LUND UNIVERSITY

Improving Risk Analysis Practices in Governmental Organizations

Sulaman, Sardar Muhammad

2015

Link to publication

Citation for published version (APA):
Sulaman, S. M. (2015). *Improving Risk Analysis Practices in Governmental Organizations*. [Licentiate Thesis, Department of Computer Science].

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Improving Risk Analysis Practices in Governmental Organizations

Sardar Muhammad Sulaman

Licentiate Thesis, 2015
Department of Computer Science
Lund University

ISSN 1652-4691
Licentiate Thesis 2, 2015
LU-CS-DISS:2015-2

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, 2015

© *2015 Sardar Muhammad Sulaman*

To my parents,

Zahid Aziz

&

Shafqat Sultana

Without whom none of my success would be possible.

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 system or service, e.g., water, power supply, transportation, is dependent on IT systems, and failure of these systems has serious and negative effects on society. In general, governmental organizations are responsible for delivery of these services to society. The increasing dependence on critical IT systems also makes them more and more complex. 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. Risk analysis of technical systems has a long history in mechanical and electrical engineering. 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 IT systems in governmental organizations. In this thesis the improvements in risk analysis processes are addressed in two different ways. First, by understanding what types of methods are available for IT systems and how they can be improved. Second, by developing new effective and efficient risk analysis methods that can be useful to analyze IT systems in governmental organizations.

In this thesis work, a systematic mapping study was carried out to understand existing methods and techniques used for analyzing IT systems. It found very few empirical research papers about the evaluation of existing risk analysis methods. The results of the mapping study suggest to empirically investigate risk analysis methods for analyzing IT systems to conclude which methods are more effective than others.

Based on the results of the mapping study a case study was carried out to evaluate the effectiveness and efficiency of an existing risk analysis method, System Theoretic Process Analysis (STPA). Based on the results of the mapping study a controlled experiment was carried out to evaluate the effectiveness of risk analysis methods. The effectiveness of risk analysis methods was evaluated 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.

The work presented in 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.

A case study was carried out to save historical information about IT incidents to be used later for risk analysis. This study investigates how difficult it is to find relevant risks from the available sources and the effort required to set up such a system. It also investigates how accurate the found risks are. It is believed that this 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.

The presented research work in this thesis provides research about methods and tools for governmental organizations to improve their 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 thank to my supervisor, Prof. Martin Höst, for his great support, valuable feedbacks, discussions, continuous support and encouragement. Thank for always being ready to take discussions and for guiding me. I learned so many things from you, thanks for everything. 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, they deserve it for providing an inspiring and motivational work environment. In addition, I would like to thanks my fellow researchers especially those in the Software Engineering Research Group for sharing valuable experiences and supporting me towards my goal (PhD). In particular, to Dr. Krzysztof Wnuk for always being ready to take discussions. And special thank to Alma, Markus, Elizabeth, Usman, Hassan, Mehmet and Jörn for their discussions and shared experiences about both research relevant and non-research relevant topics. A special thank to Dr. Taimoor Abbas and Atif Yaqoob for their support and continuous invitations for delicious food.

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 to pursue my goal (PhD).

Sardar Muhammad Sulaman

List of Publications

This licentiate thesis consists of two parts. The first part 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 four 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. Association for computing machinery (ACM) 2013.

II **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'14), Penn State University, Pennsylvania, USA, pages 424-428, 2014.

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'14), London, UK. Association for computing machinery (ACM) 2014.

IV **Identification of IT Incidents for Improved Risk Analysis by Using Machine Learning**
Sardar Muhammad Sulaman, Kim Weyns, Martin Höst
Submitted to a conference.

In this thesis the included papers will be referred as [I], [II], [III], and [IV].

Contribution Statement

Sardar Muhammad Sulaman is the first author of all included papers. He was the main inventor and designer of the studies, and was responsible for running the research processes. Also, he conducted most of the writing.

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.

The hazard analysis reported in Paper II started as a project report of a PhD course (Safety Critical Software-Intensive Systems). The study was mainly designed and carried out by Sardar Muhammad Sulaman. Sardar Muhammad Sulaman was responsible for carrying out the hazard analysis and he wrote a majority of the paper. Dr. Taimoor Abbas contributed by discussions about the application of hazard analysis method and by reviewing hazard analysis results. Dr. Krzysztof Wnuk and Prof. Martin Höst contributed as active reviewers for the study.

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 a majority of the paper, with committed assistance from Dr. Krzysztof Wnuk and Prof. Martin Höst.

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

List of Related Publications

I have also contributed to the following publications. However, these publications are not included in this thesis:

5. **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'14), Verona, Italy, Pages 17-24. IEEE Computer Society 2014.

6. **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, 2014.

CONTENTS

PART I

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 [33]. 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 [29]. Moreover, failures in these complex safety-critical systems are often results of multiple interacting decisions and errors [27].

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 a possible failure [33]. 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 [1, 2, 42]. The US Department of Homeland Security issued national strategy documents [42] for the protection of physical and cyber infrastructures that make risk and vulnerability assessments mandatory. As today almost all societal critical processes and operations are dependent on IT systems, therefore 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 a necessary activity that protects 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 and prioritized decisions by selecting one from different available options. Management of risks helps in increasing opportunities and decreasing threats to an organization or system [3, 16].

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 [3] suggests a few principles for an 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 and usually that is in the form of historical information about different incidents, expert opinions, and user's/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 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 RA methods [27]. 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 should be used that is sufficiently effective and efficient. Therefore, there is a need to investigate existing RA methods empirically. The work presented in this thesis mainly focuses on improving the risk analysis process for IT systems in public or private organizations. This work introduces improvements in risk analysis processes in two different ways. First, by understanding current practices related to risk analysis of IT systems and how they can be improved. Second, by suggesting improvements in existing risk analysis methods and developing new effective and efficient risk analysis methods to analyze IT 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 methodology 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 the research objective with the research questions. It also discusses the research methodology used. Section 3 presents related work in the field of risk analysis and management. Section 4 presents the summary of the included papers in this thesis. Next, Section 5 synthesizes the results of the research carried out in this thesis. Section 6 concludes the results of this thesis. Finally, Section 7 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

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, driving, etc. 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 impact.

There is no general definition for risk. According to ISO 31000 [3] *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 [20].

The definitions used in this thesis are that, *risk* is the chance that an undesired/negative event might happen. *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 [27]. *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 [3].

Risk analysis can be performed during the development 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. This thesis mainly focuses on risk analysis methods for *operational*

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

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 is used to direct an organization to control risks that can affect its ability to achieve objectives [3]. The coordinated set of activities consists of: risk identification, risk analysis, risk assessment, risk prioritization, and risk mitigation [I] shown in Figure 1. It tries to find a balance between loss prevention and cost associated with countermeasures.

Risk management 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 [I].

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.

1.3 Types of Risk Analysis

There are mainly two types of risk analysis methods, *quantitative* and *qualitative* [I].

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 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.

Qualitative methods, on the other hand, use descriptive values such as 'high', 'medium' or '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.

Semi-quantitative methods, which are intermediary risk analysis techniques that classify the probability and consequences by using quantitative categories such as 'financial loss between 10.000 USD and 100.000 USD' or 'less than once

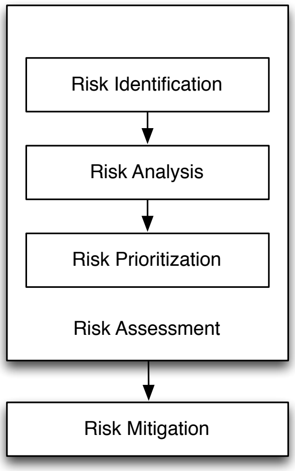


Figure 1: Risk management

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

An *IT* or *Information System* is a combination of hardware, software, data-bases, infrastructure and IT support organized to facilitate decision making in an organization. Hardware contains 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 softwares. Finally, IT support consists of help facilities provided for the proper functioning of IT system such as IT support personals, manuals, documentation or trainings [43]. 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" [25].

Critical IT systems are the systems that provide or support critical services, e.g., water, power supply, transportation, etc. to the society and failure of these systems have serious and negative effects on society. The Swedish Civil Contingencies Agency (MSB) has given examples of critical services or infrastructures [21]:

- 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 failure of these systems has 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 [43],

"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

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

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

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 [19,43].

The focus of the research presented in this thesis is mainly on proactive crisis management, i.e., mitigation and preparedness. Proactive crisis management consists of risk and vulnerability analysis of IT systems that is carried out to assess dependability of these systems. IT systems are becoming more and more complex by having many sub-systems and more and more interconnecting components that make them more vulnerable and unreliable. By assessing and mitigating risks pertaining to these complex IT systems we can improve societal services that are, directly or indirectly, dependent on these systems.

2 Research Design

2.1 Research Objectives

The research presented in this thesis was carried out as a part of PRIVAD, Program for Risk and Vulnerability Analysis Development, project funded by the Swedish Civil Contingencies Agency (MSB). The overall objective of the PRIVAD project 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 improve the analysis process of risks pertaining to IT systems in governmental organizations. The improvements in risk analysis process are addressed in two different ways.

First, by understanding current practices related to risk analysis of IT systems. There exist a number of risk analysis methods for technical systems consisting 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 approaches. This means that there is a need to understand what types of methods are available for IT systems and how they can be improved.

Second, by suggesting improvements in existing risk analysis methods and developing new effective and efficient risk analysis methods that can be useful to analyze IT systems in governmental organizations. Because the dependence on IT systems in governmental organizations is very crucial. Today almost every system or service, e.g., water, power supply, transportation, etc. are dependent on IT systems, and failures of these systems 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.

2.2 Research Questions

The research objective is broken down into the following more detailed research questions:

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

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

RQ3: How can we improve the effectiveness of a risk analysis method by using different perspectives?

RQ4: How can we identify and save historical information about IT incidents to improve risk analysis process?

RQ1 is important to investigate because it will give an idea about the existing risk analysis methods and the main empirical research that has been conducted in the area of risk analysis for IT systems.

RQ2 is also important to investigate since a general goal of any risk analysis method is to find an as complete set of risks as possible. However, it is not clear what measures should be used to evaluate or compare existing risk analysis methods. Therefore, it is important to investigate different measures that can be used to compare risk analysis methods.

RQ3 is relevant to investigate since the use of different perspectives in risk analysis and management has been suggested [5, 7, 16, 27, 30, 39, 44] but it is not empirically assessed. Therefore, it is important to empirically assess the potential of different perspectives in risk management processes. 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.

Finally, RQ4 is an exploratory investigation that leads to how can we identify and save historical information about IT incidents that can be later used for risk analysis to improve the risk analysis process.

