidelines should be used. To mitigate the risk of a software failure during operating, code reviews as well as functional, performance and security tests are performed during development. The intensity of testing takes the complexity of the components, the severity, the probability of the occurrence and the effect of failure in respect of the safety and security of the system into account. About 71% (15) of all potential failures were identified as 'catastrophic' or 'critical', 29% (6) as moderate and none as marginal failure. It can be noticed from Table 1, that potential causes of failures are software faults, erroneous HW/SW interfaces or missing services. Thus, the majority of the identified hazards and their causes correspond in the first place to software faults, insufficient reliability, performance and security. FMEA supports, similar to STPA, risk analysis. However FMEA fosters also preventive measures during the development of a product or when the system is in operation. The quality of a complex embedded system is monitored by the interpretation of the FMEA, to issue defect detection measures before going into operation. To support also an efficient maintenance of the product, the FMEA worksheet should be updated regularly.

5.2 Safety Analysis Using STPA

For hazard analysis using STPA the detailed control structure diagram of the system was acquired. Then, the first author of this study analyzed the forward collision avoidance system and identified inadequate control commands or events (for detail see [41]). Table 2 shows the inadequate control commands or events that could lead to hazardous states. During step 1 of STPA, 14 inadequate control commands or events have been identified in the forward collision avoidance system. Then, these control commands or events were analyzed, one by one, to identify their associated hazards. As one can see from Table 2, not provided control commands lead the system under consideration to hazardous states, in most cases of catastrophic level. Similarly, all identified control commands or events provided too late lead to, in most cases, hazardous states of catastrophic level. On the other hand, none of the events provided too early lead to catastrophic hazardous states; three lead to moderate and one to negligible level hazards.

It can be noted that one hazard can have more than one inadequate control action, e.g., hazard 2a in Table 2 exist because of vehicle sensor complex signal is not provided, provided unsafe and provided too late. For all these three inadequate control actions there is a single hazard.

From the identified 14 inadequate control commands or events, 22 hazards were identified. Table 3 shows the causal factors for all identified hazards in step 1 with their severity levels. The first column of Table 3 shows the identified hazards, the next column shows the severity levels, and the third column shows the causal factors for all hazards. The hazards were classified in three severity levels; catastrophic, moderate, and negligible. Over 70% (16) of all the hazards were classified as catastrophic with potentially fatal consequences. Only three hazards were classified as moderate severity level that may lead to severe accidents and have risk of serious injury. The remaining three hazards have negligible severity level. The negligible hazards do not have any serious consequences if the pertaining component fails alone and the other components of the system work properly. Therefore, based on the results of [41], it is possible to hypothesize that the STPA method efficiently supports risk analysts with limited domain experience (in our case maximum 5 years) in the identification of complete set of catastrophic hazards.

From Table 3, it can be noticed that the causal factors associated with component failures, communication errors, and software faults (dynamic behavior) were identified. Thus, the majority of the identified hazard and their causes correspond to the software faults of the studied system.

Table 2: Inadequate control actions/commands

No.	Command/Event	Not Provided	Provided Unsafe	Provided			Stopped Too Soon
				Too Early	Too Late	Out of Sequence	
1	Object Detection Signal	Catastrophic- System dysfunction [Collision] (1a)	Catastrophic- System malfunctioning (1b)	N/A	Catastrophic- System dysfunction [Collision] (1a)	N/A	N/A
2	Vehicle Complex Signal	Catastrophic- Problem in calculation of Vehicle status and collision probability (2a)	Catastrophic- Problem in calculation of Vehicle status and collision probability (2a)	N/A	Catastrophic- Problem in calculation of Vehicle status and collision probability (2a)	N/A	N/A
3	Collision Warning Signal	Negligible- (If every thing is working properly then the active safety will save from collision) (3a)	N/A	Negligible- (If every thing is working properly then the active safety will save from collision) (3a)	Negligible- (If every thing is working properly then the active safety will save from collision) (3a)	N/A	Negligible-(Warning will be stopped too soon that can cause accident. If everything works properly then the active safety will save collision) (3b)
4	System Reset Signal (Response from driver by using brakes)	Negligible- (If everything is working then the active safety will save from collision) (4a)	Negligible- (If everything is working then the active safety will save from collision) (4a)	N/A	Negligible- (If everything is working then the active safety will save from collision) (4a)	N/A	N/A
5	Vehicle Status Signal	Catastrophic- (Wrong brake pressure determination) (5a)	Catastrophic- (Wrong brake pressure determination) (5a)	N/A	Catastrophic- (Wrong brake pressure determination and decrease in reaction time) (5a)	N/A	N/A
6	Object Status Signal	Catastrophic- (Wrong brake pressure determination) (6a)	Catastrophic- (Wrong brake pressure determination) (6a)	N/A	Catastrophic- (Wrong brake pressure determination and decrease in reaction time) (6a)	N/A	N/A
7	Collision Assessment Signal	Catastrophic- System will not work [Collision] (7a)	Catastrophic- System will not work as intended [Collision] (7b)	Moderate- False signal due to system malfunctioning [Application of automatic brakes with out need] (7c)	Catastrophic- System will not work [Collision] (7a)	N/A	N/A
8	Reduce Torque	Moderate- Collision with divider, other things and vehicle can slip (8a)	N/A	N/A	Moderate- Collision with divider, other things and vehicle can slip (8a)	N/A	N/A
9	Brake Signal with Required Pressure	Catastrophic- System dysfunction [Collision] (9a)	Catastrophic- System malfunctioning [Collision] (9b)	Moderate- False signal due to system malfunctioning [Application of automatic brakes without need] (9c)	Catastrophic- System dysfunction [Collision] (9a)	N/A	N/A
10	Apply Brakes Signal	Catastrophic- System dysfunction [Collision] (10a)	N/A	Moderate- False signal due to system malfunctioning [Application of automatic brakes with out need] (10b)	Catastrophic- System dysfunction [Collision] (10a)	N/A	N/A
11	Accelerator Signal	Catastrophic- (Wrong brake pressure determination) (11a)	Catastrophic- (Wrong brake pressure determination) (11b)	N/A	Catastrophic- (Wrong brake pressure determination) (11a)	N/A	N/A
12	Change Transmission Signal	Catastrophic- Torque will not be reduced (12a)	N/A	N/A	Catastrophic- Torque will not be reduced (12a)	N/A	N/A
13	Limit air and fuel Supply Signal	Catastrophic- Torque will not be reduced (13a)	N/A	N/A	Catastrophic- Torque will not be reduced (13a)	N/A	N/A
14	Switch Off Engine Signal	Catastrophic- Torque will not be reduced (14a)	N/A	N/A	Catastrophic- Torque will not be reduced (14a)	N/A	N/A

Table 3: Causal factors of the identified hazards

No.	Step1 #	Hazards	Severity	Causal Factors
1	1a	System Dysfunction due to failure of Object detection system	Catastrophic	Object detection component failure (camera, radar or motion sensors) Communication error (no signal)
2	1b	Malfunctioning of the System due to Incorrect input from Object detection System	Catastrophic	Corrupted communication (wrong signal) and Malfunctioning of camera, radar and motion sensors Delayed communication (System will not work on time)
3	2a	Incorrect and missing calculation of Vehicle status and collision Probability due to Failure or malfunctioning of Vehicle Complex sensors	Catastrophic	Failure of vehicle sensors Communication error (no signal) Delayed communication (System will not work on time) Malfunctioning of sensors (Incorrect values sent by sensors)
4	3a	Missing collision warning signal - If rest of the System is working properly then the Active Safety will prevent from collision	Negligible	Inadequate collision assessment algorithm, Failure of warning indicator Malfunctioning of warning indicator, Incomplete controller process model Failure of collision estimator, Malfunctioning of collision estimator Incorrect vehicle or object status, Communication error (no signal) Delayed communication (System will not work on time)
5	3b	If Warning stopped too soon then it can cause accident-If everything else will work then the Active Safety will handle the situation	Negligible	Failure of warning indicator Malfunctioning of warning indicator Communication error
6	4a	Missing system reset signal can cause collision with divider or other objects due to unwanted auto braking	Negligible	Brake pedal sensor failure Communication error (no signal) Delayed communication (System will not reset on time and will apply brakes)
7	5a	Incorrect brake pressure determination due to missing vehicle status signal	Catastrophic	Failure of vehicle sensor complex (2a) Malfunctioning of collision controller due to incomplete process model Communication error (no signal) Delayed communication (System will not work on time)
8	6a	Incorrect brake pressure determination due to missing Object status signal	Catastrophic	Failure of Object detection (1a) Malfunctioning of collision controller due to incomplete process model Communication error (no signal) Delayed communication (System will not work on time)
9	7a	System Dysfunction due to missing collision assessment signal	Catastrophic	Component failures in Object detection and vehicle complex signal (1a and 2a) Failure of collision probability estimator Communication error (no signal) Delayed communication (System will not work on time)
10	7b	System will not work as intended due to unsafe (incorrect) Collision Assessment Signal	Catastrophic	Malfunctioning of Collision probability estimator Incorrect input by vehicle and object status providers Delayed communication (System will not work on time)
11	7c	Unwanted/Undesired auto braking due to False collision assessment signal	Moderate	Malfunctioning of Collision probability estimator Malfunctioning of collision controller due to incomplete process model
12	8a	Collision with the road divider, other things and also vehicle can slip due to Missing Reduce Torque signal	Moderate	Malfunctioning of brake controller due to incomplete process model (Incorrect brake pressure (safe brake pressure) will cause not to send reduce torque signal) Incorrect input by collision-assessment signal (7b) Communication error (no signal), Delayed communication (System will not work on time)
13	9a	System Dysfunction due to missing brake signal with appropriate (required) pressure	Catastrophic	Failure of brake Controller components Brake pressure determination fails, Communication error (no signal) Missing collision assessment signal, vehicle and object status signals
14	9b	System failure/malfunctioning as intended due to unsafe (incorrect) Brake signal	Catastrophic	Incomplete controller process model Malfunctioning of collision controller due to incomplete process model Delayed communication (System will not work on time)
15	9c	Unwanted/Undesired auto braking due to False Braking signal	Moderate	Malfunctioning of brake controller due to incomplete process model (Generation of false signal)
16	10a	System Dysfunction due to missing Apply Brakes signal	Catastrophic	Connection broken between brake pedal and brake actuator Failure of braking system Communication error (no signal)
17	10b	False signal due to brake system malfunctioning [Application of automatic brakes with out need]	Moderate	Malfunctioning of brake system (generation of false signal)
18	11a	Incorrect brake pressure determination due to missing Accelerator signal	Catastrophic	Sensor failure Communication error (no signal) Delayed communication (System will not work on time)
19	11b	System malfunctioning due to Missing Accelerator Signal	Catastrophic	Malfunctioning of sensor (Incorrect reading by sensor)
20	12a	Torque will not be reduced due to missing Change Transmission signal	Catastrophic	Component failure in the Torque Controller Missing reduce torque signal (8) Communication error (no signal) Delayed communication (System will not work on time)
21	13a	Torque will not be reduced due to missing limit air or/and fuel supply signal	Catastrophic	Component failure in the torque controller Malfunctioning of controller due to incorrect process model Missing reduce torque signal (8) Communication error (no signal) Delayed communication (System will not work on time)
22	14a	Torque will not be reduced due to missing Engine Switch off signal	Catastrophic	Component failure in the torque controller Malfunctioning of controller due to incorrect process model Missing reduce torque signal (8) Communication error (no signal) Delayed communication (System will not work on time)

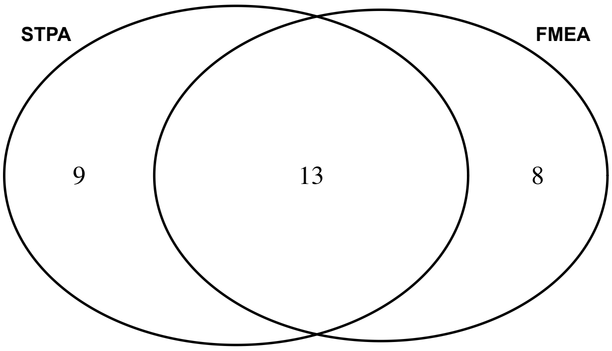


Figure 3: Number of common and distinct hazards identified by FMEA and STPA

6 Analysis

This section presents the analysis of the results that encompasses the main comparison results of the both methods, FMEA and STPA and their evaluation results based on the developed criteria.

6.1 Common and Distinct Hazards Identified by Both Methods

Table 4 shows the mapping of the hazards identified by both analysis methods. The identified hazards are represented in Table 4 by using their numbers used in Tables 1, 2, and 3. As it can be noticed from Table 1, the analysis by FMEA found 21 hazards. On the other hand, STPA found 22 hazards shown in Table 2 and Table 3. In total, both analysis methods found 30 unique or distinct hazards. As shown in Table 4, there are some hazards identified by STPA which are on a more abstract level compared to corresponding hazards identified by FMEA. For example, hazards 2a and 12a (identified by STPA) are mapped to two hazards each identified by FMEA. As it can be noticed from Table 4, there are some identified hazards that are only identified by one analysis method, either FMEA or STPA. Table 4 and Figure 3 also show 13 common hazards identified by both analysis methods.

6.2 Classification of The Identified Hazards

The identified hazards are classified into the following five error categories.