2.3 Research Methodology

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 [8]. Empirical research can be characterized as *exploratory* or *evaluative*.

In *exploratory* research the aim is to understand, with more or less prejudice, a specific phenomenon [37]. 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 [9]. 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.

In *evaluative* research the aim is to assess the effects and effectiveness of innovations, interventions, practices etc. [37]. It involves a systematic collection of data, which can be of both qualitative and quantitative type.

This thesis mainly contains *exploratory* and *evaluative* empirical research, based on studies using systematic mapping study, experiment and case study research methodologies.

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 [22]. On the other hand, the goal of a systematic literature review (SLR) is to analyze and aggregate the base of empirical evidence [23]. 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 [22] table I).

In Paper II a case study is presented that is carried out as evaluative research. *Case study* is an in-depth study of a specific phenomenon or artifact in a real life context. Case study is a suitable research method when the boundaries between phenomenon and context cannot be clearly specified. In Paper II a risk analysis method, System Theoretic Process Analysis (STPA), has been evaluated by applying it on an automobile safety application. The used primary data was of third degree [26], which is published literature and system description of qualitative nature.

Paper III presents results from evaluative research based on a controlled experiment as research method. *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 Paper IV a case study is presented that is carried out as an exploratory research. The research in Paper IV was initiated by an idea of automatic identification of IT incidents reported in online news sources that can later be used for

Table 1: Research type and method used in the included papers

Work	Research type	Research method
Paper I	Exploratory	Systematic mapping study
Paper II	Evaluative	Case study
Paper III	Evaluative	Experiment
Paper IV	Exploratory	Case study

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

3 Related Work

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 [16] and ISO/IEC 27002 [15], by national governmental organizations, such as the National Institute of Standards and Technology (NIST) [39] or the British Central Communication and Telecommunication Agency (CCTA) [13], by non-governmental organizations such as Club de la Sécurité de l'Information Français (CLUSIF) [31] or by research institutes such as the Carnegie Mellon Software Engineering Institute (SEI) [4]. A detailed comparison of some of these frameworks is conducted by ENISA [11] and Syalim et al. [40].

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) [10], Failure Mode and Effect Analysis (FMEA) [30] and Hazard and operability study (HAZOP) [35]. 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 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 consists of a logical diagram that shows the relation between the system components and their failures. Ericson [10] presented a review of the research performed on FTA with its advantages and shortcomings.

FMEA is a risk or hazard analysis technique that can be applied as both a top-down and a bottom-up approach [30]. The top-down approach (usually function oriented) is mainly used in an early design phase before deciding the whole system structure. However, FTA is a suitable choice for the top down approach. 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. [14] 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 planning phase of a system. 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 [29].

Leveson [27] 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 [28]. Nakao et al. [32] evaluated the STPA technique in a case study where it is applied on an operational crew return vehicle design. The feasibility and usefulness of STPA technique is also evaluated thoroughly for early system design phase by Ishimatsu et al. [17]. These studies [17, 32] conclude that with STPA it is possible to recognize safety requirements and constraints of the system before the detailed design. Several authors [27, 34, 41] reported positive outcomes from applying STPA on various systems.

To improve the risk analysis process, researchers and practitioners have introduced some improvements in current practices. For example, to tackle the lack of information in early design problem, Johannessen et al. [18] 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 [12] 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 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 [44]. RBRA was presented on an illustrative example from the computer software engineering domain but not experimentally investigated. Leveson [27] and McDermott et al. [30] 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 [6, 36]. Perspective-based reading was also applied for object oriented design inspections [38], code reviews [24] and usability inspections [45]. 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.

4 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

In this paper, 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 1086 unique papers were identified. Then 57 out of 1086 papers were identified as relevant for this study. 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 conduct research where already available methods are evaluated. 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: Hazard Analysis of Collision Avoidance System using STPA

In this paper, we present experiences gained by applying the System Theoretic Process Analysis (STPA) method for hazard analysis on a forward collision avoidance system in a case study. Our main objectives were to investigate effectiveness (in terms of the number and quality of identified hazards) and time efficiency (in terms of required efforts) of the studied method. Based on the findings of this study STPA is shown as an effective and efficient hazard analysis method for assessing the safety of a safety-critical system and it requires a moderate level of effort. STPA consists of two steps, identification of inadequate control commands or events and identification of their causal factors.

Using STPA we identified 14 inadequate control commands or events in the analyzed system with their associated hazards. We believe that the reason of the effectiveness of STPA is that it considers and focuses in step 1 on the control commands or events and their feedbacks instead of only individual component failures. Regarding effort required to apply STPA on a safety-critical system, based on the results found in this study, it can be concluded that STPA requires moderate effort in relation to the level of experience of the study participants. We believe that the reason for its effort efficiency is that the use of STPA for hazard analysis allows domain experts and hazard analysts to complement each other because of its simplicity.

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 [5, 7, 16, 27, 30, 39, 44]. 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: Identification of IT Incidents for Improved Risk Analysis by Using Machine Learning

In this paper, 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 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.

5 Synthesis

This section summarizes the main results in relation to the research questions and reported studies. Moreover, the main validity threats to the results are also 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 risk analysis methods and approaches exist for analyzing IT systems? Is there any empirical research that compares or evaluates existing risk analysis methods?

This research question 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

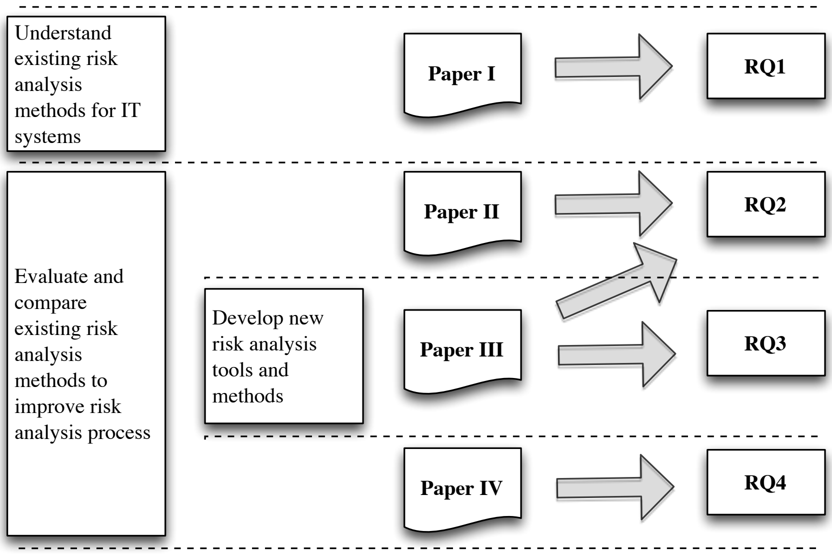

Understand existing risk analysis methods for IT systems
Paper I
RQ1
Evaluate and compare existing risk analysis methods to improve risk analysis process
Paper II
RQ2
Develop new risk analysis tools and methods
Paper III
RQ3
Paper IV
RQ4

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

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.

Concerning 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.

The main validity issue for Paper I concerns missing possible relevant studies due to some 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 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 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.

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

This research question is addressed by Paper II and Paper III. It is not easy to compare or evaluate risk analysis methods because of their subjective nature. 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. In Paper II we investigated effectiveness and time efficiency of the System Theoretic Process Analysis (STPA) hazard analysis method by applying it on a system from the software intensive safety-critical domain (forward collision avoidance system). In Paper III we compared two risk analysis methods (perspective based and traditional risk analysis) and measured their effectiveness by counting identified relevant and non-relevant risks, ease of use and confidence of participants on their identified risks.

Based on the results of Paper II and III, we conclude that risk analysis methods can be evaluated or compared by counting the *number and quality 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. The *time efficiency*, investigated in Paper II, 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 paper II 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).

The main validity issue for Paper II is the measurement of required effort for hazard analysis. To measure time efficiency we had to calculate person hours spent on the performed hazard analysis. We could not manage to measure it exactly because the selected system was analyzed in interruptions. Also, required effort estimates can vary based on experience of risk analysts, available system information, and system scope that is different for different cases.

RQ3: How can we improve the effectiveness of a risk analysis method by using different perspectives?

In Paper III the effectiveness of risk analysis process was investigated by performing a controlled experiment. Here, effectiveness was measured by counting the number of relevant and non-relevant risks identified by the experiment participants with difficulty level of risk analysis process and their confidence about the identified 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, 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 than risk analysis without perspectives. 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.

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.

RQ4: How can we identify and save historical information about already happened incidents to improve risk analysis process?

This research question is addressed by Paper IV. It discusses 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 IV, 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 IV. 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.

The main validity issue for Paper IV is the scalability of the proposed solution. In presented work we only proposed a prototype solution using an example dataset. In the future we are planning 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.

6 Conclusions

The main objective of the research work presented in this thesis was to improve the analysis process of risks pertaining to IT systems in governmental organizations, which is addressed in two different ways as mentioned in Section 2.

First, by understanding current practices related to risk analysis of IT systems because it 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 how they can be improved. Second, by suggesting improvements in the existing risk analysis methods and developing new effective and efficient risk analysis methods that can be useful to analyze IT systems in governmental organizations. Because the dependence on IT systems in governmental organizations is very crucial.

A systematic mapping study is presented in Paper I to answer RQ1, i.e., to understand existing methods and approaches used for analyzing IT systems. 57 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 case study is presented in Paper II to answer RQ2 that evaluated the effectiveness and efficiency of an existing risk analysis method named STPA. RQ2 was also partly answered by Paper III. Based on the results of paper II and III, 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. Experiments seem to be more suitable choice for evaluation of different risk analysis methods, however it depends on the interesting aspects. If the goal is to evaluate methods or techniques quantitatively then experiments are more suitable but if one is interested in qualitative evaluation then case studies might be more appropriate. 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. This way we can evaluate and compare different risk analysis methods to conclude which method is effective among others.

A controlled experiment is presented in Paper III to answer RQ3, i.e., to evaluate effectiveness of risk analysis process. It also presents a new risk analysis method that uses different perspectives while analyzing IT systems. The effectiveness was measured by counting the number of relevant and non-relevant risks identified by the experiment participants. The difficulty level of risk analysis process and the experiment participants confidence about the identified risks were also investigated. 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.

Another case study is presented in Paper IV to answer RQ4, i.e., to save historical information about IT incidents to be used later for risk analysis. Based on the results of Paper IV, 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 IV. 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.