- Component interaction error
- Software error

Table 4: Mapping of the identified hazards

No.	Hazards identified by STPA	Hazards identified by FMEA
1	1a	6
2	1b	Not identified
3	2a	1 & 2
4	3a	Not identified
5	3b	3
6	4a	12
7	5a	13
8	6a	14
9	7a	Not identified
10	7b	Not identified
11	7c	Not identified
12	8a	18
13	9a	Not identified
14	9b	Not identified
15	9c	Not identified
16	10a	20
17	10b	Not identified
18	11a	8
19	11b	15
20	12a	11 & 17
21	13a	Not identified
22	14a	Not identified
23	Not identified	4
24	Not identified	5
25	Not identified	7
26	Not identified	9
27	Not identified	10
28	Not identified	16
29	Not identified	19
30	Not identified	21

- Human error
- Component error
- System error

These categories were selected because it is claimed that the STPA method identifies these types of hazards [10, 11, 19]. According to [19] and [10, 11] STPA can find more component interaction, software and human hazards than traditional methods.

Figure 4 shows three bar plots showing classification for the common and distinct identified hazards by the both methods. The first bar plot shows the hazards only identified by the STPA method. The second bar plot shows the hazards only identified by the FMEA method. Finally, the third bar plot shows the classification of the hazards identified by the both methods (common hazards).

Each bar plot shows the percentage of the hazards that were classified in each way. For example, for hazards only identified by STPA, 18% were classified as component interaction hazards, 72% as software hazards, 54% as component failure, and 18% as system hazards. It can be noticed that the percentage of the classified hazards in the classification exceeds 100% since some hazards are classified in more than one category.

Just by looking at the "distributions" of hazard types in the three different cases, it is not possible to clearly find any major differences.

The second bar plot in Figure 4 shows the classification of the hazards only identified by FMEA. 62% as software hazards, 75% as component failure hazards, and 25% as system hazards. As it can be noticed in Figure 4, FMEA did not find any unique hazard of component interaction and human error type that is not identified by the STPA method. Here, one interesting result is that FMEA identified as many software error type hazards as STPA. It should be noted that the data points in this study are few and the focus of the study is not on quantitative comparison of the methods. However, as noted above, there is almost no difference regarding the identified software error type hazards by both methods. One positive result in favor of STPA, based on the experience of the authors of this study, is that it identified clear software error type hazards because of its keywords ('provided', 'not provided', etc.), which make it simple and easy to identify software error type hazards.

Finally, the third bar plot in Figure 4 shows classification of the common identified hazards (identified by both the FMEA and the STPA methods). 27% were classified as component interaction hazards, 72% as software hazards, 81% as component failure, and 18% as system hazards.

There are no common identified hazards of human error type. Apparently, none of the methods could find any human error type hazard in this study. The reason for this can be that the analyzed system does not involve much human input or interaction.

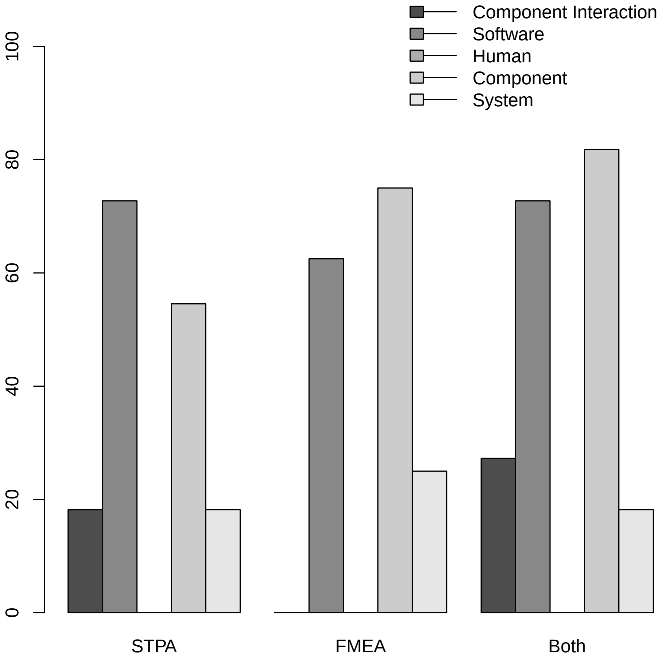


Figure 4: Classification of the identified hazards

6.3 Comparison of The Causal Factors of The Identified Hazards

This section presents the comparison of the common hazards identified by the both hazard analysis methods, FMEA and STPA. As shown in Table 4, 13 hazards are identified as common hazards. There is a clear difference in the identified causal factors by the two hazard analysis methods. For example, the causes identified by STPA are more detailed and cover more aspects (see Table 3). Furthermore, as one can see in Table 3, the potential causes identified by STPA cover hardware failures, communication errors including delayed communication, and software errors. However, FMEA did not find potential causes in detail. However, it found the causes that also cover hardware and software errors like 'Architecture erroneous or software failure' in No. 9 (see Table 1) and the potential causes are detailed enough to assess the probability of a failure.

The main reason behind the detailed identification of potential causes by STPA is the used keywords during analysis such as, 'provided', 'not provided', 'provided unsafe', etc. (see step 1 of STPA in Table 2). The keywords used in the STPA analysis help to identify detailed potential causes, in particular they help to find communication error causes. Based on this interpretation it can be concluded that STPA covers more component interaction hazard causes. Here, it should be noted that FMEA also found communication error type causes, but as compared to STPA they are not identified for all hazards and also are not detailed.

The findings of this study regarding the identification of causal factors of the identified hazards corroborates with the findings of [10, 11], who have compared, qualitatively, STPA with bottom-up and other top-down analysis techniques. According to [10, 11], STPA can find more types of causes than traditional methods, and STPA has a structured process to follow in doing the analysis that is a likely reason to result in a more complete result.

6.4 Mapping of The Analysis Steps of FMEA and STPA

This section presents the comparison of the analysis process of both methods. For this purpose, steps of both methods are mapped to each other in Table 5 to find the common steps. Then, the mapped common steps of both methods are compared based on the qualitative criteria derived from TAM. Here, it should be noted that the output of the mapping performed in this section is further used to compare the analysis process of both methods to yield evaluation results presented in Section 6.5.

As it can be noticed in Table 5, step 1 for the both methods (FMEA and STPA) is mapped as a same step called Map-A. In step 1 of FMEA, decomposition of the system is performed and on the other hand in STPA step 1 acquisition of the system's functional control diagram along with its safety requirements is performed. Moreover, step 1 of STPA demands for a high-level system hazards identification.

Table 5: Mapping of the analysis steps for FMEA and STPA

FMEA	**STPA**	**Mapping comments**
Step 1: Decomposition of the system to be analysed into subsystems and components	**Step 1**: Acquisition of functional control diagram of the system to be analysed as a whole, and identification of some high-level system hazards to start with	**Map-A**: Step 1 of both methods are mapped as a same step in the analysis process because FMEA based on reliability theory (decomposition required) and STPA is based on system theory (system required as a whole)
Step 2: Assigning the application function to each subcomponent and subsystem	N/A	**Map-B**:This step of FMEA does not map to any STPA step
Step 3: Determine and analyse the • Potential failure modes • Causes of failure • Failure effects that can lead system to a hazardous state	**Step 2**: Identify the potential inadequate control commands or events (potential hazards) **Step 3**: Determine how each potential hazardous control action (potential hazards) identified in Step 2 could occur (causal factors of identified potential hazards)	**Map-C**: Step 3 of FMEA is mapped to step 2 and step 3 of STPA, which consists of identification of potential failures (or hazards), their causes and effects
Step 4: Evaluate risk and calculate risk priority number (RPN)	N/A	**Map-D**: This tep of FMEA does not map to any STPA step
Step 5: Specify defect avoidance or risk mitigation measures	**Step 4**: Design controls and countermeasures if they do not already exist or evaluate existing	**Map-E**: Step 5 of FMEA and step 4 of STPA are mapped to each other because they are both about designing and evaluating countermeasures

We mapped step 1 of FMEA and STPA as a same step because it is an initiating step for the analysis process in both methods.

Further, step 2 of FMEA shown in Table 5 does not correspond to any step of STPA in the analysis process that is about assigning the application function to each sub-component and sub-system. Here, it is interesting to see that STPA performs this task (task of step 2 of FMEA) in its step 2 but without making it explicit. In step 2 of STPA, identification of all control commands or events of a system is performed that is more or less same as assigning the application function in step 2 of FMEA. Here, step 2 of FMEA can be mapped to identifying system's control commands or events activity in STPA but STPA's guideline does not distinguish it explicitly from other steps. In this study, the authors do not intend to modify the existing methods for sake of mapping or any other research activity instead they evaluate the analysis process of both methods based on how the methods are developed and presented along with their guidelines and application instructions.

Then, step 3 of FMEA is mapped to step 2 and 3 of STPA (Map-C) that is about identifying hazards and their causes and effects. In FMEA, identifying hazards (failure modes) and their causes and effects is carried out in a single step (step 3). Nevertheless, in STPA identifying hazards and their causes is carried out in two steps (step 2 and 3).

After this, step 4 of FMEA that is about calculating risk priority number does not map to any STPA step. Here, this step gives some estimation, mostly quantitative, about the identified failure modes' severity and then based on this their prioritization is carried out. On the other hand, STPA does not have any step that deals with identified hazards' severity and their prioritization.

Finally, step 5 of FMEA is mapped to step 4 of STPA (Map-E) that is about designing and evaluating risk mitigation measures. These steps of the analysis process that deal with countermeasures for identified hazards are exactly the same in both methods.

6.5 Evaluation of The Analysis Process of FMEA and STPA

In this section, the analysis process of the FMEA and STPA methods are compared based on the Technology Acceptance Model (TAM) [4, 5]. TAM is used to investigate how users accept and use new technologies for information systems and has also successfully been applied to assess risk analysis and treatment processes [30]. TAM was originally proposed for information systems (IS) acceptance and usage. TAM uses 'perceived usefulness' and 'perceived ease of use' that estimates the beliefs in information technology acceptance and usage. The perceived usefulness means that the degree to which a person believes that using a particular information system would enhance his or her job performance. Furthermore, the perceived ease of use means that the degree to which a person believes that using a particular system would be free of effort [4, 5].

Table 6: Assessment criteria for STPA and FMEA derived from the Technology Acceptance Model (TAM) [4, 5]

TAM constructs	Derived qualitative criteria
Perceived ease of use	- How easy or hard - Why was it easy or hard
Perceived usefulness	- Provided support by the method - Confidence about the results - Applicability for software

In this study, the criteria that we defined for evaluation of the analysis process of FMEA and STPA are derived and inspired by TAM as shown in Table 6. Here, it should be noted that the criteria was defined in a meeting after having discussion among all the authors and it is originally based on TAM.

The qualitative criteria are defined as follows:

1. **How easy or hard**: This criterion is used to assess how easy or hard it is to apply a specific step of the analysis process. For this criterion a five point Likert scale with values ‘very easy’, ‘easy’, ‘moderate’, ‘hard’, and ‘very hard’ is used.

2. **Why was it easy or hard**: This criterion is used to assess if a method’s particular step was easy or hard to apply then why it was easy or hard. This criterion is a follow-up criterion linked with previous criterion.

3. **Support by method**: This criterion assesses how much support is provided by method, i.e., method application guidelines and support for analysis by method itself. Here, the provided support by a method affects its effectiveness, i.e., how well a method performs in the analysis process. This support can be provided by guidelines to carry out analysis, tools to apply method or any thing that helps analysts to carry out analysis.

4. **Confidence about the results**: This criterion is used to assess the confidence of the risk analysts about the carried out risk analysis and its results. If a used method is good enough then the performed analysis by applying that method will yield some degree of confidence in analysts about the method application and yielded results.

5. **Applicability for software**: This final criterion used to assess how applicable a particular step of a method is for identifying software hazards and their

> causes. Today, almost every system has software in it that makes the system to behave dynamically. An effective analysis method must identify problems pertaining to the dynamic behavior of a system, i.e., software relevant hazards and their causes.

Table 7 and 8 shows the evaluation of the analysis process of FMEA and STPA based on the afore-mentioned criteria. Both methods were evaluated by using the afore-mentioned derived qualitative criteria.

For example, step 1 of FMEA, shown in Table 7, was easy to apply because of available detailed requirements and architecture of the system. Then, the provided support by FMEA in step 1 was the structural decomposition. After this, the analyst is confident about the results of application of step 1 of the FMEA method. Finally, step 1 was evaluated for the applicability for software and in this case it is well suited because it fosters risk-based development and testing.

On the other hand, step 1 of STPA was easy to apply because of the available detailed functional control diagram and safety requirements and constraints of the system. After this, the STPA method provides explicit support in this step and the analyst is confident about the analysis results. Finally, this step is well suited for software because the design of STPA is only for software that can easily be seen in the analysis. In this way, all steps of FMEA and STPA are evaluated for these qualitative criteria that are shown in Table 7 and 8, respectively.