7 Further research agenda

In this section, the future research work is discussed. Based on the results of Paper I we found that there is a need for empirical investigation of different risk analysis

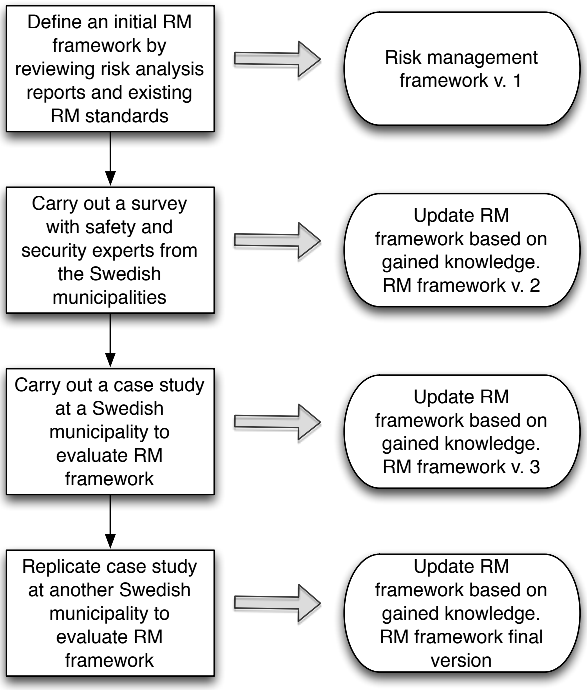

Define an initial RM framework by reviewing risk analysis reports and existing RM standards
Risk management framework v. 1
Carry out a survey with safety and security experts from the Swedish municipalities
Update RM framework based on gained knowledge. RM framework v. 2
Carry out a case study at a Swedish municipality to evaluate RM framework
Update RM framework based on gained knowledge. RM framework v. 3
Replicate case study at another Swedish municipality to evaluate RM framework
Update RM framework based on gained knowledge. RM framework final version

Figure 3: The development of risk management framework

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 in governmental organizations.

A first important continuation of the work in future is to conduct another case study following the case study presented in Paper II focusing on FTA and FMEA and then comparing new results with the results of Paper II. It is also relevant to investigate and explore the effects of the application of STPA in groups with more domain experts and hazard analysts and to compare the results with other traditional hazard analysis methods, i.e., FTA and FMEA.

Another possible research route 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.

In the future we also plan to develop a risk management framework for the Swedish municipalities (governmental organizations) that will be developed by following a few research steps shown in Figure 3.

In Step 1, we have analyzed a number of available risk analysis and management reports of different municipalities of Sweden. In these reports we did not find any detailed analysis of risks pertaining to IT systems. These reports mention the importance of IT systems and the dependence of Swedish municipalities' processes and operations on them. We believe that performing a detailed risk analysis or management of IT systems in Swedish municipalities is a very time and effort consuming activity especially if one follows existing risk management frameworks. Therefore, there is a need for a risk management framework that is less technical than the already proposed risk management frameworks and it also takes into account all critical resource of an organization with large-scale IT systems. The new risk management framework shall require less effort and time to analyze IT systems by focusing on more critical resources that are necessary for critical operations. The new risk management framework will be suitable to analyze IT systems from a higher level by skipping some technical detailed analysis. After Step 1 we have developed an initial version of risk management framework.

In Step 2, the requirements for the planned risk management framework to further refine it will be gathered by conducting a survey with practitioners working in the safety and security domain in different municipalities of Sweden. The main objective of the future survey research is to investigate the following factors, e.g., resources required for risk analysis and management, used risk analysis methods, required competence to perform risk analysis, and definition of clear roles and responsibilities. Step 2 will help in the development of the second version of risk management framework.

In Step 3, we are planning to evaluate the planed risk management framework by conducting a case study in a Swedish municipality by analyzing and managing IT risks. Step 3 will help in the development of the third version of risk management framework.

Finally, in Step 4, we are planning to replicate case study carried out in Step 3 to further evaluate and refine the planned risk management framework. Step 4 will help in the development of the final version of risk management framework.

BIBLIOGRAPHY

[1] SFS (1997:857). *Ellag, Swedish Code of Statutes, Stockholm.* (In Swedish).

[2] 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).

[3] ISO 31000:2009. *Risk Management - Principles and Guidelines*. 2009.

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

[5] J. S. Christopher Alberts, A. Dorofee, and C. Woody. Introduction to the OCTAVE approach. Technical report, Networked Systems Survivability Program, Carnegie Mellon Software Engineering Institute, Pittsburgh, PA 15213-3890, 2003.

[6] 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*, 1:133–164, 1996.

[7] M. H. Noureddine Boudriga and J. Krichene. Netram: A framework for information security risk management. Technical report, Techno-parc El Ghazala, Route de Raoued, Ariana, 2083, Tunisia, 2007.

[8] S. Easterbrook. Empirical research methods for software engineering. In *proceedings of the 22:nd IEEE/ACM International Conference on Automated Software Engineering*, ASE '07, pages 574–574, New York, NY, USA, 2007. ACM.

[9] S. Easterbrook, J. Singer, M-Anne Storey, and D. Damian. Selecting empirical methods for software engineering research. In Forrest Shull, Janice Singer, and Dag I.K. Sjoberg, editors, *Guide to Advanced Empirical Software Engineering*, pages 285–311. Springer London, 2008.

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

[11] European Union Agency for Network and Information Security (ENISA), working group on risk assessment and risk management. Inventory of risk assessment and risk management methods, 2006.

[12] M. Gleirscher. Hazard Analysis for Technical Systems. In *proceedings of the 5:th International Conference on Software Quality, SWQD '13, vol. 133 of LNBIP*, pages 104–124. Springer Berlin Heidelberg, 2013.

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

[14] 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*, pages 119–128, Sept 2007.

[15] International Organization for Standardization. ISO/IEC 27002:2005 - Information technology - Security techniques - Code of practice for information security management, 2005.

[16] International Organization for Standardization. ISO/IEC 27005:2011 - Information technology - Security techniques - Information security risk management, 2011.

[17] 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, September 2010.

[18] P. Johannessen, F. Torner, and J. Torin. Actuator based hazard analysis for safety critical systems. In *Computer Safety, Reliability, and Security*, volume 3219 of *Lecture Notes in Computer Science*, pages 130–141. Springer Berlin Heidelberg, 2004.

[19] J. Johansson. *Risk and Vulnerability Analysis of Interdependent Technical Infrastructure*. PhD thesis, Lund University, 2010.

[20] S. Kaplan and B. J. Garrick. On the quantitative definition of Risk. *Risk Analysis*, 1(1):11–27, 1981.

[21] KBM. Hot och riskrapport. Technical report, Västerås Sweden, 2005.

[22] 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.

[23] 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.

[24] O. Laitenberger, K. El Emam, and T. G. Harbich. An internally replicated quasi-experimental comparison of checklist and perspective based reading of code documents. *IEEE Trans. Software Engineering (USA)*, 27(5):387 – 421, 2001.

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

[26] T. C. Lethbridge, S. E. Sim, and J. Singer. Studying software engineers: Data collection techniques for software field studies. *Empirical Software Engineering*, 10(3):311–341, 2005.

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

[28] 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*, 5(1), 2012.

[29] 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*, pages 37–48, 1995.

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

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

[32] H. Nakao, M. Katahira, Y. Miyamoto, , and N. 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*, pages 497–501, 2011.

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

[34] 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.

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

[36] B. Regnell, P. Runeson, and T. Thelin. Are the perspectives really different? further experimentation on scenario-based reading of requirements. *Empirical Software Engineering*, 5(4):331–356, December 2000.

[37] C. Robson. *Real world research*. Blackwell, 2nd edition, 2002.

[38] 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*, pages 148 – 57, Los Alamitos, CA, USA, 2002.

[39] 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.

[40] 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. IEEE Computer Society, 2009.

[41] 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.

[42] U.S. Dept. of Homeland Security. The national strategy for the physical protection of critical infrastructures and key assets. Technical report, 2003.

[43] K. Weyns. *IT Dependability Management in Governmental Organisations*. PhD thesis, Lund University, 2011.

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

[45] Z. Zhang, V. Basili, and B. Shneideman. Perspective-based usability inspection: an empirical validation of efficacy. *Empirical Software Engineering*, 4(1):43 – 69, 1999.

PART II

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 analyses the project management risks in software development projects.

Section 2 presents related work in the field of risk analysis and systematic literature reviews. Section 4 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 6 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) [5], by non-governmental organizations such as Club de la Sécurité de l'Information Français (CLUSIF) [16] 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) [3], Failure Mode and Effect Analysis (FMEA) [15] 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. [2]). 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 2.

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?

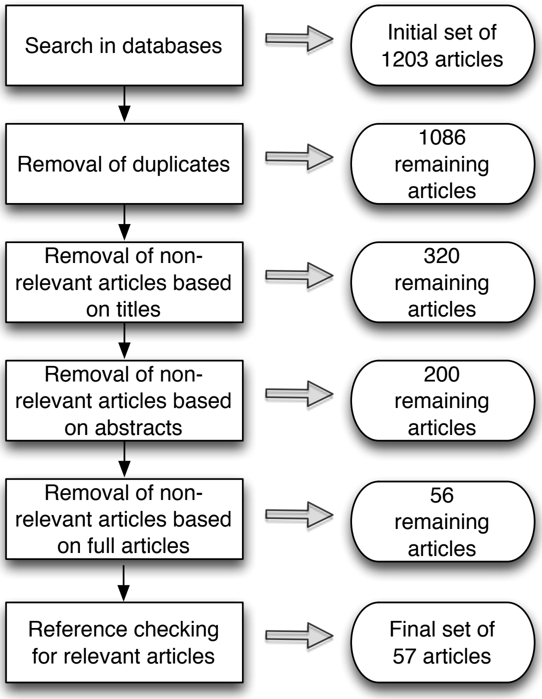

Search in databases
Initial set of 1203 articles
Removal of duplicates
1086 remaining articles
Removal of non-relevant articles based on titles
320 remaining articles
Removal of non-relevant articles based on abstracts
200 remaining articles
Removal of non-relevant articles based on full articles
56 remaining articles
Reference checking for relevant articles
Final set of 57 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 on 23 May 2012 and 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.

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 4.1. The results of this data extraction and classification are discussed in Section 6.

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.

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.

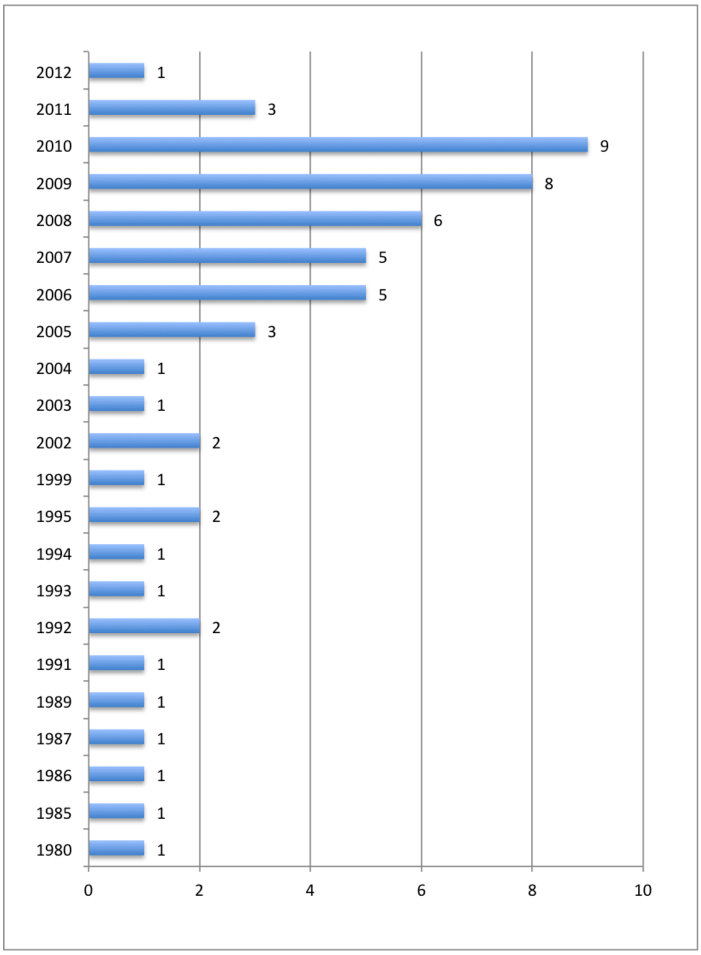


Figure 2: Histogram of publication year for the identified articles

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	5
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	34

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 articles that are about presenting new frameworks, methods and models for risk analysis.

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, are about risk analysis of IT systems in general. This means that the paper does not specify 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.

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 identi-

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	49
Hospital systems	2, 32, 57	3
E-Commerce	25, 50	2
Cloud computing	16	1
E-government	51	1
Web-service systems	21	1

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	36
Empirical -Case study	2, 7, 15, 16, 17, 21, 26, 28, 32, 34, 35, 36, 39, 41, 46, 51, 52, 54, 55, 56, 57	21

fied as analytical and 21 as empirical research. These 21 articles all presented case studies on risk analysis methods, no surveys or experiments were identified.

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 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,

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	28
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	1
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	7
Quantitative	4, 5, 15, 19, 23, 24, 25, 26, 28, 31, 33, 36, 37, 39, 40, 41, 44, 49, 52, 53, 54, 55, 56	23
Combined approach	9, 10, 18, 42, 45, 46	6
Semi-Quantitative	3, 16	2

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.

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 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. 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.

6 Discussion

First of all it can be observed that 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.

Table 6: Paper area vs. empirical or not

Paper area	No	Yes
BCP	0	1
General	1	1
Risk Analysis	12	6
Risk Analysis and Assessment	1	0
Risk Analysis and Management	7	2
Risk Assessment	6	7
Risk Assessment and management	1	0
Risk Assessment and mitigation	0	1
Risk Assessment, prioritization, and management	0	1
Risk Identification, analysis, and management	0	1
Risk management	7	1
Risk Mitigation	1	0
SUM	36	21

Table 7: Risk analysis approach vs. empirical or not

Approach	No	Yes
Combined approach	5	1
General	14	5
Qualitative	3	4
Quantitative	13	10
Semi-quantitative	1	1
SUM	36	21

Only when they are applied in practice they are adapted for specific classes of systems.

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 (36 articles), and fewer articles (21 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 modelling 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) FitŮ, J.O., MacŠas, 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., FernĞndez-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 procedings 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 modelling 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., Fuzzy based threat analysis in total hospital information system, Advances in Computer Science and Information Technology, In Joint Proceedings AST/UCMA/ISA/ACN 2010 Conferences, pp. 1-14, 2010.

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

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

[4] European Union Agency for Network and Information Security (ENISA), working group on risk assessment and risk management. Inventory of risk assessment and risk management methods, 2006.

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

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

[7] International Organization for Standardization. ISO/IEC 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] R. E. McDermott, R. J. Mikulak, and M. R. Beauregard. *The Basics of FMEA, 2nd Edition*. Taylor & Francis, 1996.

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

[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. IEEE Computer Society, 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. IEEE Computer Society, 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

Hazard Analysis of Collision Avoidance System using STPA

Abstract

As our society becomes more and more dependent on IT systems, failures of these systems can harm more and more people and organizations both public and private. Diligently performing risk and hazard analysis helps to minimize the potential harm of the IT systems failures on the society and increases the probability of their undisturbed operation. In this paper, we present experiences gained by applying System Theoretic Process Analysis (STPA) method for hazard analysis on forward collision avoidance system. Our main objectives are to investigate effectiveness (in terms of the number and quality of identified hazards) and time efficiency (in terms of required efforts) of the studied method. Based on the findings of this study STPA has proved to be an effective and efficient hazard analysis method for assessing the safety of a safety-critical system and it requires a moderate level of effort.

Sardar Muhammad Sulaman, Taimoor Abbas, Krzysztof Wnuk and Martin Höst, *In Proceedings of the 11:th International Conference on Information Systems for Crisis Response and Management (ISCRAM '14)*, pages 424–428, 2014.

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, i.e., both for individuals and organizations (private and public), failures of these IT systems can have serious negative consequences and effects on the society. Diligently performing risk and hazard analysis helps to minimize the potential harm of the IT system failures on the society [10, 17]. However, the risk/hazard analysis of a modern socio-technical system is 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 [10]. 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 desynchronized command that may lead to severe accidents. Therefore, new methods for performing risk and hazard analysis optimized for dynamic systems are highly required.

This study presents experience gained by applying the System Theoretic Process Analysis (STPA) [11] method for hazard analysis on forward collision avoidance system as an example of a socio-technical safety-critical system. The main objective of this study is to investigate effectiveness (in terms of the number and quality of identified risks) and time efficiency (in terms of required effort) of STPA hazard analysis method in the software intensive safety-critical system domain by addressing the following study goals:

RG1: How can STPA assess and improve the safety of a software-intensive safety-critical system? This research goal is achieved by applying the STPA method on a forward collision avoidance system.

RG2: How much effort is required to apply STPA on a system for hazard analysis? This goal is achieved by measuring the effort required to apply STPA on our sample system.

The European Telecommunications Standards Institute (ETSI) has identified collision avoidance as one of the most distinctive safety-regulated applications of intelligent transport systems (ITS) [5]. These safety related applications rely on situational awareness such as pre-crash control loss warning and distance to collision warning [1] making this system highly context dependent [10]. Moreover, for the development of ITS collision avoidance applications the National highway Traffic Safety Administration (NHTSA) has identified and prioritized 37 pre-crash scenarios [14].

In this study, STPA method is applied on an ITS application (collision avoidance) with the goal to identify potential hazards with their causal factors. Based on the gained experience from this study it can be concluded that STPA is an effective method as in its first step it identified 14 inadequate control commands or events with their associated hazards. Regarding effort required to apply STPA on a safety-critical system it can be concluded that STPA requires moderate level of effort.

The remainder of this paper is structured as follows: The next section presents related work pertaining to the application of different hazard analysis methods on safety-critical systems. Then it has a brief explanation of the forward collision avoidance system and the used hazard analysis method, i.e., STPA. After this, it presents the results of the performed hazard analysis. Then, the next section discusses the results of the performed analysis with advantages and limitations of the used method. Finally, it summarizes the results of the performed analysis and presents the future work.

2 Related work

In this section, we browse the related work regarding the Fault Tree Analysis (FTA), Failure Mode and Effect Analysis (FMEA), and Hazard and Operability Analysis (HAZOP) hazard analysis techniques as they are considered the most commonly used. FTA is a top-down risk or 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 consists of a logical diagram that shows the relation between the system components and their failures. A fault tree that only contains AND and OR gates can alternatively be represented by Reliability Block Diagram (RBD). Ericson [4] presented a review of the research performed on FTA with its advantages and shortcomings.

FMEA is a risk or hazard analysis technique that can be applied as both the top-down and the bottom-up approach [13]. The top-down approach (usually function oriented) is mainly used in an early design phase before deciding the whole system structure. It analyzes the system by identifying the main system functions and then the potential failures of these functions with causes. The functional failures with significant effects are usually prioritized in the analysis. The top-down approach may also be used on an operational system to identify existing risks. The bottom-up approach is used when a system concept has been decided. During the bottom-up approach, each component on the lowest level of the system design is studied one-by-one and the analysis gets complete after revisiting all components. Grunske et al. [7] 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 technique commonly used in planning phase of a system. 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 multidisciplinary 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 [12].

To tackle the lack of information in early design problem, Johannessen et al. [9] 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 [6] 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.

To summarize, hazard analysis techniques, such as FTA, FMEA and HAZOP mainly focus on component failures and in the system design phase the component failures cannot be considered. Even at later stages it is very hard to identify the causes of a hazard if it is not directly associated to a specific component failure. Thus, in addition to afore-mentioned risk/hazard analysis methods a new hazard analysis technique can be applied, named STPA that considers safety as a control problem rather than a component failure problem [10]. Nakao et al. [15] evaluated the STPA technique in a case study where it is applied on an operational crew-return vehicle design. The feasibility and usefulness of STPA technique is also evaluated thoroughly for early system design phase by Ishimatsu et al. [8]. These studies [8, 15] conclude that with STPA it is possible to recognize safety requirements and constraints of the system before the detailed design. Several authors [10, 16, 18] reported positive outcomes from applying STPA on various systems.

3 Background

3.1 System theoretic process analysis (STPA)

The System Theoretic Process Analysis (STPA) method for hazard analysis developed by Leveson [11], focuses on analyzing the dynamic behavior of the systems and therefore provides significant advantages over the traditional hazard analysis methods. The STPA is a top-down method, just like the FTA method. On the contrary, the 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. The STPA is based on system theory rather than reliability theory and it considers safety as a system's control (constraint) problem rather than a component failure problem. Among the most prominent benefits of using the STPA, Ishimatsu et al. [8] listed the efficiency of the later phase of the STPA when the broader scenarios are analyzed. According to Ishimatsu et al. 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.