To summarize, the FMEA analysis process defined in Table 7 consists of five steps. Step 1 and Step 2 are easy to perform, because of the bottom up analysis of the system. However, experience in the development of dependable systems is needed to identify failure modes (step 3) and its potential risk (step 4). Moreover, the introduction of defect prevention measures (step 5) in product development is a common task in every project. On the other hand, the STPA analysis process defined in Table 8 consists of four steps. Step 1 and 2 are easy to apply because of the available detailed information about the system and the STPA keywords used to identify inadequate controls in the system. Moreover, step 3 is hard to perform because it finds causal factors in large amount that can be challenging sometimes. Finally, step 4 is also hard to carry out because there is no explicit support provided by the method. STPA is a simple method that does not require high level of experience by the analysts to apply the method.

7 Validity Evaluation

The validity of a study represents the trustworthiness of its results, which means for example that the results are not biased by the researcher's own opinion or point of view [35]. Validity of this kind of study (a case study) can be assessed regarding construct validity, internal validity, external validity, and reliability [35,43]

Construct validity considers the studied artifacts and concerns if they represent what the researcher have in mind and also if the studied artifacts are investigated

Table 7: Analysis of the FMEA process for safety analysis

FMEA steps	**How easy or hard?**	**Why was it easy or hard?**	**Provided support by the method**	**Confidence about the analysis results**	**Applicability for Software**
Step 1	Easy	Requirements and architecture of the system on an abstract level are well defined	Structural decomposition is supported	Experience in the application of FMEA in safety-critical systems such as railway interlocking systems was the basis	Very well suited for software, because for instance risk-based development and testing is fostered
Step 2	Easy	Functions of the systems are defined	Supported by templates	Method is easy to apply	Yes, on the basis of the requirements
Step 3	Moderate	The identification of failure causes may be challenging	Yes, taking domain specific failure data into account	Confident, because a potential failure can be assigned to each task of a component	Yes, on the basis of requirements and design specification of software systems
Step 4	Hard	It is not easy to assign the potential risk to avoid a risk scenario	Yes, by assessing the complexity of a component and the probability of a failure	Taking qualitative interpretation of the RPN into account, gives confidence	The method fosters the application of risk-based testing in software development
Step 5	Moderate	Efficiency of the measures have to be assessed	Yes	20 years experience in industry	The application of FMEA fosters the improvement of the software development process

Table 8: Analysis of the STPA process for safety analysis

STPA steps	**How easy or hard?**	**Why was it easy or hard?**	**Provided support by the method**	**Confidence about the analysis results**	**Applicability for Software**
Step 1	Easy	The functional control diagram and requirements of system with its safety constraints are available in detail	Method does not explicitly support in this step instead it requires detailed functional control diagram and other system descriptions	Confident about the results of this step based on the reviewed literature about STPA and by studying advanced level safety course	Very well suited for software because the main focus of STPA is on dynamic behavior of systems, which covers mainly the software part
Step 2	Easy	Identification of inadequate safety controls is easy because of the STPA keywords, i.e., not provided, provided unsafe, provided too late or early, and stopped too soon	Systematic approach by using STPA keywords identified almost complete set of potential hazards	Confident because all components in system's functional control diagram are one by one evaluated against the keywords to find complete set of hazards	Very well suited as the main focus of STPA is on software and the dynamic behavior of system. It identifies majority of the hazards relevant to software
Step 3	Hard	Identification of causal factors can be challenging	Keywords to evaluate system's dynamic deviation from required safety	Confident, because STPA yielded almost a complete analysis result for both the potential hazards and their causal factors	Very well suited as it identified majority of the software relevant causal factors
Step 4	Hard	Designing new countermeasures and evaluating existing ones can be difficult or challenging	No explicit support by the method	Researcher in safety domain having 5 years of research experience in analyzing methods and tools used for the analysis of safety critical systems	It identifies problems in software and suggest improvements depending on the stage, i.e., design, development and operation

according to the research questions of the study. In this study, the collision avoidance system was analyzed to identify hazards. The analysis was done by two persons, the first and the second author of this study. The hazard analysis performed by the first author is already published in a previous paper [41]. In the current study, the second author analyzed the system using FMEA and all the documentation and information were available, which were used during the hazard analysis performed by the first author. There could be a risk of not understanding the analyzed system and its description by the authors. To decrease this threat a simple system was selected and also its detailed description was acquired and made available to all the authors of this study. Moreover, there could be another risk of not understanding the investigated methods by the authors. To decrease this threat the experienced persons were selected to apply methods and also the suitable method guidelines and instructions were followed during the analyses. Also, regular meetings were carried out to eliminate any existing ambiguities in understanding of the system and its description and investigated methods. This could also impact the evaluation of the analysis process. The effect was however limited by dividing the analysis into steps that were assumed to be known and understood.

Internal validity is important and mostly applicable in studies of causal relationships. In this study, there can be a chance of *history internal validity* threat. To decrease the chance of history threat the following measure were taken. The second author of this study was selected to apply FMEA on the collision avoidance system. The first author already knows the existing hazards in the selected system because he has applied STPA on the selected system in the previous study [41]. Therefore, in the current study, to improve research validity it was decided that the first author would not apply the FMEA method on the selected system. Instead another author did that. The second author of this study did not have access or review the previous study results [41]. Furthermore, it was also considered that the same system information and description is available to the second author to analyze the system.

All stages of risk and hazard analysis process involve subjectivity [33]. There is always a chance of uncertainty, the need for judgment, considerable scope for human bias, and inaccuracy. It is highly likely that the results obtained by one risk analyst are not same to the results obtained by other risk analysts starting with the same information [33]. In our case study both the authors (first and second) analyzed the collision avoidance system independently by applying different hazard analysis methods (STPA and FMEA). Moreover, both the authors have sufficient level of experience of analyzing safety critical systems and it is believed that in this case study there is a little or no chance of this threat. Here, in this study the objective is not to compare hazard analysis methods based on just the numbers of identified hazards instead the objective is to compare them based on the types of hazards found.

During the study the results, and the written formulations of the results, were studied and discussed by all authors in order to limit the risk that results from one method was treated and formulated more positively than the results of the other.

External validity is concerned with to what extent it is possible to generalize the findings, and to what extent the findings are of interest to people outside the investigated case. In this study, the selected system is a real software-intensive safety critical system, therefore it is believed that the results of this study will be applicable and helpful in analysis of such type of safety critical systems. Moreover, the results of this study can be further used to compare different analysis methods using other safety critical systems. Furthermore, there might be a threat of difference in results of both hazard analyses because of different level of experience of the first and second authors of this study who performed hazard analyses.

Reliability is concerned with to what extent the data and the analysis are dependent on the specific researchers. The reliability was addressed by conducting both the data collection and analysis as a group of researchers instead of one single researcher. In this study there are less chances of this threat because the data used for the analysis is of third degree [17], e.g., documentation, description, published literature. Moreover, the first-degree data collected in this study is the hazard analysis results or identified hazards. To decrease the chances of reliability threats, guidelines for the both methods were properly used. Special measures were taken during the hazard analysis process and continuously reviewed by the co-authors. For example, the first author of this study performed an initial comparison of the identified hazards yielded from the FMEA analysis and STPA analysis. The first author created a list of the common and unique hazards that were identified by the both analysis methods. After this the second author reviewed the initial comparison performed by the first author. The second author identified one more hazard (no. 18 in Table 4) as common hazard identified by the both analysis methods. Then, the first author classified all the identified hazards into five error categories, component interaction error, software error, human errors, component error, and system error. This classification was also reviewed by the second author of this study for the researcher triangulation.

8 Discussion

The common identified hazards are classified as software error and component error type mainly, as shown in Figure 4. Moreover, there are some common identified hazards classified as component interaction types hazards. From the 13 common identified hazards it can be observed that both methods found software error type hazards covering the dynamic behavior of the system. In [10,11,16,19,26,42], the authors have mentioned that the traditional analysis methods (FMEA, FTA, etc.) cannot identify software errors. However, FMEA is still used in many safety critical hardware software systems and was extended to detect software hazards [22] as well as vulnerabilities [36], and has even been applied for security testing [27]. Another difference to STPA is that it does not start from an undesired state but from a malfunctioning hardware or software component. However, in our case study there is no major difference between the types of identified hazards by the both applied methods on the collision avoidance system.

Furthermore, both methods also found some hazards that are unique to them (identified by only one method, either FMEA or STPA). STPA identified 9 unique hazards that are not identified by FMEA of which majority of the identified hazards are of software and component failure type hazards. On contrary, FMEA identified 8 unique hazards that were not identified by STPA. Interestingly, the majority of the uniquely identified hazards by FMEA are also of software and component failure hazards like STPA.

Moreover, a small difference can be noticed in the unique identified hazards by the two analysis methods regarding component failure type hazards. That means, FMEA identified more component failure hazards as compared to STPA. This shows the basic philosophy behind both methods, FMEA focuses more on components, their failures and risk mitigation measures, whereas STPA focuses on delivery of control commands and their feedbacks.

One more interesting factor in our case study is that STPA found fewer unique system error type hazards than FMEA. Because STPA is developed considering system engineering and thinking, which consider whole system as a single unit instead splitting it in several parts. One potential reason of finding few system type hazards by STPA can be that the analyzed system in this study does not have many system type hazards. On the other hand, FMEA identified 2 out of 8 hazards of system error type. Another interesting difference can be observed regarding the component interaction error type hazards, STPA identified 18% hazards of component interaction error type. On the other hand, FMEA did not identify any hazard of component interaction error type. This result corroborates with the results of the previous studies.

[10, 11] mention that the main difference of STPA from bottom-up analysis methods like FMEA is that bottom-up analysis techniques start by identifying all possible failures. This can result in a very long list of potential failures if there are a lot of components to consider in the analysis. However, this long list is produced because FMEA takes the architecture and complexity of components into account [38]. Moreover, this long list of potential failures can be managed by introducing a hierarchical structure in FMEA. Furthermore, FMEA fosters propositions for the structure of a hardware and software system and generates preventive measures during development and operating [38]. However, in our case study both methods were applied independently by the different authors on a collision avoidance system to find operational hazards of the system that yielded in almost the same number of identified hazards (21 by FMEA and 22 by STPA).

However, one clear difference where STPA seems to outperform FMEA is finding causal factors of identified hazards. According to [10, 11], STPA considers more types of hazard causes than the other traditional hazard analysis methods. Therefore, STPA is more complete than existing traditional hazard analysis methods [10]. In our case study, the results corroborate with the findings of [10, 11] regarding STPA's complete causal factors identification.

On the other hand, FMEA (which is based on reliability theory) is stronger with respect to risk assessment of software failures by calculation of a risk priority number based on the complexity of a component or system. In STPA (which is based on system theory) there is no corresponding process step. Also assigning the application function to each sub-component and subsystem is not covered in STPA. However, the steps of (a) system decomposition and acquisition, respectively, (b) identification of potential failures, their causes and effects, as well as (c) definition of countermeasures map to each other. Especially, the definition of countermeasures is according to the technology acceptance model hard to perform and requires experience.

9 Conclusions and Future Work

In this paper, we present a qualitative comparison of the two hazard analysis methods, Failure Mode and Effect Analysis (FMEA) and System Theoretic Process Analysis (STPA) using case study research methodology. Both methods have been applied on the same forward collision avoidance system to compare the effectiveness of FMEA and STPA. Moreover, the analysis process of both methods is also evaluated by applying a qualitative criteria derived from Technology Acceptance Model (TAM) [4, 5].

It can be observed that almost all types of hazards that were identified in the study were found by both methods. That is, both methods found hazards classified as component interaction, software, component failure and system type. With regard to component failure hazards, FMEA identified more component failure hazards than STPA. With regard to software error type hazards, STPA found more hazards than FMEA of unique hazards. With regard to component interaction error type hazards, STPA found some hazards however FMEA did not find any of unique hazards. Finally, with regard to system type error hazards, FMEA found slightly more hazards than STPA.

Both FMEA and STPA consider system decomposition (FMEA decomposes and STPA considers whole system for analysis), identification of potential failures, their causes and effects, as well as definition of countermeasures. But STPA does not consider risk assessment in terms of risk priority number calculation and assignment of the application function to each subsystem.

The methods have different focuses. FMEA especially takes the architecture and complexity of components into account, whereas STPA is stronger in finding causal factors of identified hazards.

It can be concluded that, in this study, there were no type of hazard that was not found by any of the methods, which means that it is not possible to point out any significant difference in the types of hazards found. However, it can be observed that none of the methods in this study was effective enough to find all identified hazards, which means that they complemented each other well in this study. Further research, especially in terms of case studies and experiments, is needed in order to investigate differences, but also combinations of the methods and possible extensions of them. In addition, safety has been defined as an important risk driver for testing [9], but the number of risk-based testing approaches taking safety analysis into account is limited [6]. Comparing different safety analysis methods like FMEA and STPA with respect to test planning, design, execution and evaluation is another suggested topic for further research that could help to increase adoption of safety analysis methods for risk-based testing.

Bibliography

[1] J. C. Becker and G. Flick, "A practical approach to failure mode, effects and criticality analysis (FMECA) for computing systems," in *Proceedings of the IEEE High-Assurance Systems Engineering Workshop*, Oct 1996, pp. 228–236.

[2] Bond et al., "Collision mitigation by braking system," *US Patent 6607255B2*, 2003.

[3] E. Coelingh, A. Eidehall, and M. Bengtsson, "Collision warning with full auto brake and pedestrian detection - a practical example of automatic emergency braking," in *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2010, pp. 155–160.

[4] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989.

[5] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: A comparison of two theoretical models," *Management Science*, vol. 35, no. 8, pp. 982–1003, 1989.

[6] G. Erdogan, Y. Li, R. K. Runde, F. Seehusen, and K. Stølen, "Approaches for the combined use of risk analysis and testing: a systematic literature review," *International Journal on Software Tools for Technology Transfer*, vol. 16, no. 5, pp. 627–642, 2014.

[7] C. A. Ericson, "Fault Tree Analysis - A History," in *proceedings of The 17:th International System Safety Conference*, 1999.

[8] C. F. Eubanks, S. Kmenta, and K. Ishii, "Advanced failure modes and effects analysis using behavior modeling," in *Proceedings of Design Engineering Technical Conferences and Design Theory and Methodology Conference, Sacramento, California, USA*, 1997.

[9] M. Felderer and I. Schieferdecker, "A taxonomy of risk-based testing," *International Journal on Software Tools for Technology Transfer*, vol. 16, no. 5, pp. 559–568, 2014.

[10] C. H. Fleming, M. Spencer, N. G. Leveson, and C. Wilkinson, "Safety assurance in NextGen," NASA/CR-2012-217553, Tech. Rep., 2012.

[11] C. H. Fleming, M. Spencer, J. Thomas, N. Leveson, and C. Wilkinson, "Safety assurance in NextGen and complex transportation systems," *Safety Science*, vol. 55, pp. 173–187, 2013.

[12] L. Grunske, R. Colvin, and K. Winter, “Probabilistic model-checking support for FMEA,” in *Proceedings of the 4:th International Conference on the Quantitative Evaluation of Systems, QEST 2007*, Sept 2007, pp. 119–128.

[13] C. M. Haissig and R. Brandao, “Using TCAS Surveillance to Enable Legacy ADS-B Transponder Use for In-trail Procedures,” in *Proceedings of the 31st Digital Avionics Systems Conference (DASC), Williamsburg, Virginia, USA*, 2012, pp. 5D5:1–5D5:114.

[14] IEC 60812:2006, “Analysis techniques for system reliability - procedure for failure mode and effects analysis (FMEA),” 2006. [Online]. Available: http://www.iec.chAugust2014

[15] T. Ishimatsu, N. G. Leveson, J. Thomas, M. Katahira, Y. Miyamoto, and H. Nakao, “Modeling and hazard analysis using STPA,” in *proceedings of the International Conference on Association for the Advancement of Space Safety*. NASA, Sep. 2010.

[16] T. Ishimatsu, N. G. Leveson, J. P. Thomas, C. H. Fleming, M. Katahira, Y. Miyamoto, R. Ujiie, H. Nakao, and N. Hoshino, “Hazard analysis of complex spacecraft using Systems-Theoretic Process Analysis,” *Journal of Spacecraft and Rockets*, vol. 51, no. 2, pp. 509–522, 2014.

[17] T. C. Lethbridge, S. E. Sim, and J. Singer, “Studying software engineers: Data collection techniques for software field studies,” *Empirical Software Engineering*, vol. 10, no. 3, pp. 311–341, 2005.

[18] N. G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*. The MIT Press, 2012.

[19] N. G. Leveson, C. H. Fleming, M. Spencer, J. Thomas, and C. Wilkinson, “Safety assessment of complex, software-intensive systems,” *SAE International Journal of Aerospace*, vol. 5, no. 1, 2012.

[20] M. O. Mackel, “Mit blick auf’s risiko. software-fmea im entwicklungsprozess softwareintensiver technischer systeme,” *Qualität und Zuverlässigkeit*, vol. 1, no. 46, pp. 65–68, 2001.

[21] J. A. McDermid, M. Nicholson, D. J. Pumfrey, and P. Fenelon, “Experience with the application of HAZOP to computer-based systems,” in *proceedings of the 10:th Annual Conference on Computer Assurance, 1995. COMPASS ’95, Systems Integrity, Software Safety and Process Security*, 1995, pp. 37–48.

[22] M. McDonald, R. Musson, and R. Smith, *The Practical Guide to Defect Prevention*. Microsoft Press, 2008.

[23] G. Menkhaus and B. Andrich, "Metric suite directing the failure mode analysis of embedded software systems." in *ICEIS (3)*, 2005, pp. 266–273.

[24] R. J. Mikulak, R. McDermott, and M. Beauregard, *The Basics of FMEA*. Productivity Press, paper back, 2008.

[25] MIL-STD-1629A, "Procedures for performing a Failure Mode, Effects and criticality Analysis, U.S. Department of Defense," 1980.

[26] H. Nakao, M. Katahira, Y. Miyamoto, and N. G. Leveson, "Safety guided design of crew return vehicle in concept design phase using STAMP/STPA," in *proceedings of the 5:th International Association for the Advancement of Space Safety (IAASS) Conference*, 2011, pp. 497–501.

[27] B. Peischl, M. Felderer, and A. Beer, "Testing security requirements with non-experts: Approaches and empirical investigations," in *IEEE International Conference on Software Quality, Reliability and Security (QRS)*. IEEE, 2016, pp. 254–261.

[28] S. J. Pereira, G. Lee, and J. Howard, "A system-theoretic hazard analysis methodology for a non-advocate safety assessment of the ballistic missile defense system," in *Proceedings of the AIAA Missile Sciences Conference, Monterey, California*, 2006.

[29] K. H. Pries, "Failure mode & effects analysis in software development," SAE Technical Paper, Tech. Rep., 1998.

[30] R. Ramler and M. Felderer, "A process for risk-based test strategy development and its industrial evaluation," in *International Conference on Product-Focused Software Process Improvement*. Springer, 2015, pp. 355–371.

[31] C. Raspotnig and A. Opdahl, "Comparing risk identification techniques for safety and security requirements," *Journal of Systems and Software*, vol. 86, no. 4, pp. 1124–1151, 2013.

[32] F. Redmill, M. Chudleigh, and J. Catmur, *System Safety : HAZOP and Software HAZOP*. John Wiley & Sons, 1999.

[33] F. Redmill, "Risk analysis - a subjective process," *Engineering Management Journal*, vol. 12, pp. 91–96(5), April 2002.

[34] RTCA/DO-312, "Safety, Performance, and Interoperability Requirements Document for the In-Trail Procedure in Oceanic Airspace (ATSA-ITP) Application," RTCA Incorporate, Washington DC, Tech. Rep., 2008.

[35] P. Runeson, M. Höst, A. Rainer, and B. Regnell, *Case Study Research in Software Engineering: Guidelines and Examples*. Wiley, 2012.

[36] C. Schmittner, T. Gruber, P. Puschner, and E. Schoitsch, "Security application of failure mode and effect analysis (fmea)," in *Computer Safety, Reliability, and Security*. Springer, 2014, pp. 310–325.

[37] G. Sindre and A. L. Opdahl, "Eliciting security requirements with misuse cases," *Requirements Engineering*, vol. 10, no. 1, pp. 34–44, 2005.

[38] J. J. Stadler and N. J. Seidl, "Software Failure Modes and Effects Analysis," in *Reliability and Maintainability Symposium (RAMS), 2013 Proceedings-Annual*. IEEE, 2013, pp. 1–5.

[39] T. Stålhane and G. Sindre, "A comparison of two approaches to safety analysis based on use cases," in *Conceptual Modeling - ER 2007*, ser. Lecture Notes in Computer Science, vol. 4801. Springer, 2007, pp. 423–437.

[40] S. M. Sulaman, K. Weyns, and M. Höst, "A review of research on risk analysis methods for it systems," in *proceedings of the 17:th International Conference on Evaluation and Assessment in Software Engineering (EASE '13)*. ACM, 2013, pp. 86–96.

[41] S. M. Sulaman, T. Abbas, K. Wnuk, and M. Höst, "Hazard analysis of collision avoidance system using STPA," in *International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, 2014, pp. 424 – 428.

[42] J. Thomas and N. G. Leveson, "Performing hazard analysis on complex, software and human-intensive systems," in *proceedings of the 29:th ISSC Conference about System Safety*, 2011.

[43] R. K. Yin, *Case Study Research: Design and Methods*. SAGE Publications, 2003.

[44] S. Yu, Q. Yang, J. Liu, and M. Pan, "A comparison of FMEA, AFMEA and FTA," in *Proceedings of the 9th International Conference on Reliability, Maintainability and Safety (ICRMS)*, June 2011, pp. 954–960.

Paper V

Identification of IT Incidents for Improved Risk Analysis by Using Machine Learning

Abstract

Today almost every system or service is dependent on IT systems, and failure of these systems have serious and negative effects on the society. IT incidents are critical for the society as they can stop the function of critical systems and services. Therefore, it is important to analyze these systems for potential risks before becoming dependent on them. Moreover, in a software engineering context risk analysis is an important activity for the development and operation of safe software-intensive systems. However, the increased complexity and size of software-intensive systems put additional requirements on the effectiveness of the risk analysis process. This means that the risk analysis process needs to be improved and it is believed that this can be done by having an overview of already occurred IT incidents. This study investigates how difficult it is to find relevant risks from available sources and the effort required to set up such a system. It also investigates the accuracy of the found risks. In this study 58% of texts that potentially can contain information about IT incidents were correctly identified from an experiment dataset by using the presented method. It is concluded that the identifying texts about IT incidents with automated methods like the one presented in this study is possible, but it requires some effort to set up.

Sardar Muhammad Sulaman, Kim Weyns and Martin Höst,
In Proceedings of the 41:st Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pages 369- 373, 2015.

1 Introduction

Both researchers and practitioners often talk about "IT incidents" that either have happened or may happen in the future. This can either be incidents with critical IT services, or incidents with IT systems that support other critical functions in our society. Risk analysis and management are important activities for the safe development of modern software-intensive systems, since there are a number of factors that make these systems increasingly complex and critical. The factors effecting modern software-intensive systems are the fast changing technology, limited ability to learn from experiences, increasing complexity and coupling, more complex relationships between humans and automation, changing regulatory and public view of safety, and increasing dependability on such systems [7]. Risk analysis and management are important activities for most of the project management tasks. In a software engineering context, these activities focus on the software development process and ensure its integrity. The activities try to ensure there are no or little unforeseen negative impacts on the software development project. At the very least, they help to keep all identified potential risks under the effective management control [9]. There exist many, low and high level, risk analysis methods and frameworks that complements in identification and management of risks [12]. By using these methods and frameworks it is possible to foresee the potential consequences of future possible IT incidents that later can be decreased or mitigated. Risk analysis includes a step where potential risks are identified, e.g. through "brain storming" activities. In this step it is assumed to be valuable to understand what IT incidents that have already occurred. This means that information about already happened IT incidents can be used as an input to risk analysis and management processes. However, historical data about such unwanted events is not easily accessible and it is not available at a single place. Therefore, there is a need for an intelligent system that automatically identifies already happened IT incidents and then saves them in a database.

In this paper we discuss and evaluate an approach for automatically collecting information about IT incidents from online news sources. In online news sources there are texts available that includes relevant incidents. A system for automatic identification should find as many relevant articles as possible, but it should falsely identify as few now relevant articles as relevant as possible.

The approach is general and could be used to collect information about other topics, but the approach is of special interest in relation to IT incidents, as much information about them is available online and although they are a critical infrastructure, there is less coordination in the collection of information about incidents than is the case for, for example, nuclear or aviation incidents.

This means that a basic assumption underlying the research conducted in this paper is that there is information available about IT incidents in texts available on for example Internet news sources, but that it is too costly to identify and sort out relevant texts manually. Using machine-learning techniques offers greater flexibility and accuracy than a simple keyword search. In recent years, much progress has been made in the field of machine learning and techniques for automatic or semi-automatic information retrieval are being used in practice in widely different areas [4,8,10,11,16].

When a large set of relevant articles with IT incidents has been identified this can be used to synthesize a list of typical types of incidents, and it can also serve as a bank of examples. These sources can be used in training, and as examples in the identification step of the risk analysis process.

2 Related Work

There exist a few systems that are relevant to the research carried out in this study. The first relevant system is GDACS[1] (Global Disaster and Alert Coordination System), which provides alerts and ways of calculating consequences of sudden disasters that help in improvements of emergency response and capabilities [13]. The GDACS website also monitors the media and social media for news about each disaster, although this uses keywords and not machine learning. GDACS was developed in a joint project by the United Nations, the European Commission and disaster managers worldwide.

Another relevant system is EMM[2] (Europe Media Monitor), developed by the Joint Research Centre, which collects news from news portals worldwide in 60 languages. It also performs classification of the collected news articles. After the classification it analyses the news to find different kinds of alerts (e.g. earthquake, storm, lightning strike, flooding) and presents these alerts in a visual representation. This system uses clustering techniques (grouping the objects in a way that all objects in one group are more similar to each other as compared to objects in other groups) and keywords for the identification of events and their graphical display [3].

The research presented in this paper is carried out using text classification and information filtering techniques. A number of studies have discussed text classification in general and presented results by using different machine learning algorithms. For example, Sebastiani et al. [11] present an overview of different available machine learning approaches for automatic text classification. In the study, the authors discuss different methods, their applications, their effectiveness and recent progress that has been made in the field.

Machine learning algorithms have also been used in spam e-mail filtering [1,2]. It has been concluded that the use of machine learning algorithms is better than the use of simple keyword search for spam filtering. This indicates that Machine Learning techniques also are better for the purpose of this paper than keyword search.