It is important to mention that STPA can be applied at any stage; design phase, operational phase and for hazard assessment, 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 can cause some loss, i.e., life, mission, and 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) A control action required is not provided

    (b) An unsafe (incorrect) control action is provided

    (c) A control action is provided too early or too late (wrong time or sequence)

    (d) A control action is stopped too early or applied too long.

The term provided, mentioned above, means correct communication 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 application of STPA, the functional control structure diagram of the system is required and all control loops in the system are identified from the functional control diagram. After this, in each control loop all components that contribute to unsafe behavior of system are identified.

In this study, STPA is applied on a socio-technical system that has three controllers, which are the critical components because they all contain a process model [11]. Controller receives input from almost all components of the system, e.g., sensors, actuators, etc. and then it performs internal calculations to issue a command.

3.2 Forward collision avoidance system

Forward collision avoidance system alerts the driver of a vehicle about a crash situation and applies automatic brakes after a certain time period if the driver does not respond to the warning alert (provides passive and active safety). The system performs two main functions: (1) the object/obstacle detection (by using forward-looking sensors that detect hindrance in front of the vehicle) and (2) the generation of warning or applying auto breaks (passive/active response). The forward-looking

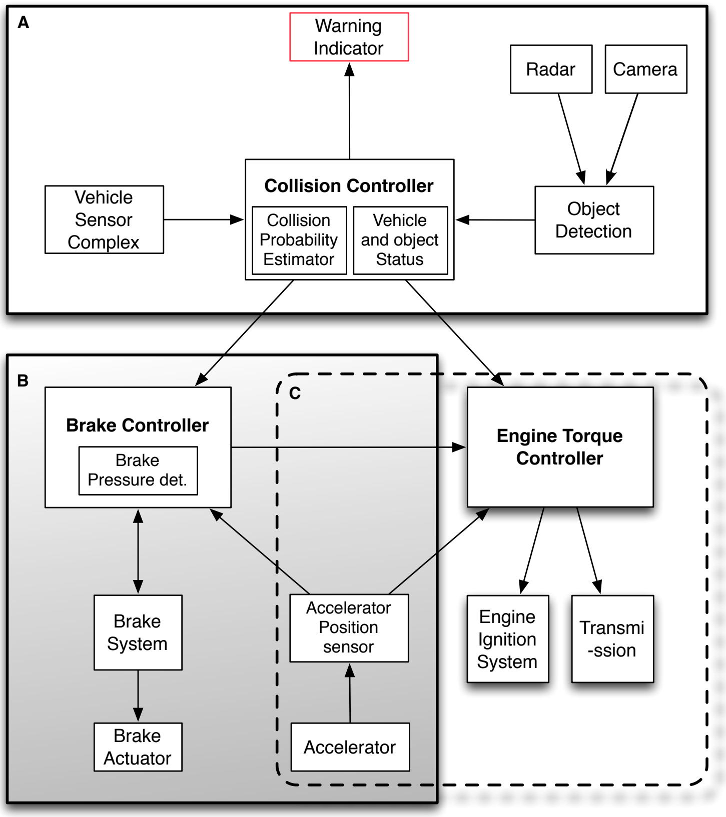


Figure 1: Forward collision avoidance system with autonomous braking [2]

sensors could have few 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 study, we suppose that it uses more than one motion sensors to complement radar and 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 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.

4 Hazard analysis

For hazard analysis 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. After this, the identified inadequate control commands or events were analyzed for their causal factors. Then, the second author analyzed and reviewed the identified inadequate control commands or events and their causal factors. Finally, the results (both inadequate control commands or events and their causal factors) were analyzed and reviewed by the third and the fourth author according to the guidelines of STPA presented in [10, 11].

Table 1 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 it can be noticed from Table 1, all identified control commands or events if not provided 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. Interestingly, similar to a previous study [8], only one of the stopped too soon control commands or events could lead to a hazardous state. One possible interpretation of this result could be that STPA method should be further evaluated on systems that contain more operations that are not only triggers but require time for completion. We assume that both our system and the system presented by Ishimatsu et al. [8]

have a limited number of such cases. Therefore, assessing the sensitivity of the STPA method in identifying these potential hazards should be further explored.

From the identified 14 inadequate control commands or events, we identified 22 hazards. The hazards were classified in three severity levels, i.e., catastrophic, moderate and negligible (see Tables 1 and 2). 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, e.g., 3a, and 4a. 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, it is possible to hypothesize that 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.

Table 2 shows the causal factors for the all identified hazards in step 1 with their severity levels. The first column of Table 2 shows the identified hazards and the next column shows the severity levels, and the third column shows the causal factors for all hazards. Looking at two example hazards, i.e., 5a and 6a, we can notice that they are caused by inadequate control commands from the vehicle and object status signals. Hazard 5a is the incorrect brake pressure determination due to missing vehicle status signal. Hazard 6a is the incorrect brake pressure determination due to missing object status signal.

For example, the causal factors for the hazard 5a could be:

- Failure of vehicle sensor complex (2a)
- Malfunctioning of collision controller due to incomplete process model
- Communication failure or error (no signal)
- Delayed communication (System will fail to provide active safety on time)

The causal factors for the hazard 6a could be:

- Failure of Object detection (1a)
- Malfunctioning of collision controller due to incomplete process model
- Communication failure or error (no signal)
- Delayed communication (System will fail to provide active safety on time)

It can be noticed from Table 2 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 dynamic behavior of the studied system. We report that our results corroborate with the findings presented by Ishimatsu et al. [8] and Pereira et al. [16].

5 Discussion

STPA worked well for the identification of hazards or risks in this study. Specifically, the initial phase (Step 1) of STPA is effective and it does not take too much time and effort. Our experiences show that persons with limited domain experience (maximum 5 years) required one week of effort (interrupted by other activities conducted in parallel) to perform the first step of the analysis and two weeks of interrupted effort (not full time) to perform the second step. In these effort estimates, we assume that the detailed functional diagram is already available. STPA method is suitable for the situations when both domain expert and hazard analyst with limited experience have to complement and supervise each other that yields better results.

However, we have noticed that the straightforward application of STPA on any safety-critical system (especially socio-technical) greatly depends on the availability of the control structure (structural and functional) diagram. Therefore, the quality of the results of STPA method is directly dependent on the quality of the control structure diagram and the amount of included system functional information.

In order to achieve the best possible quality of the entire hazard analysis process from STPA method the main focus should be delivered on step 1. Step 2 of the method is similar to the traditional hazard analysis methods i.e. FTA. The reason of the effectiveness of STPA is that it considers and focuses in step 1 on the control commands or events and their feedbacks instead of only individual component failures. The assumption that a not provided, provided unsafe, provided too early or late, stopped to soon command or feedback is a result of either component failure or communication failure that covers dynamic behavior of the system is the main strength of STPA. This way, STPA also takes into consideration component interactions in the system.

Further actions (deriving constrains and safety requirements) after identification of hazards and their causal factors were not performed in this study because of two reasons. Firstly, the STPA method description [10] and available literature [11] provide limited guidelines of transforming hazards and their causal factors into a robust set of safety requirements and constraints. Therefore, we suggest extending the guidelines of the later phases of STPA. Secondly, the scope of this study was beforehand limited to identification of hazards and their causal factors. Therefore, deriving system constraints or safety requirements remains in the scope of future work.

To summarize, our results corroborate with previously reported positive experiences from STPA application in several domains, e.g., space [8], air traffic [11], defense [16], rail transportation [18], and extend these positive outcomes by an example from the software-intensive automotive domain.

6 Conclusions

This study presented the results of a hazard analysis performed using STPA hazard analysis method on a safety-critical system; forward collision avoidance system. To the best of our knowledge, this is the first study that reports the positive outcomes from the application of STPA on a software-intensive automobile system to assess and improve its safety.

Based on the findings of this study STPA has proved to be an effective and efficient hazard analysis method for assessing the safety of a safety-critical system from automotive domain. Using STPA we mostly seamlessly identified 14 inadequate control commands or events in the analyzed system with their associated hazards. We believe that the reason of the effectiveness of STPA is that it considers and greatly focuses in step 1 on the control commands or events and their feedbacks instead of only individual component failures.

Regarding effort required to apply STPA on a safety-critical system, based on the results found in this study, it can be concluded that STPA requires moderate effort in relation to the level of experience of the study participants. We believe that the reason of its effort efficiency is that the use of STPA for hazard analysis allows domain experts and hazard analysts to complement each other because of its simplicity.

However, there are some shortcomings pertaining to the application guidelines [10, 11] and there is some missing detail about the deriving constrains and safety requirements as a further action after identification of hazards and their causal factors. These shortcomings can easily be overcome by writing the detailed instructions or guidelines for STPA application.

Our positive experiences with STPA suggest that performing both steps 1 and 2 in a group of both domain experts and risk analysts greatly increase the discussion opportunities that lead to more effective and in-depth results. We experienced that the identified risks and hazards had more technical depth and constituted a better view on the analyzed system safety. Future studies are planned to explore the effects of the application of STPA in groups with more domain experts and hazard analysts and to compare the results with other traditional hazard analysis methods, i.e., FTA and FMEA.