[1] http://www.gdacs.org

[2] http://www.emm.newsbrief.eu

3 Research Methodology

3.1 Research Objectives

The main objective of this study is to prepare and evaluate a prototype of a system that automatically identifies information pertaining to IT incidents reported in online news sources. The main research question is to explore how IT incidents can be identified from available news sources on the Internet. The main research question has been broken down into the following more detailed research questions:

- RQ1: What search and identification methods should be used for this type of search?
- RQ2: Which steps are important to effectively use machine learning for searching and identification of IT incidents?
- RQ3: What kind of data sources should be used for the identification of IT incidents?
- RQ4: How much effort is required to perform this identification of IT incidents in practice?

3.2 Research Approach

This is an explorative study initiated by an idea of automatic identification of IT incidents reported in online news sources that can later be used for risk analysis. The research in this study is carried out in a number of following steps.

1. An appropriate method or technique for the identification problem was selected. After reviewing the literature and techniques it was found that using a machine learning technique is the best solution for this study [1].
2. A data source containing relevant articles, about IT risks or incidents, in a large number and in a relatively low frequency for the example to be realistic was found. After searching available and accessible news web sources, one web source was found that contains thousands of articles of interest for this study.
3. Retrieval of text (IT risk or incident articles) from the selected data source and its cleaning was performed. The data retrieval and initial data cleaning was performed by specially written java code. Then, the further data cleaning and processing was performed by using a machine learning tool.
4. Two datasets were manually prepared to be used as learning and evaluation datasets. Only after these steps, the selected machine learning techniques could be applied on a training dataset for the creation of a classifier.

5. Finally the selected machine learning techniques are applied on the training dataset for the creation of a classifier. Then, cross validation was performed within the training dataset as well as by applying the learned classifier on an independent evaluation dataset.

For this study, machine learning and data pre-processing algorithms were applied through the WEKA[3] (Waikato Environment for Knowledge Analysis) datamining tool. It is an open source software developed in java by researchers of the University of Waikato, New Zealand. It has a well-designed GUI and it has a wide variety of machine learning algorithms implemented, and allows direct application of algorithms on datasets [5].

4 proposed IT Incident Identification Method

This paper proposes a method for the identification of IT incidents in the context of risk and vulnerability analysis by using machine learning algorithms. With the processing of text written in natural language, and by using machine learning techniques it is possible to classify or identify IT incidents automatically.

It can be noticed that there is an overlap between the research approach and the proposed identification method. However, the research approach includes some steps that are carried out for research purpose only, like the first step that is used only for research. The research includes using the proposed approach, and evaluating it. In the research the proposed approach for identification is only used once, and with archived data. In a real setting it will be used continuously and new texts are constantly investigated.

Below is a brief description of steps required for the identification of IT incidents. This study has been conducted by developing a prototype solution in the following steps:

1. Retrieval of unstructured data from the web source.
2. Preparing a training dataset by manually classifying a smaller set of documents.
3. Pre-processing of retrieved data performed to convert unstructured data to structured form required for the machine learning tool.
   (a) Cleaning of data performed by removing stop words and stemming.
   (b) Selection of features with the help of feature selection algorithms.
4. Build a classifier.
   (a) Build a classifier by training on the manually classified training dataset.

[3] http://www.cs.waikato.ac.nz/ml/weka

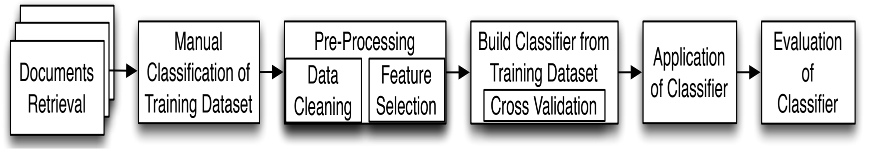

Documents Retrieval
Manual Classification of Training Dataset
Pre-Processing
Data Cleaning
Feature Selection
Build Classifier from Training Dataset
Cross Validation
Application of Classifier
Evaluation of Classifier

Figure 1: Identification process of IT incidents

   (b) Evaluate the selected classifying method internally on the training data (cross-validation).

5. Application of classifier on a separate evaluation data set.

6. Evaluation of classifier by applying it on the complete dataset downloaded from the data source for the identification of relevant IT incidents.

The steps involved in the proposed method for identification of IT incidents are summarized in Figure 1.

4.1 Retrieval of Data

For the development of a prototype that will identify IT incidents a large number of possible candidate texts need to be retrieved. As the focus of this study is on IT incidents that have happened, a suitable source for this is the e-news stories written by reporters from all over the world. For the evaluation of the approach proposed in this study, we selected "The Risk Digest"[4] as data source. This is a newsgroup about IT-related risks and incidents moderated by Peter G. Neumann. It consists of 27 volumes published from 1985 to 2013 and each volume contains a varying number of issues (between 45 and 98). For the retrieval of these documents a simple software tool was used written in Java. By using this tool, 25,500 records were downloaded and, after removal of the surrounding HTML code, stored in a format suited for the machine learning tool.

4.2 Preparing a Training Dataset

A small set of documents from the retrieved data were selected for the training dataset (dataset X) and manually classified. It contains 200 documents, which were selected randomly from the large set of downloaded documents. As a large proportion of this dataset contains articles relevant to IT incidents, which is unlikely to be the case in other news sources it was decided to limit the scope of the example used in this study to identifying only those documents that were about IT incidents in commercial aviation. This reduces the percentage of relevant articles

[4] http://www.catless.ncl.ac.uk/Risks/

from over 50 percent to less than 10 percent, which is more realistic in this type of information retrieval problem. That is, the purpose of focusing on commercial aviation was to obtain a dataset with a realistic number of interesting articles.

Therefore the training dataset was manually classified into the following two classes:

A) Articles of interest: documents about IT incidents in commercial aviation, i.e. everything about non-military aircraft, airports, airline ticket systems, flight control, baggage handling, design of aircraft, etc.

B) Articles not of interest: documents not about IT incidents in commercial aviation, also including related, but separate, fields of military aircraft, space technology, etc.

After manual classification, class A contained 12 documents and the remaining 188 documents were classified in class B.

4.3 Pre-Processing of Data

Machine learning algorithms can extract useful information from a huge amount of available data semi-automatically or automatically. The first step for this task is to perform the transformation of unstructured data into structured form. The unstructured data in the form of strings of characters must be transformed into a machine-readable representation in this case a vector of words. This transformation was performed by using a machine learning tool that leads to a feature value representation. Pre-processing is crucial to attain useful results [6]; it includes management of missing feature values, data cleaning, and feature selection.

Data cleaning detects and removes errors and incompatibilities from the retrieved data to achieve more accurate results by improving the quality of data. Data cleaning consists of removal of stop words, converting words to lowercase, and stemming. For the stemming, in this study the *Lovins* stemming algorithm was used [5].

Feature selection is the next important step to select the most significant and correlated features pertaining to the class attributes. The class attribute is a special attribute that defines the classes (attribute used for the outcome of classification). Feature selection algorithms select a subset of suitable features from original large dataset. In this study, the feature selection was performed by using the *CfsSubsetEval* algorithm.

4.4 Build Classifier

After preparing a training dataset, by manually classifying 200 randomly selected documents, and performing the above mentioned steps the classifier was trained by using the Naive Bayes machine learning algorithm in WEKA. Then, the built classifier was evaluated by performing cross validation in 10 folds. (10-folds divide dataset in 10 equal parts and then use 9 for learning and 1 for evaluation, and repeat this process by using all parts one by one) on the training dataset.

4.5 Application of Classifier

Next, the classifier was applied on the remaining large dataset. After applying the classifier, the documents in the large dataset were classified in two classes, about commercial aviation and not about commercial aviation.

4.6 Evaluation of Identification

After identifying relevant documents with the application of the built classifier an evaluation of the identification results was performed. For the evaluation of the results an evaluation dataset (dataset Y) was prepared by manually classifying a smaller set of documents like the training dataset. Then, the built classifier was applied on the manually classified evaluation dataset for measuring the performance and accuracy.

5 Results and Discussions

This section presents the results of applying the proposed method on the example data used for this study. It presents results of both before and after carrying out stemming and stop words removal. First dataset X was used as training and cross-validation set and dataset Y as evaluation set, then the order was reversed.

Table 1 and Table 2 present classification results with and without carrying out stemming and stop words removal. The results without parentheses are with carrying out stemming and stop words removal. The results within parentheses are without carrying out stemming and stop words removal. Table 1 presents the results of classifier I that is built from dataset X and then applied on dataset Y for evaluation.

Classifier I correctly classified 9 of the training documents as interesting (TP, true positives) and 3 documents incorrectly classified as not-interesting (FN, false negatives). It also correctly classified 187 documents as not-interesting (TN, true negatives) and one document incorrectly classified as interesting (FP, false positives). This gives an *accuracy* ($(TP + TN)/(TP + TN + FP + FN)$) of 98%, a *precision* ($TP/(TP + FP)$) of 90%, and a *recall* ($TP/(TP + FN)$) of 75%.

Then, the cross-validation was performed by using 10-folds and classifier I obtained similar results as in the learning phase. Classifier I was applied on dataset Y, with the results as shown in the column to the right in Table 1. Although there is an increase in the accuracy after performing stemming and stop words removal, the results are not optimal in the sense that some interesting documents are missed.

Table 1: Classifier I results (built on dataset X, evaluated on dataset Y)

	Training		**Cross-Validation**		**Evaluation**	
	Predicted		**Predicted**		**Predicted**	
Actual	Class-A	Class-B	Class-A	Class-B	Class-A	Class-B
Class-A	9 (10)	3 (2)	9 (10)	3 (2)	4 (2)	8 (10)
Class-B	1 (4)	187 (184)	1 (4)	187 (184)	1 (1)	187 (187)
Accuracy	98% (97%)		98% (97%)		95.5% (94.5%)	
Precision	90% (71%)		90% (71%)		80% (66%)	
Recall	75% (83%)		75% (83%)		33% (16%)	

Table 2: Classifier II results (built on dataset Y, evaluated on dataset X)

	Training		**Cross-Validation**		**Evaluation**	
	Predicted		**Predicted**		**Predicted**	
Actual	Class-A	Class-B	Class-A	Class-B	Class-A	Class-B
Class-A	11 (9)	1 (3)	10 (9)	2 (3)	7 (7)	5 (5)
Class-B	1 (3)	187 (185)	4 (3)	184 (185)	7 (14)	181 (174)
Accuracy	99% (97%)		97% (97%)		94% (90.5%)	
Precision	91% (75%)		71% (75%)		50% (33%)	
Recall	91% (75%)		83% (75%)		58% (58%)	

Table 2 presents the results of classifier II built from dataset Y and then applied on dataset X for evaluation. Classifier II correctly classified 11 documents as interesting and one document incorrectly classified as not-interesting. It also correctly classified 187 documents as not-interesting and one document incorrectly classified as interesting. Here, a small increase in accuracy compared to learning results of classifier I can be noticed.

After performing cross-validation, classifier II correctly classified 10 documents as interesting and 2 documents incorrectly classified as not-interesting. It also correctly classified 184 documents as not-interesting and 4 documents incorrectly classified as interesting. Here, a decrease in the precision compared to the results of classifier I can be noticed. The accuracy has also decreased. The cross-validation results (with stemming) of classifier II are probably good for the IT incident identification system.

After applying classifier II on the evaluation dataset X, it correctly classified 7 documents as interesting and 5 documents incorrectly classified as not-interesting. Classifier II also correctly classified 181 documents as not-interesting and 7 documents incorrectly classified as interesting. Here, the increase in the numbers of true positives as compared to the evaluation results presented in Table 1 can be noticed. The number of true negatives has decreased and due to this there is an decrease in accuracy, but it performed well for the interesting document class (Class-A), which is a requirement for the proposed system.

As mentioned by Sebastiani et al. [11] and Yu [15], stemming has both the positive and negative effects on accuracy for text classification results. Toman et al. [14] also mentioned that the stemming even decreases the accuracy of a text classifier. However, in the results presented in Table 1 and Table 2, it can be noticed that after carrying out stemming and stop words removal both classifiers performed well and obtained the results with an increase in accuracy as compared to without stemming and stop words removal.

6 Conclusions and Future Work

Based on the reviewed literature and related work we conclude that using a machine learning technique as a search and identification method (RQ1) is the best solution for this type of search because it is too costly to identify and sort out relevant risks or IT incidents manually. Regarding the steps to effectively use machine learning for searching and identification of IT incidents (RQ2), a method is proposed in this study in Section 4 (see Figure 1). The proposed method worked rather well in this study by identifying potential texts about IT incidents. This indicates that it is possible to support the work of identifying texts about IT incidents with automated methods like this. This support could be an important aid in the process of building a database of occurred IT incidents. Regarding data source (RQ3), we conclude that the best data source for this study could be the e-news stories written by reporters from all over the world. In this study we selected “The Risk Digest" as data source. It is an example of real data containing relevant articles, about IT risks or incidents, in a large number and in a relatively low frequency for the example to be realistic. This, however, needs further research. Regarding required effort (RQ4), based on the results of this study, we conclude that it is possible to identify interesting texts from a large number of potential texts but it requires a substantial effort to set up. With one of the two investigated data sets, 33% of all the relevant articles were found, and with the other data set, 58% of the relevant articles were found. This means that a large number of relevant texts can be found with this support, even if not all texts are found.