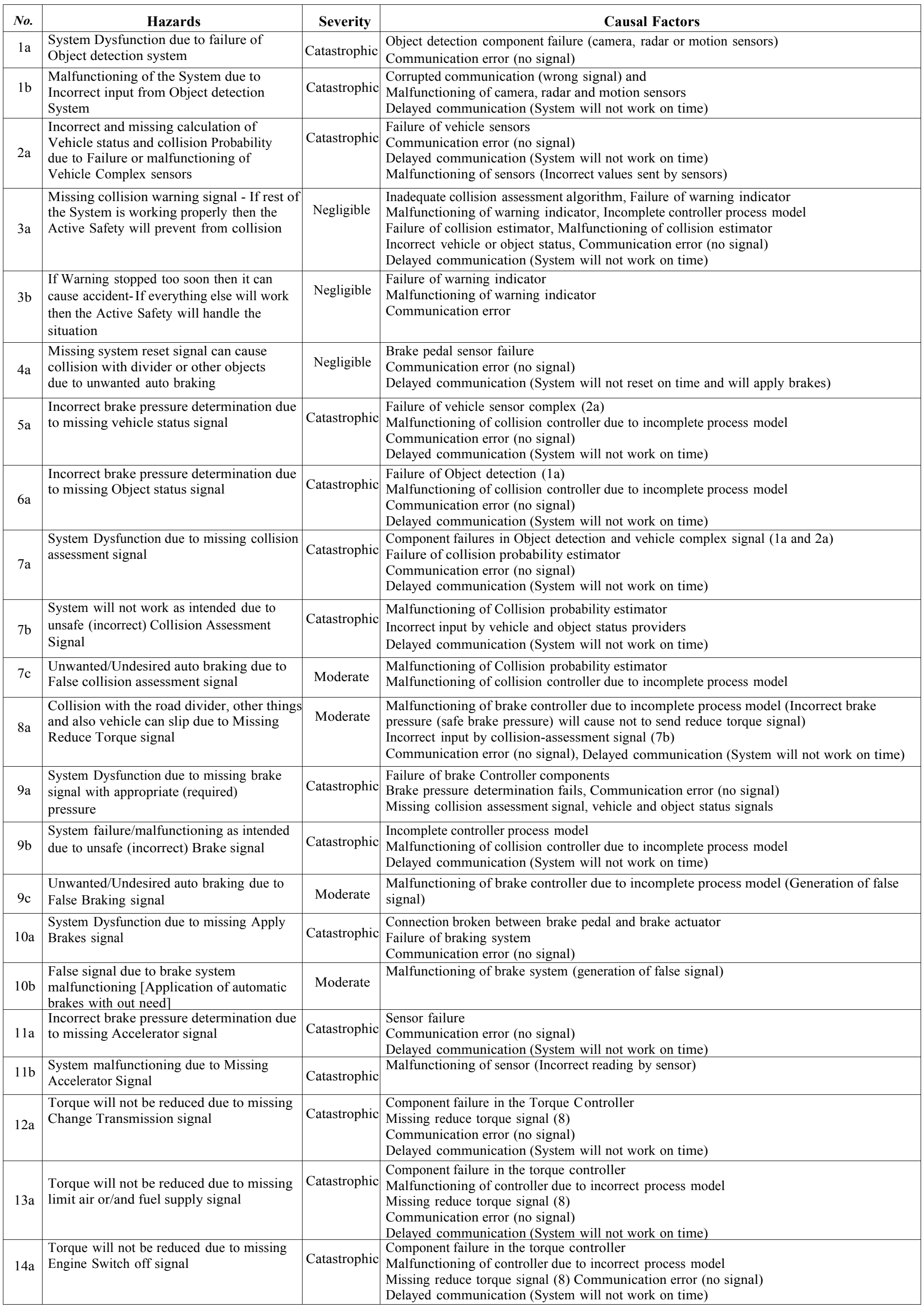
7 Acknowledgment

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

Table 1: Inadequate control actions/commands (Table 1 can be downloaded at http://serg.cs.lth.se/stpa) (OOS = Out of Sequence)

#	Event/Command	Not Provided	Provided Unsafe	Provided			Stopped Too Soon
				Too Early	Too Late	OOS	
1	**Object Detection Signal**	Catastrophic - System disfuntion [Collision] (1a)	Catastrophic - System malfunctioning [Collision] (1b)	N/A	Catastrophic - System disfuntion [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 but if every thing is working properly then the active safety will save collision) (3b)
4	**System Reset Signal (Response from driver by using brakes)**	Negligible - (If every thing is working properly then the active safety will save from collision) (4a)	Negligible - (If every thing is working properly then the active safety will save from collision) (4a)	N/A	Negligible - (If every thing is working properly 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)	Critical - 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**	Catastrophic - Collision with divider, other things and vehicle can slip [Dangerous] (8a)	N/A	N/A	Catastrophic - Collision with divider, other things and vehicle can slip [Dangerous] (8a)	N/A	N/A
9	**Brake Signal with Required Pressure**	Catastrophic - System disfuntion [Collision] (9a)	Catastrophic - System malfunctioning [Collision] (9b)	Critical - False signal due to system malfunctioning [Application of automatic brakes with out need] (9c)	Catastrophic - System disfuntion [Collision] (9a)	N/A	N/A
10	**Apply Brakes Signal**	Catastrophic - System disfuntion [Collision] (10a)	N/A	Critical - False signal due to system malfunctioning [Application of automatic brakes with out need] (10b)	Catastrophic - System disfuntion [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 [Dangerous] (12a)	N/A	N/A	Catastrophic - Torque will not be reduced [Dangerous] (12a)	N/A	N/A
13	**Limit air and feul Supply Signal**	Catastrophic - Torque will not be reduced [Dangerous] (13a)	N/A	N/A	Catastrophic - Torque will not be reduced [Dangerous] (13a)	N/A	N/A
14	**Switch Off Engine Signal**	Catastrophic - Torque will not be reduced [Dangerous] (14a)	N/A	N/A	Catastrophic - Torque will not be reduced [Dangerous] (14a)	N/A	N/A

Table 2: Causal factors of the identified hazards (Table 2 can be downloaded at http://serg.cs.lth.se/stpa)



No.	**Hazards**	**Severity**	**Causal Factors**
1a	System Dysfunction due to failure of Object detection system	Catastrophic	Object detection component failure (camera, radar or motion sensors) Communication error (no signal)
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)
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)
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)
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
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)
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)
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)
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)
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)
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
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)
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
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)
9c	Unwanted/Undesired auto braking due to False Braking signal	Moderate	Malfunctioning of brake controller due to incomplete process model (Generation of false signal)
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)
10b	False signal due to brake system malfunctioning [Application of automatic brakes with out need]	Moderate	Malfunctioning of brake system (generation of false signal)
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)
11b	System malfunctioning due to Missing Accelerator Signal	Catastrophic	Malfunctioning of sensor (Incorrect reading by sensor)
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)
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)
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)

Bibliography

[1] T. Abbas, L. Bernado, A. Thiel, C. Mecklenbrauker, and F. Tufvesson, “Radio channel properties for vehicular communication: Merging lanes versus urban intersections,” *IEEE Vehicular Technology Magazine*, vol. 8, no. 4, pp. 27–34, Dec 2013.

[2] J. V. Bond, G. H. Engelman, J. Ekmark, J. L. Jansson, M. N. Tarabishy, and L. Tellis, “Collision mitigation by braking system,” US Patent 6607255B2, 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 13:th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Sept 2010, pp. 155–160.

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

[5] ETSI-TR-102638, “Intelligent transport sytems (ITS), vehicular communications, basic set of applications, definitions,” Tech. Rep., V1.1.1, 2009.

[6] M. Gleirscher, “Hazard Analysis for Technical Systems,” in *proceedings of the 5:th International Conference on Software Quality, SWQD ’13, vol. 133 of LNBIP*. Springer Berlin Heidelberg, 2013, pp. 104–124.

[7] 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.

[8] 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.

[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 Berlin Heidelberg, 2004, pp. 130–141.

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

[11] 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.

[12] 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.

[13] R. E. McDermott, R. J. Mikulak, and M. R. Beauregard, *The Basics of FMEA, 2nd Edition*. Taylor & Francis, 1996.

[14] W. G. Najm, J. D. Smith, and M. Yanagisawa, “Pre-crash scenario typology for crash avoidance research,” Tech. Rep., Technical Report DOT HS 810 767, U.S. Department of Transportation Research and Innovative Technology Administration, Washington, DC, 2007.

[15] H. Nakao, M. Katahira, Y. Miyamoto, , and N. 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.

[16] 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.

[17] 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.

[18] 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.

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, 6, 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 3 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 7.

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, 6, 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 [7, 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

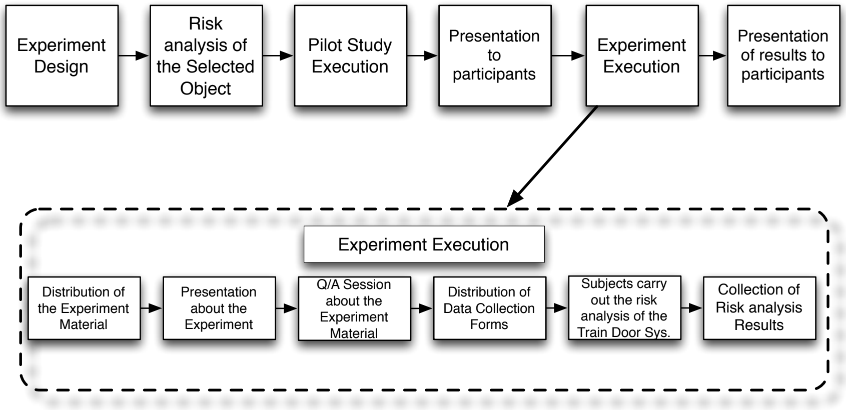

Experiment Design
Risk analysis of the Selected Object
Pilot Study Execution
Presentation to participants
Experiment Execution
Presentation of results to participants
Experiment Execution
Distribution of the Experiment Material
Presentation about the Experiment
Q/A Session about the Experiment Material
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/

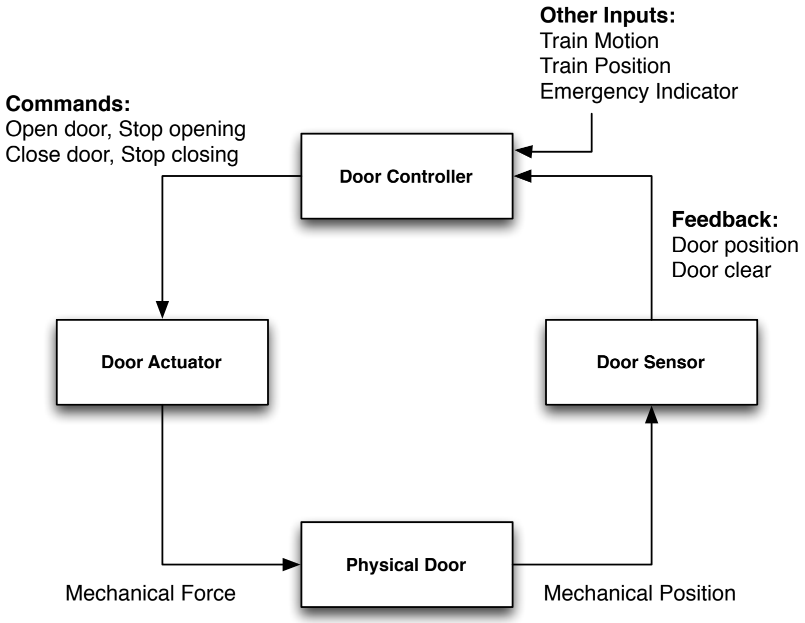


Figure 2: Functional diagram of a Train Door System [21]

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

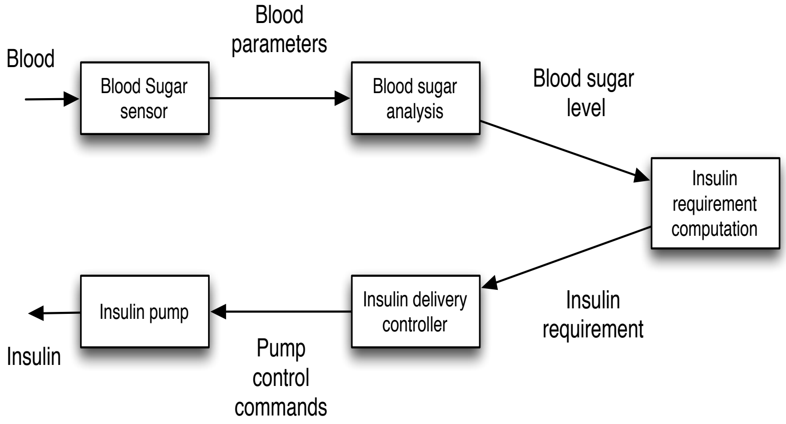


Figure 3: Functional diagram of an Insulin Pump [17]

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

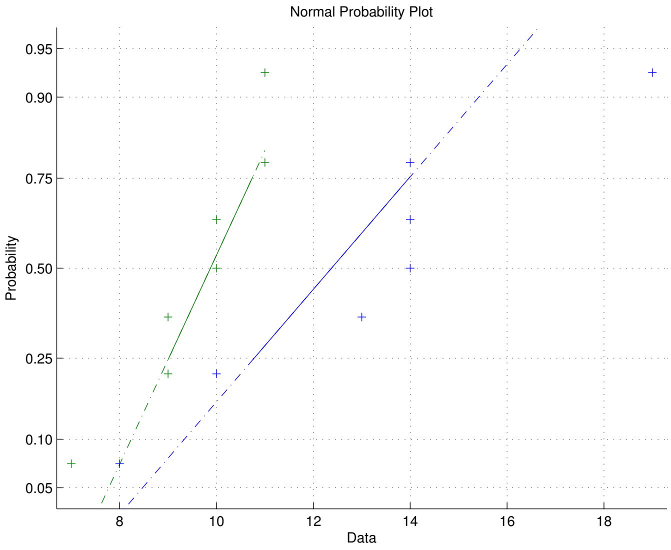


Figure 4: Normal distribution plot for the data

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

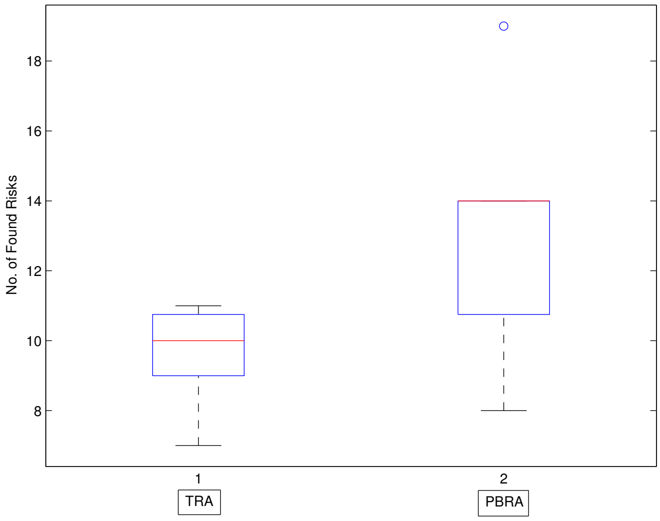


Figure 5: Box Plot of Found Risks

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, 6, 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. J. Alberts and A. J. 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, pp. 133–164, 1996.

[3] M. H. N. Boudriga 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] International Organization for Standardization, “ISO/IEC 27005:2011 - Information technology - Security techniques - Information security risk management,” 2011.

[7] 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.

[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 Berlin Heidelberg, 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 (USA)*, 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. E. McDermott, R. J. Mikulak, and M. R. 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*, Los Alamitos, CA, USA, 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 (first edition by Kluwer in 2000), 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

IDENTIFICATION OF IT INCIDENTS FOR IMPROVED RISK ANALYSIS

Abstract

Today almost every system or service, e.g., water, power supply, transportation, etc. 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. 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. Therefore, the risk analysis process needs to be improved and it is believed that by having an overview of already occurred IT incidents, the risk analysis process can be improved. The saved information about IT incidents can be used as an input to risk analysis, which can help to correctly estimate the consequences of potential risks. This study investigates how difficult it is to find relevant risks from the available sources and the effort required to set up such a system. It also investigates how accurate the found risks are. It presents a prototype solution of a system that automatically identifies information pertaining to IT incidents, from texts available online on Internet news sources, that have happened. This way IT incidents can be saved semi-automatically in a database and the saved information can be used later as an input to risk analysis. 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,
Submitted to a conference.

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 the important activities for the safe development of modern software-intensive systems, because there are few factors that make these systems more and more 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 [11].

Risk analysis and management are important activities for most of the project management tasks. Risk analysis and management in a software engineering context focus on the software development process and ensure its integrity. They try to ensure there is 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 [14].

There exist many, low and high level, risk analysis methods and frameworks that complements in identification and management of risks [17]. By using these methods and frameworks we can foresee the potential consequences of future possible IT incidents that later can be decreased or mitigated. A risk analysis includes a step where potential risks are identified, e.g. through "brain storming" activities. As a preparation to a step like this it can 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. To perform risk analysis the historical data of already happened unwanted events is required to correctly estimate the consequences of potential risks. However, historical data about such unwanted events is not easily accessible and it is not available at a single place. The saved information can also be useful for people working with IT-incident management, both researchers and practitioners, to have an overview of recent unwanted events that have occurred. 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. The approach is general and could be used to collect information about other topics, but the approach is off 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. The long-term objective is to be able to automatically identify relevant texts based on a person's particular field of interest, both in already available archives and in real time from newly published texts. If this is possible, it is possible to understand more about what type of IT incidents that have occurred, and, based on this, to improve our understanding of the type of IT incidents that are important to include in risk analyses.

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, 6, 13, 15, 16, 24]. This paper presents an explorative study on the practical steps necessary to set up a system, combining a number of well-established machine learning techniques, for automatic identification of online articles about IT incidents. It investigates how difficult it is to find relevant risks from the available sources and the effort required to set up such a system. It also investigates how accurate the found risks are.

The outline of this paper is as follows. Section 2 presents some relevant systems and reviews related work. Section 3 presents the background of machine learning tool and algorithms. In Section 4 the methodology and research questions are discussed. Section 5 presents the proposed method for automatic identification of IT incidents. Next, Section 6 presents the obtained results of the application and the evaluation of our proposed method. Finally, Section 7 summarizes the results of this study.

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 [18]. 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

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

[2] http://www.emm.newsbrief.eu

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. [16] present an overview of different available machine learning approaches for automatic text classification. In that study, the authors discuss different methods, their applications, their effectiveness and recent progress that has been made in the field.

The studies by Ikonomakis et al. [8], Yang [22], Yu [23], and Joachims [9] have performed automated text classification and compared the performance and accuracy of different machine learning algorithms such as Naive Bayes, Support Vector Machines (SVM), Decision Tree, k-Nearest Neighbors algorithm (k-NN) and Linear Least Square Fit (LLSF). They concluded that Naive Bayes and SVM performed well for text classification problems.

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.

3 Background

3.1 Naive Bayes

For the identification of IT incidents the Naive Bayes machine-learning algorithm was selected [21]. Naive Bayes is a classifying technique that predicts the probability of a document belonging to a certain class by first, for each selected feature, estimating the probability that a document containing this feature belongs to each class based on the occurrences in the training data. Then, by assuming independence between the features [12], the likelihood that a given document belongs to each class can be calculated based on the occurrence or non-occurrence of each of the features. The document can then be assigned to the class for which this value is the highest.

Although it makes the assumption of independence between features, which is obviously not the case in natural language, it performs well in many text classification and information filtering (e.g. spam filtering) applications [21].

Table 1: Confusion matrix

Actual	**Predicted Classes** Class-A (Documents of interest)	 Class-B (Documents not of interest)
Class-A	TP = *True Positive*	FN = *False Negative*
Class-B	FP = *False Positive*	TN = *True Negative*

3.2 Performance measures

Confusion matrix

Table 1 shows the "confusion matrix" that is used to calculate accuracy, precision and recall. TP (*True Positive*) is the number of documents correctly classified as interesting and FP (*False Positive*) is the number of not-interesting documents classified as interesting. In next column the FN (*False Negative*) is the number of actually interesting documents classified as not-interesting and TN (*True Negative*) is the number of not-interesting documents correctly classified as not-interesting. The performance measures that can be calculated directly from the confusion matrix are: [20]

Accuracy

This is the correct classification rate of a classifier. It is the ratio of the number of correctly classified documents to the total number of documents.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

Precision

This is also a correctness measure that only takes into account the documents classified as interesting. It is the ratio of the number of correctly classified documents as interesting to the total number of documents classified as interesting (correctly and incorrectly).

$$Precision = TP / (TP + FP)$$

Recall

This correctness measure focuses only on the documents that are actually of interest. It is the ratio of the number of correctly classified documents as interesting to the total number of documents that are actually interesting.

$$Recall = TP \, / \, (\, TP + FN \,)$$

ROC Curve

The ROC (receiver operating curve), as shown in Figure 3, illustrates the relative tradeoff between the benefits (true positive) and the costs (false positive) for different thresholds in the classification algorithm. It depicts only the threshold points used by classifier for classification. It shows the recall, or true positive rate, on the vertical axis and on the horizontal axis the false positive rate, the number of false positives (not-interesting documents incorrectly classified) divided by the total number of not-interesting documents.

$$False\ Positive\ Rate = FP \, / \, (FP + TN)$$

In Figure 3, the ROC plot is divided in two triangles. A ROC curve of a classifier in the upper left triangle shows good performance (better than random classifying) and a curve in the lower right triangle illustrates poor performance. A perfect classifier (100% accuracy) would be at the point (0,1) in the ROC plot [5].

3.3 The WEKA tool

For this study, machine learning and data pre-processing algorithms were applied through the WEKA[3] (Waikato Environment for Knowledge Analysis) data-mining 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. It is a complete data mining and machine learning solution that provides tools and algorithms for data pre-processing, classification, regression, clustering, association rules, and visualization [7].

4 Research methodology

4.1 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:

1. What search and identification methods should be used for this type of search?

[3] http://www.cs.waikato.ac.nz/ml/weka

2. Which steps are important to effectively use machine learning for searching and identification of IT incidents?

3. What kind of data sources should be used for the identification of IT incidents?

4. How much effort is required to perform this identification of IT incidents using in practice?

4.2 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 steps as shown in Figure 1.

In Step 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].

In Step 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.

In Step 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.

In Step 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.

Finally in Step 5, 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.