However further research is needed to understand if it is possible to transfer these conclusions to other texts, especially for texts that are taken from e.g. news papers, and if it is possible increase the recall of the method with further training of the method.

7 Acknowledgement

This work was funded by the Swedish Civil Contingencies Agency under a grant for PRIVAD, Program for Risk and Vulnerability Analysis Development.

Bibliography

[1] I. Androutsopoulos, G. Paliouras, V. Karkaletsis, G. Sakkis, C. D. Spyropoulos, and P. Stamatopoulos, "Learning to filter spam e-mail: A comparison of a naive bayesian and a memory-based approach," in *proceedings of the workshop on Machine Learning and Textual Information Access*, 2000, pp. 1–13.

[2] W. A. Awad and S. M. ELseuofi, "Machine learning methods for E-mail classification," *International Journal of Computer Applications*, vol. 16, no. 1, pp. 39–45, February 2011.

[3] C. Best, B. Pouliquen, R. Steinberger, E. Goot, K. Blackler, F. Fuart, T. Oellinger, and C. Ignat, *Intelligence and Security Informatics*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2006, vol. 3975, ch. Towards Automatic Event Tracking, pp. 26–34.

[4] T. S. Guzella and W. M. Caminhas, "A review of machine learning approaches to spam filtering," *International Journal of Expert Systems with Applications*, vol. 36, no. 7, pp. 10 206–10 222, 2009.

[5] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: An update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, Nov. 2009.

[6] S. B. Kotsiantis, D. Kanellopoulos, and P. Pintelas, "Data preprocessing for supervised learning," *International Journal of Computer Science*, vol. 1, no. 2, pp. 111–117, 2006.

[7] N. G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*. The MIT Press, 2012.

[8] I. Paparrizos, B. B. Cambazoglu, and A. Gionis, "Machine learned job recommendation," in *Proceedings of the 5:th ACM conference on Recommender systems*. ACM, 2011, pp. 325–328.

[9] G. G. Roy, "A risk management framework for software engineering practice," in *proceedings of the Australian Software Engineering Conference*, 2004, pp. 60–67.

[10] N. S. Roy and B. Rossi, "Towards an improvement of bug severity classification," in *proceedings of the 40:th EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA '14)*, Aug 2014, pp. 269–276.

[11] F. Sebastiani and C. N. Ricerche, "Machine learning in automated text categorization," *ACM Computing Surveys*, vol. 34, pp. 1–47, 2002.

[12] S. M. Sulaman, K. Weyns, and M. Höst, "A review of research on risk analysis methods for it systems," in *proceedings of the 17:th International Conference on Evaluation and Assessment in Software Engineering (EASE '13)*. ACM, 2013, pp. 86–96.

[13] D. G. Tom, "Global disaster alert and coordination system: More effective and efficient humanitarian response," in *proceedings of the 14:th International Emergency Management Society (TIEMS) Annual Conference*, 2007, pp. 324–334.

[14] M. Toman, R. Tesar, and K. Jezek, "Influence of word normalization on text classification," in *proceedings of Multidisciplinary Approaches to Global Information Systems*, 2006.

[15] B. Yu, "An evaluation of text classification methods for literary study," Ph.D. dissertation, University of Illinois at Urbana-Champaign, Champaign, IL, USA, 2006.

[16] S. Zander, T. Nguyen, and G. Armitage, "Automated traffic classification and application identification using machine learning," in *proceedings of the 30:th IEEE Conference on Local Computer Networks*, 2005, pp. 250–257.

Paper VI

A method for assessing resilience of socio-technical IT-systems

Abstract

Modern society is increasingly dependent on IT-systems. Due to this dependence it is important that IT-networks are designed to be resilient, meaning that they will either maintain or quickly recover their functionality when exposed to strain. Simulation-based methods that consider supply network topology as well as system responsible for repairing supply network have previously been used and found to be beneficial for assessing resilience of electricity and water distribution systems. A method of this kind is here applied for IT-networks. In effect the IT-system is studied as a socio-technical system, here broadly understood as a system whose functionality is dependent on technical as well as organizational sub-systems. The aim of the present research is to test if such a method is applicable for assessing resilience of IT-systems, meaning that: 1) it is possible to use based on available data, in this case gathered mainly through interviews with focus groups, 2) the results are relevant for users/owners/maintainers. The method is tested in a case study on the IT-network of one department of Lund university as well as the university core network. Results show that the method is applicable for the studied IT-network and that we can obtain the resilience metrics sought for. It is found that the method can enable system owners to see if and for what levels of strain they are presently reaching their desired targets concerning system resilience. Concerning the relevance of the method, feedback from system experts indicates that the method might primarily be useful for IT-systems whose failure would 1) result in large economic losses (e.g. IT-system of major private companies) or 2) lead to loss of health or safety (e.g. IT-systems of governmental organizations and

hospitals).

Finn Landegren, Sardar Muhammad Sulaman, Peter Möller, Martin Höst and Jonas Johansson,
In Proceedings of the 26th European Safety and Reliability Conference (ESREL '16), pages 2199–2206. Taylor & Francis Group.

1 Introduction

Modern society is increasingly dependent on IT-systems. Due to this dependence it is important that IT-networks are designed to be resilient, meaning that they will either maintain or quickly recover their functionality when exposed to strain. This calls for new and better approaches for assessing IT-network resilience. Simulation-based methods that consider network topology as well as system responsible for repairing supply network have previously been used and found to be beneficial for assessing resilience of electricity [4] and water distribution systems [10]. In effect in these cases the infrastructure system is studied as a socio-technical system, here broadly understood as a system whose functionality is dependent on technical as well as organizational sub-systems. To the authors' knowledge a method of this kind has not been applied for IT-networks. The aim of present research is to test whether a simulation based method addressing socio-technical system aspects is applicable for assessing resilience of IT-systems, meaning that: 1) it is possible to use based on available data, in our case gathered mainly through interviews with focus groups, 2) the results are relevant for users/owners/maintainers. Resilience is here broadly understood as the ability of a system to withstand sudden shocks [2]. The here proposed method enables assessment of three crucial aspects of resilience: 1) robustness, meaning the ability to withstand strain without loss of function, 2) rapidity, meaning the ability to quickly regain function in case of strain and 3) resilience loss, here understood as the overall number of user hours of service that are lost due to a given disruption. Fig. 1 shows how these metrics are related to the so called resilience curve showing functionality of a system through time. As suggested by [12] recovery time (T) provides a measure of rapidity, and $F_{norm} - X$ provides a measure of robustness, where F_{norm} is the normal functionality of the system and X is the initial loss in functionality. X and T are, similarly to [12], calculated using equations 1 and 2.

$$X = F_{norm} - F(t_0) \tag{1}$$

$$T = t_1 - t_0 \tag{2}$$

Where $F(t_0)$ is system functionality immediately after disruption, t_1 is the time point at which the system is fully restored and t_0 is the time point of the disturbance. Functionality can potentially be measured in several ways. In this paper it is understood as number of users that have service. In accordance with [1], resilience loss is calculated through equation 3.

$$RL = \int_{t_0}^{t_1} [F_{norm} - F(t)]dt \tag{3}$$

Where $F(t)$ is the system functionality at time t.

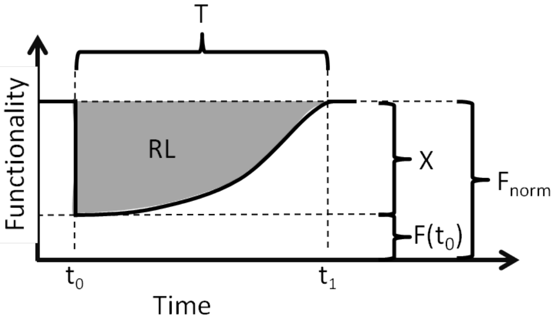


Figure 1: Resilience curve showing level of functionality of a disrupted system over time.

2 Related Work

A number of approaches can be found for assessing IT network resilience. [11] proposes a risk-based approach for designing a resilient network. The approach includes three risk management design techniques. The first technique is to minimize the maximum damage that could occur in the network, the second technique is to minimize the maximum risk in the network and the third technique is to minimize the root mean squared damage. [7] presents a risk assessment process to identify the challenges with the highest potential impact to a network and its users. The outcome of the presented process is a prioritized list of challenges and associated system faults, which can guide network engineers towards the mechanisms that have to be built into the network to ensure network resilience. [9] evaluates IT-network resilience based on the operational state of the network and the service delivered in which disturbances are simulated as node and link failures. [8] presents a systematic approach for building resilient IT-networks. The authors first presents fundamental elements at the framework level such as metrics, policies, and information sensing mechanisms. Then, the authors presents a case study to show how the developed framework and mechanisms can be applied to enhance resilience. [5] presents a model to construct awareness of resilience issues that consists of four stages. The authors model the behavior of defender and attacker in the proposed model by using Extended Generalized Stochastic Game Nets (EGSGN) which combines Game theory and Stochastic Petri Nets. In a case study the authors show how to use EGSGN to depict the network resilience situation in the proposed model. Existing methods for assessing IT-network resilience are generally not considering socio-technical aspects in great detail. In particular, no identified approach is combining explicit modeling of IT-network as well as repair system. To answer to this gap this study presents a method based on Monte Carlo simulation for assessing resilience of socio-technical IT-systems. Simulations are done using a hybrid modeling approach which considers the technical network, represented using graph theory, as well as the repair system, represented by a queuing model.

3 Research Methodology

The study was conducted as a case study [6] where the simulation based method was evaluated by applying it on a real IT system. The method for evaluation of resilience had, previously to the present study, been employed on an electricity network [4]. The model has not been evaluated for IT systems.

The objective of this study is to investigate the following:

1. Is it possible to apply the model on IT systems?
2. Are the results of applying the model on IT systems relevant?
3. What additional factors would be important to introduce in the model in order to make the relevance for IT systems higher?

Question 1 concerns if it is possible to apply the model on IT systems, which is determined by investigating if all required data is available for a real IT system and how hard it is to collect this data. Question 2 concerns how useful the results of using the model are for IT managers responsible for large IT systems.

The study was conducted in three main steps, see Fig. 2. The steps were carried out in sequence. In the first step, data was collected from the case organization. This was conducted by first having a meeting with representatives from the organization followed by continued contacts where detailed information was collected. At the meeting the ideas of the model were presented by the authors of this paper, and the representatives of the organization presented their IT system and its architecture, history, requirements, etc. At the meeting the authors of this paper and two representatives from the organization participated.

In the second step the model was used with the data that was collected in the previous step. This was based on the original, general, model and it meant that some aspects of the model was adapted to fit to this situation.

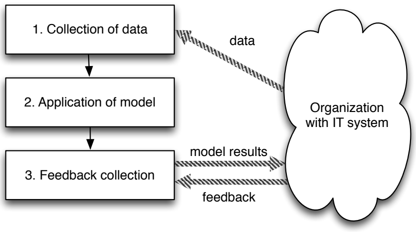


Figure 2: Methodology

In the third step results from applying the model were presented to the organization. This was carried out in an informal setting where representatives from the organization participated and were able to give feedback on the usefulness of the approach. The researchers actively asked for information about the usefulness of the results at the meeting.

Compared to the research questions it can be concluded that question 1 was mainly answered in the first and second step and question 2 and 3 in the third step.

4 Simulation Model

A hybrid modeling approach is used which considers the technical network, represented using graph theory, as well as the repair system, represented by a queuing model. Here the previously developed hybrid-model is presented and described in an IT-system context. The infrastructure network is represented as a graph $G(V,E)$ where V consists of N nodes and E consists of M edges (see e.g. [3]).

$$V = [\, n_1 \, n_2 \, ... \, n_N \,] \tag{4}$$

$$E = [\, e_1 \, e_2 \, ... \, e_M \,] \tag{5}$$

The complete set of components in the network, C, is constituted by the sets V and E, i.e .:

$$C = [\, n_1 \, n_2 \, ... \, n_N \, e_1 \, e_2 \, ... \, e_M \,] = [\, c_1 \, c_2 \, ... \, c_{N+M} \,] \tag{6}$$

An adjacency matrix A is used to represent the connections in the network, see equation 7.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix} \tag{7}$$

In IT-networks, especially in the core network, two nodes often have multiple connections. Hence in the connectivity matrix used here a_{ij} is k if n_i is connected to n_j by means of k separate edges. Each connecting cable could potentially be represented as a separate network edge. Here, however, cables that are geographically co-located are represented as one single network edge. This way of representing network edges will be reasonable concerning some hazard types, such as excavator accidents, flooding or sabotage, in which case all components located in proximity are likely to be damaged. Other hazard sources may not be as accurately reflected through this model assumption. These types of hazards can easily be simulated by representing each geographically co-located cable as an individual edge.

Further, two type of services need to be modeled: connection to internet and connection to server. This is here considered by performing two separate analyses, one assessing restoration of internet connection, the other assessing restoration of server connection. Two kinds of node faults can occur. Partial hardware failure of access switch means that the connectivity of the network will remain unchanged while some of the customers connected to the access switch loose access to the network. Remaining faults mean that the network connectivity is changed so that no information passes through the failed component. Two vectors are used to represent node faults. One vector provides the number of faults that has occurred in the network nodes. This number may be either 0, 1 or 2 (the latter being possible only in access switches). The other vector is a boolean and provides information concerning if a given component is an access switch that has had hardware failure. If node i has had a failure of type 1 the i:th element in the two vectors will be 1. If it has had a failure of type 2 the i:th element will be 1 and 0 (meaning that it has had one fault which is not a hardware fault) or 2 and 1 (meaning that it has had 2 hardware faults). Concerning edges, a failure of edge e_i connecting n_j and n_k is simulated by setting a_{kj} and a_{jk} to 0.

Capacity is not considered in the network model. A user is therefore considered to be connected to a source if there is at least one unbroken path leading from the access switch at which the customer is located to at least one source. A breadth-first search strategy is used to find all supplied nodes.

The repair system is represented as a queuing system (see Fig. 3) in which installation jobs and component faults are served by a chosen number of repair teams p, of i different types, $[\, p_1\, p_2 \,...\, p_i \,]$, using materiel m, of j different types, $[\, m_1\, m_2 \,...\, m_j \,]$. Each type of team has a certain competence concerning what types of repairs they can perform. Failure modes and repair times of components are stochastic variables.

Model entities: The repair system model has four types of entities: jobs, queues, repair teams and stock. Jobs have a repair time, and a vector specifying resources needed for repair. One queue holds repair jobs and an additional queue holds completed jobs, thereby simplifying post-simulation analysis. Repair teams serve the first failure in the queue that is serviceable by the given repair team with resources in stock. The stock holds materiel of different amounts (specified

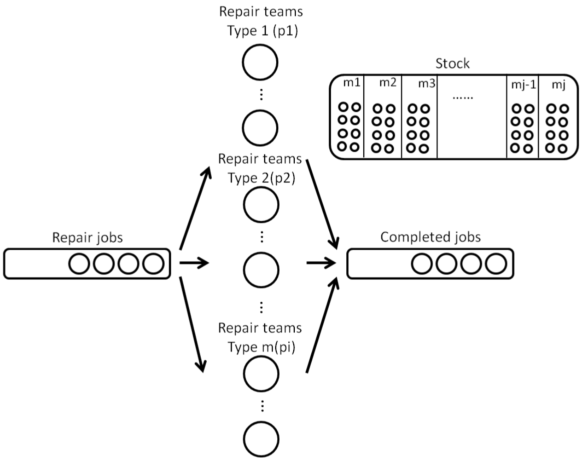


Figure 3: Overview of repair system model, including repair teams of i different types, $[\, p_1\, p_2 \ldots p_i \,]$, two queues and stock containing materiel of j different types, $[\, m_1\, m_2 \ldots m_j \,]$.

by a vector).

Process Overview and Scheduling: On each time step it is checked if the stock inventory and the number of repair teams should be updated (a matrix specifies when and by how much the inventory should be refilled). If a repair team is not currently working, it begins repair on the first job in queue that can be serviced by the repair team with the available materiel.

Job prioritization: Though variations may occur between IT-systems, order of repair tasks is likely to be decided to a high extent so that customer service hours is maximized. This goal is reached by prioritizing jobs that will bring back most network users per hour of work time. Utility of repair jobs is thus calculated through equation 8.

$$U_i = C_{itot}/T_i \tag{8}$$

Where U_i is the utility of repairing component i, C_{itot} is the number of customers brought back by repairing component i and T_i is the expected repair time of component i. C_itot is calculated through equation 9.

$$C_{itot} = (C_i + C_{I1} + C_{I2} + \ldots + C_{Ix}) * b_s \tag{9}$$

Where C_i is the number of customers connected to component i, C_{Ij} is the total number of customers in the j:th non-supplied network region with member components connected to component i, b_s is a Boolean being 1 if component i is connected to a supplied region in the network, otherwise 0. A network region denotes a set of nodes and edges that are directly or indirectly connected. A region is supplied with a given service if there is at least one non-failed source supplying the service among its member components. The hybrid model was implemented in MatlabÃĆÂő. Object oriented programming, influenced by agent based modeling, was used for modeling the repair system.

Outages are simulated using the network model. Sampled scenarios are used due to the excessive simulation times that would result if a complete scenario set was used. Sampled scenarios are represented as in equation 10.

$$SM_i = \begin{pmatrix} c_{11} & \cdots & c_{x1} \\ \vdots & \ddots & \vdots \\ c_{1S} & \cdots & c_{xS} \end{pmatrix} \tag{10}$$

Where x is the number of failed components and S is the sample size. Each row in the strain matrix thus represents one outage scenario.

5 Case Study

The method for studying resilience is tested in a case study on the core IT-network of Lund university as well as the network of one university department. Data was collected through interviews with focus groups and through document analysis, and the results were assessed for relevance through an interview. In the model, the core network is considered as well as the network of one department, altogether consisting of 22 nodes (routers, firewalls, distribution switches, access switches and storage server) and 46 edges (cables above or below ground), see Fig. 4. All cables in the core network are below ground while all cables in the department network are above ground. It is assumed that access switches are fully utilized, meaning that they serve 60 users each. This gives a total of 720 users.

Interviews were carried out with system experts, in order to collect information about fault modes, their relative probabilities, repair times and resources needed for repair. There are three possible fault modes for network nodes: hardware, operator and software faults, with estimated relative probabilities being 10, 70 and 20% respectively. It was decided that software faults would not be considered due the irregular consequences and repair times associated with these faults. Also, fault in access point to internet is not considered. When only hardware and operator faults are considered they get the relative probabilities 12.5 and 87.5% respectively. Repair data is shown in Table 1 and 2.

Repair occurs only during work hours. This may not be the case in other, more safety critical IT-systems. For this reason when restoration time is calculated, we only get the number of work hours that is required to restore the system, not the overall time that passes from initial disruption to complete restoration. As was remarked on previously, access switches may fail partially or completely. The former happens when one single hardware fault occurs in the component. In the

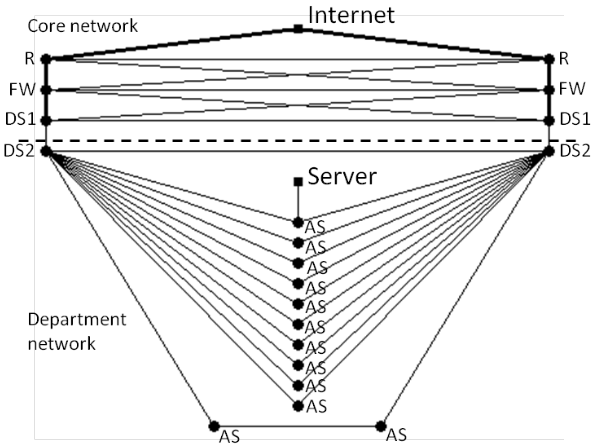


Figure 4: IT-network. Thick lines indicate that nodes are connected by two edges, thin lines indicate connection by single edge. Dashed line indicates boundary between core network and department network. R=router, FW=firewall, DS1/2=distribution switch1/2, AS=access switch.

studied network a partial failure will result in loss of service for 2/3 of the customers at the access switch. A complete failure occurs either when the component has an operator fault, or when it has two simultaneous hardware faults.

Three types of repairers can be distinguished: 1) central IT management personnel, 2) department IT management personnel, 3) excavator teams. Central IT management personnel repair routers, firewalls, access switches 1 and 2, access switches as well as above ground cables. Department IT management personnel repairs servers and above ground cables. Excavator teams repair below ground cables. Central IT management repairers work alone, while repairs of type 2 and 3 work in teams of two. Furthermore seven different kinds of materiel resources can be distinguished. If the amount of a certain resource is assessed to be sufficient for any type of repair it is set to be infinite. Table 3 shows the amount of resources that becomes available over time. It was found that central IT management repairers will not at all times follow the prioritization rule given by equation 8. If a fault in the core network is causing outages this will always be prioritized by central IT management repairers before any faults in the department network.

The present analysis is performed for six levels of network strain: N-1, N-2, N-3, N-6, N-9 and N-12, where N-k denotes the failure of k out of the total N network components. It is also performed with respect to two different network services, connection to internet and connection to server. Since fault scenarios as well as fault modes and repair times are stochastic variables many samples must be simulated in order to reach a convergent result. In order to decide the sample size a convergence analysis is performed. The analysis is performed for three levels of strain N-1, N-6 and N-12, which cover the range of strains that are used in the analysis and it is performed for restoration of internet as well as server connection. For each level of strain three series of simulations are run, each encompassing 300,000 samples. It is investigated how mean, 95%-percentile and maximum restoration time converges. Convergence is considered to have occurred when the error relative to the final value, obtained after 300,000 samples, is permanently below 0.5 hours for all three series. It is found that the mean value converges after 1290 and 19 samples for restoration of internet and server respectively, that the 95%-percentile converges after 37,620 and 7940 samples for restoration of internet and server connection respectively and that the maximum does not converge within the studied interval. Based on this information the sample size of the analysis is chosen to be 50,000. This is sufficient for estimating mean value and 95%-percentile with desired precision, although not for estimating the maximum restoration time.

Table 1: Fault modes of network nodes.

Component	Failure mode	Relative probability	Repair time (h)	Resources needed
Router, Firewall, Dist. Switch 1 and 2, Acc. Switch	Hardware	12.5%	2	Central repairer + Hardware
	Operator	87.5%	2	Central repairer
Server	Hardware	12.5%	0.5	Department repair team + Hardware
	Operator	87.5%	2	Department repair team

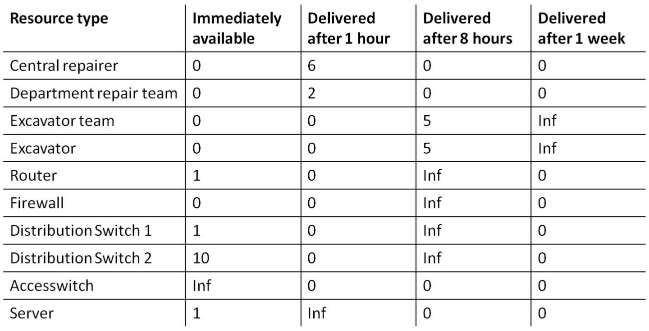
Table 2: Fault modes of network edges.

Component	Repair time (h)	Resources needed
Below ground (core network)	[16,40]	Excavator team +excavator
Above ground (department)	0.5	Central repairer or department repair team

6 Result

In this section results from applying the model on the case study network are first presented. Then the feed-back from the IT-system personnel is presented.

Table 3: Resources available over time.



Resource type	Immediately available	Delivered after 1 hour	Delivered after 8 hours	Delivered after 1 week
Central repairer	0	6	0	0
Department repair team	0	2	0	0
Excavator team	0	0	5	Inf
Excavator	0	0	5	Inf
Router	1	0	Inf	0
Firewall	0	0	Inf	0
Distribution Switch 1	1	0	Inf	0
Distribution Switch 2	10	0	Inf	0
Accesswitch	Inf	0	0	0
Server	1	Inf	0	0

6.1 Application of Model

In Fig. 5 we see an overview of robustness and rapidity of all simulated scenarios (i.e. 300,000 samples, since 50,000 samples are used for each of the six levels of strain). Robustness has been normalized by division with F_{norm}. Results are shown for restoration of internet and server connection respectively. It can be seen that robustness is extremely low for some scenarios, i.e. equal or close to zero. Furthermore, for some of these scenarios rapidity is very high, almost 50 hours in worst case for restoration of internet and almost 40 hours for restoration of server connection. Restoration time given here is only the work time needed for restoring the system. The total time from failure to complete recovery will be about one week in these extreme scenarios, assuming that repairers have a 40 hour work week.

Fig. 6 shows rapidity with respect to restoration of internet and server as a function of level of strain. Mean value (solid line) and 5 and 95% percentiles (dashed lines) are shown in the graphs. We see that for restoration of internet supply the mean restoration time is increasing steadily, from mean restoration time of about one hour up to about 5 hours at the N-12 level of strain. The percentiles furthermore show that for strains of N-9 or less the majority of the scenarios will be in a small interval. At N-12 level of strain this changes. Here the 95%-percentile is much higher, about 18 hours. For restoration of connection to server we see that at the N-1 level of strain the mean rapidity is about 1 hour. It then increases to a bit more than 3 hours for N-12 level of strain. Percentiles show that there is a larger variability for low levels of strain than for higher levels of strain.

Fig. 7 shows normalized robustness with respect to restoration of internet and server connection as a function of level of strain. Mean value (solid line) and 5 and 95% percentiles (dashed lines) are shown in the graph. We see that for restoration of internet connection mean robustness will at the N-1 level of strain be close to 1. It then decreases to a bit more than 0.4 for N-12 level of strain meaning that on average a bit more than 40% of the users have internet connection following

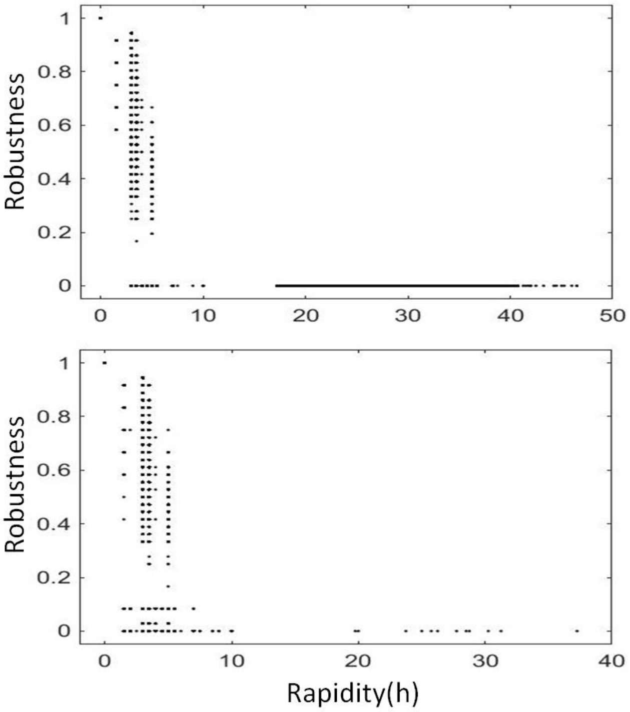


Figure 5: Robustness and rapidity with respect to restoration of internet (upper) and server connection (lower) for all simulated scenarios and all levels of strain.

a disturbance of this magnitude. The percentiles also show that the variability increases significantly with the level of strain. The mean robustness is slightly lower for all levels of strain. The lower percentile shows that variability is greater concerning restoration of server than restoration of internet for strain levels N-3 to N-6.

Fig. 8 shows RL (see equation 3) of all simulated scenarios and levels of strain, sorted in ascending order. RL is shown on a logarithmic scale meaning that scenarios with $RL = 0$ are not seen. We can thereby see that less than 2/5 of the N-1 scenarios will have $RL > 0$, while at the N-12 level of strain all scenarios have a $RL > 0$. We can also see that for restoration of internet the worst scenarios have RL larger than 10^4 while for restoration of server the worst scenarios have RL larger than 10^3.

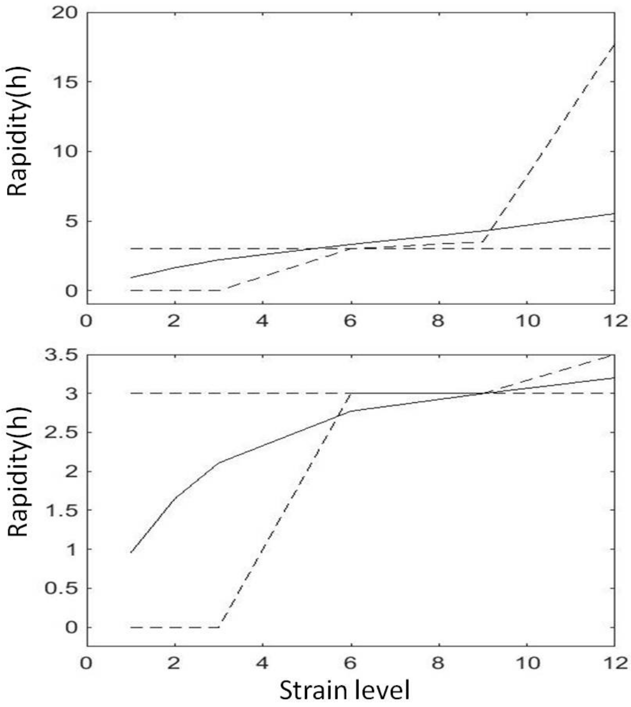


Figure 6: Rapidity concerning restoration of internet (upper) and server (lower) as function of level of strain. Solid line (–) is mean value and dashed lines (- -) are 5 and 95% percentiles.

6.2 Feedback From IT-System Personnel

Results included in this paper, except for Fig. 8 that was obtained at a later stage, were presented to the system experts. It was found that the method was not considered useful for them at present. The main reason for this is that their system is not considered to be critical enough. The system experts also explained that they had not experienced accidents involving N-2 failures or greater. For this reason they consider it less motivated to prepare for such events. They however were aware that such accidents had happened in other systems with great consequences, notably the Swedish university network (SUN) which was out of operation due to two simultaneous cable faults. Also, during a construction project at the campus a core network cable was moved and it was then found that it had not been connected properly. In effect, given that there was not such a network redundancy as was believed an N-2 event would not have been very improbable. While not finding the method useful for their own purposes, the system experts suggested

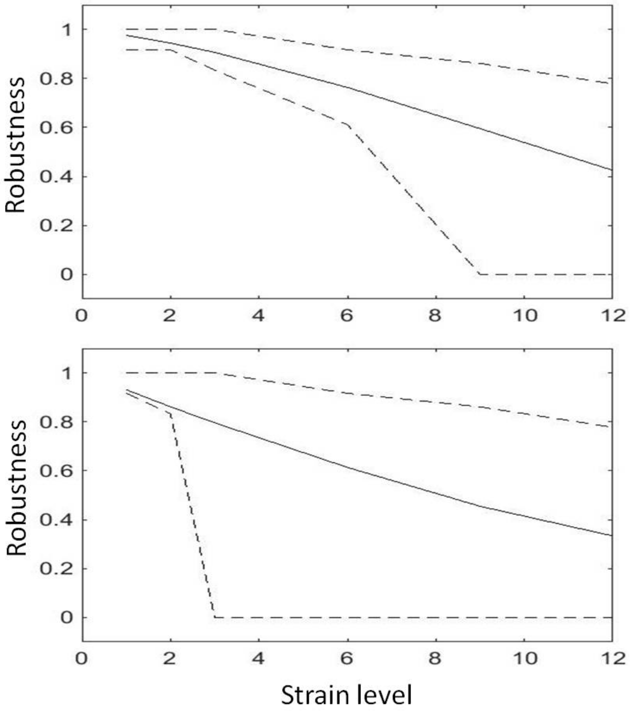


Figure 7: Robustness with respect to internet (upper) and server connection (lower) as function of level of strain. Solid line (–) is mean value and dashed lines (- -) are 5 and 95% percentiles.

that a method of the described kind could be of interest to use for IT-systems that are critical either from an economic perspective or from a health and safety perspective. Their recommendation was to test its usefulness in IT-systems of large private companies or hospitals.

7 Discussion

Simulation based methods that consider repair system as well as technical network have previously been used for assessing resilience of electricity and water distribution systems. The objective of the present paper is to test if such a method 1) is applicable within the IT-context, 2) is giving relevant results and 3) captures all relevant factors. Below these three questions are discussed.

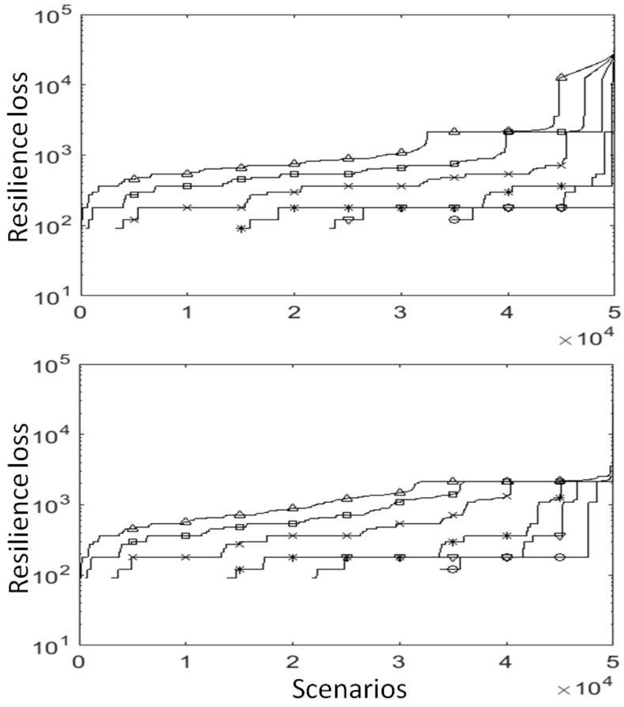


Figure 8: All scenarios sorted in ascending order based on resilience loss with respect to restoration of internet (upper) and server connection (lower). Unit of resilience loss is user hours. Strain levels N-1(o), N-2(downward pointing triangle), N-3(*), N-6(x), N-9(square), N-12(upward pointing triangle).

7.1 Applicability

Concerning objective 1) preliminary results show that the method is applicable for the studied IT-network and that we can obtain the resilience metrics sought for. Results show that disturbance scenarios for which resilience is low can be identified, based on three important resilience indicators: rapidity, robustness and resilience loss. Results of this type can be useful in the process of increasing system resilience. Results further show how system robustness and rapidity change with level of strain. This makes it possible for system owners/operators to see if and for what levels of strain they are presently reaching their desired targets concerning system resilience.

The interviewed system experts think that restoration should have occurred within a business day, 8 work hours. Results show that overall resilience of the system is generally in agreement with this goal. Concerning restoration of server the 95%-percentile of rapidity will go from about 3 hours for N-1 strain, to about 5 hours for N-12 strain. In other words in 95% of the simulated scenarios restoration time is well within the 8 hour boundary desired. For internet restoration the 95%-percentile of rapidity goes from about 3 hours for N-1 strain, and for all levels of strain except N-12 it will be less than 5 hours. It is therefore well within the the 8 hour time boundary. However, for N-12 level of strain the restoration time of internet is significantly longer, 20 hours. Also results showing rapidity for all individual scenarios show that some scenarios have restoration times that are significantly longer than 8 hours.

7.2 Relevance

Concerning objective 2) feed-back from system experts showed that the method was not found to be relevant to the personnel of the studied system, since this system is not considered to be critical enough. The system experts however thought that the method could be useful when applied to IT-systems whose outage could cause either large economic losses or risk to health and safety.

7.3 Model Completeness

Concerning objective 3) software faults are not considered in the model. Software faults are said to constitute about 20% of the total number of faults in network nodes. These faults are not considered mainly due to their irregular repair time and consequences. Considering this type of faults is a possible topic for future work. It could also be of interest to consider cost of resources in future work. This would make it possible to use optimization methods in order to find a set of repair system resources that maximizes system resilience given a specified available budget.

8 Conclusions

Preliminary results indicate that the developed simulation-based method is useful for assessing resilience of IT-systems, meaning that the method can enable system owners to see if and for what levels of strain they are presently reaching their desired targets concerning three crucial indicators of resilience; robustness, rapidity and resilience loss. Interview with system experts showed that the method might first and foremost be of importance for IT-systems that are critical for society, (e.g. IT-systems of government or hospitals) or whose failure will cause major economic losses (e.g. IT-systems of major companies). In future work it will be of interest to apply the model to IT-systems of these types. A further direction of future work is to find ways of taking software faults into account. Also it can be of interest to consider cost of resources in the model, so as to enable optimization of the repair system given a limited budget.

Bibliography

[1] M. Bruneau, S. E. Chang, R. T. Eguchi, G. C. Lee, T. D. O'Rourke, A. M. Reinhorn, M. Shinozuka, K. Tierney, W. A. Wallace, and D. von Winterfeldt, "A framework to quantitatively assess and enhance the seismic resilience of communities," *Earthquake spectra*, vol. 19, no. 4, pp. 733–752, 2003.

[2] L. K. Comfort, A. Boin, and C. C. Demchak, "The rise of resilience," in *Designing resilience: Preparing for extreme events*, A. Boin, L. K. Comfort, and C. C. Demchak, Eds. University of Pittsburgh Pre, 2013.

[3] J. Johansson, "Risk and vulnerability analysis of interdependent technical infrastructures: addressing socio-technical systems," Ph.D. dissertation, Lund University, Sweden, 2010.

[4] F. Landegren, "Technical infrastructure networks as socio-technical systems: Addressing infrastructure resilience and societal outage consequences," PhD Thesis, Lund University, Sweden, 2017.

[5] M. Liu, T. Feng, P. Smith, and D. Hutchison, "Situational awareness for improving network resilience management," in *Proceedings of the 9th International Conference on Information Security Practice and Experience (ISPEC), Lanzhou, China, May 12-14, 2013.* Springer Berlin Heidelberg, 2013, pp. 31–43.

[6] P. Runeson, M. Höst, A. Rainer, and B. Regnell, *Case Study Research in Software Engineering*. Wiley, 2011.

[7] M. Schöller, P. Smith, and D. Hutchison, "Assessing risk for network resilience," in *Proceedings of the 3rd International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Oct 2011, pp. 1–7.

[8] P. Smith, D. Hutchison, J. P. G. Sterbenz, M. Schöller, A. Fessi, M. Karaliopoulos, C. Lac, and B. Plattner, "Network resilience: a systematic approach," *IEEE Communications Magazine*, vol. 49, no. 7, pp. 88–97, July 2011.

[9] J. P. G. Sterbenz, E. K. Çetinkaya, M. A. Hameed, A. Jabbar, S. Qian, and J. P. Rohrer, "Evaluation of network resilience, survivability, and disruption tolerance: analysis, topology generation, simulation, and experimentation," *Telecommunication Systems*, vol. 52, no. 2, pp. 705–736, 2011. [Online]. Available: http://dx.doi.org/10.1007/s11235-011-9573-6

[10] T. Tabucchi, R. Davidson, and S. Brink, "Simulation of post-earthquake water supply system restoration," *Civil Engineering and Environmental Systems*, vol. 27, no. 4, pp. 263–279, 2010.

[11] K. Vajanapoom, D. Tipper, and S. Akavipat, “A risk management approach to resilient network design,” in *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Oct 2010, pp. 622–627.

[12] C. W. Zobel, “Representing perceived tradeoffs in defining disaster resilience,” *Decision Support Systems*, vol. 50, no. 2, pp. 394–403, 2011.