In Figure 1 and 2, it can be noticed that there is an overlap in few steps however both figures are different. Figure 1 shows the steps that are performed to carry out research presented in this study. Figure 2 shows the steps that are required to identify IT incidents or any other texts of interest. Since the proposed method in this study is general, it can be used to identify texts about any topic of interest.

5 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

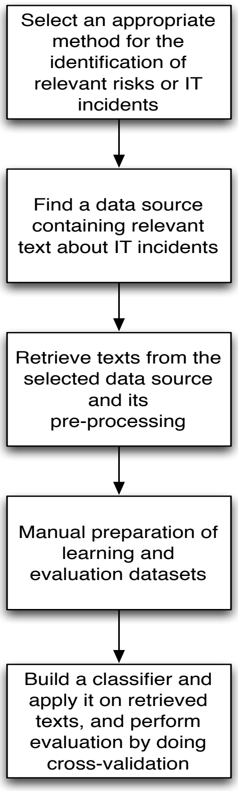

Select an appropriate method for the identification of relevant risks or IT incidents
Find a data source containing relevant text about IT incidents
Retrieve texts from the selected data source and its pre-processing
Manual preparation of learning and evaluation datasets
Build a classifier and apply it on retrieved texts, and perform evaluation by doing cross-validation

Figure 1: Research methodology

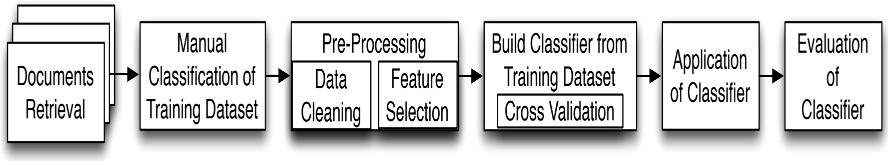

Documents Retrieval
Manual Classification of Training Dataset
Pre-Processing
Data Cleaning
Feature Selection
Build Classifier from Training Dataset
Cross Validation
Application of Classifier
Evaluation of Classifier

Figure 2: Identification process of IT incidents

techniques it is possible to classify or identify IT incidents automatically. It is important to mention that the proposed method is general and could be used to collect information about other topics. 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.
   (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 2.

5.1 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, the best source for this could be the e-news stories written by reporters from all over the world. For this the first step, data must be retrieved from the web source and then transformed into an appropriate form as required by the tools. To accomplish this task a software tool is needed that can traverse the web source and store the data into a database for further steps of IT incidents identification.

For the evaluation of the approach proposed in this study, we selected "The Risk Digest"[4] as data source. "The Risk Digest" 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. Every issue consists of around 15 selected posts on IT risks. 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 file in a format suited for the machine learning tool.

5.2 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 and these documents 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 from over 50 percent to less than 10 percent, which is more realistic in this type of information retrieval problem. 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. The classification used in this study could

[4] `http://www.catless.ncl.ac.uk/Risks/`

for example be useful for a safety manager working at an airport, wanting to stay informed about relevant incidents occurring at other airports.

5.3 Pre-Processing of data

Machine learning algorithms can extract useful information from a huge amount of available data semi-automatically or automatically. The semi-automatic extraction of information also known as supervised machine learning. In this, the training dataset should be labelled by class names for classification that could be used by machine to learn patterns or rules. The automatic extraction of information is called unsupervised machine learning that does not require labeling of text for classification. It determines classes or clusters of data by itself without knowing labels, it is also known as clustering. For the IT incident identification the semi-automatic machine learning technique was used. 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. Every unique word in each document will be a feature or attribute with the frequency of occurrence in the document as a value. Pre-processing is crucial to attain useful results [10]; 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. Stemming is a process of reducing the words to their basic form or root. It combines all the words that have the same morphological root and reduces them to one base form word. A stemmer first identifies all extended words and then converts them to their base form by removing their suffixes. For example, the words ‘describe’, ‘describes’, ‘described’, and ‘describing’ will all be stemmed to the word ‘describ’. Stemming further reduces the set of attributes. For the stemming, in this study the *Lovins* stemming algorithm was used [7].

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. This is an evaluation method that selects features that are highly correlated with the class attribute and having less or low inter-correlation.

5.4 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.

5.5 Application of classifier

Next, the classifier, built by using Naive Bayes mentioned in previous section, 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.

5.6 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.

6 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. In cross-validation, a classifier is trained based on only a part of the training set and evaluated on the remainder of the training set, but the selected features are based on the whole training set. Therefore this is different than evaluating the process on a completely separate evaluation set. In cross-validation the features for the training and evaluation are selected only from the training dataset, but in the evaluation from a completely separate dataset the features for the training and evaluation are selected from the different datasets (i.e., training and evaluation).

Table 2 and Table 3 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.

Table 2: 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%)	

- The results within parentheses are without carrying out stemming and stop words removal.

Table 2 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) and 3 documents incorrectly classified as not-interesting (FN). It also correctly classified 187 documents as not-interesting (TN) and one document incorrectly classified as interesting (FP).

Then, the cross-validation was performed by using 10-folds and classifier I obtained similar results as in the learning phase. The performance measures with stemming and stop words removal are: accuracy 98%, precision 90% and recall 75% and without stemming and stop words removal are: accuracy 97%, precision 71% and recall 83%. From these results it can be noticed that there is an increase in the accuracy, of 1%, after carrying out stemming and stop words removal.

Classifier I was applied on dataset Y, which correctly classified 4 documents as interesting (TP) and 8 documents incorrectly classified as not-interesting (FN). It also correctly classified 187 documents as not-interesting (TN) and 1 document incorrectly classified as interesting (FP).

Although there is an increase in the accuracy after performing stemming and stop words removal, the results are not good enough to be used for the proposed system. Because a good classifier for the IT-incident identification system will be one that has a high true positive rate. Classifier I performed well for correct classification of TN, but it incorrectly classified the interesting documents as not-interesting that is not good for the proposed system. Classifier I has a high accuracy, from evaluation, 95.5% with stemming and 94.5% without stemming.

Table 3 presents the results of classifier II built from dataset Y and then applied on dataset X for evaluation.

Table 3: 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%)	

Classifier II correctly classified 11 documents as interesting (TP) and one document incorrectly classified as not-interesting (FN). It also correctly classified 187 documents as not-interesting (TN) and one document incorrectly classified as interesting (FP). Here, an increase in accuracy compared to learning results of classifier I can be noticed. The performance measures for classifier II with stemming are: accuracy 99%, precision 91% and recall 91%.

After performing cross-validation, classifier II correctly classified 10 documents as interesting (TP) and 2 documents incorrectly classified as not-interesting (FN). It also correctly classified 184 documents as not-interesting (TN) and 4 documents incorrectly classified as interesting (FP). Here, a decrease in the precision compared to the results of classifier I can be noticed. The accuracy has also decreased, 97% from 98%, and this decrease is due to to the increase in false negatives (FN) but it correctly classified more, 10, documents as interesting. The cross-validation results (with stemming) of classifier II are very good for the IT incident identification system.

After applying classifier II on the evaluation dataset X, it correctly classified 7 documents as interesting (TP) and 5 documents incorrectly classified as not-interesting (FN). Classifier II also correctly classified 181 documents as not-interesting (TN) and 7 documents incorrectly classified as interesting (FP). Here, the increase in the numbers of TPs as compared to evaluation results presented in Table 2 can be noticed. The numbers of TNs have decreased and due to this there is an decrease in accuracy (94% from 95.5%) 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. [16] and Yu [23], stemming has both the positive and negative effects on accuracy for text classification results. Toman et al. [19] also mentioned that the stemming even decreases the accuracy of a text classifier. But in the results presented in Table 2 and Table 3, it can be noticed that after carrying out stemming and stop words removal both the built classifiers

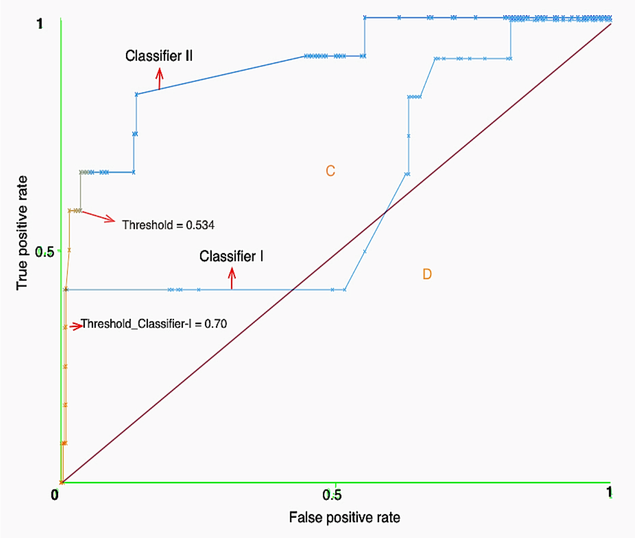


Figure 3: ROC Curve

performed well and obtained the results with an increase in accuracy as compared to without carrying out stemming and stop words removal.

Figure 3 shows the ROC curves for classifier I and classifier II applied on dataset Y and X for evaluation respectively after carrying out stemming and stop words removal presented in Table 2 and Table 3. As it can be noticed, the lower threshold curve of classifier I that has maximum accuracy at the point (.005, .333) with the threshold value 0.70. The ROC curve of classifier I passed from the lower right triangle D that is an indication of a not well performing classifier. The upper threshold curve of classifier II that performed well at the point (.03, .58) with the threshold value 0.534. It performed well by high rate of correct classification and low rate of incorrect classification. Classifier II in this study performed very well during evaluation and has a high accuracy 94% with stemming and 90.5% without stemming that is acceptable for the IT incident identification system.

7 Conclusions and Future work

This study presented a prototype solution of a system that automatically identifies and saves information pertaining to already happened IT incidents that can be used as an input to risk analysis and management. By having historical information of already happened IT incidents, risk analysis and management practices can be improved.

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 5 (see Figure 2). The proposed method worked 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.

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. It is also necessary to further investigate what happens when the relative amount of interesting articles is lower than in this study.

8 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] H. Eyal-Salman, A.-D. Seriai, and C. Dony, "Feature-to-code traceability in a collection of software variants: Combining formal concept analysis and information retrieval," in *proceedings of the 14:th IEEE International Conference on Information Reuse and Integration (IRI)*, Aug 2013, pp. 209–216.

[5] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006.

[6] 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.

[7] 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.

[8] M. Ikonomakis, S. Kotsiantis, and V. Tampakas, "Text classification using machine learning techniques," *WSEAS Transactions on Computers*, vol. 4, no. 8, pp. 966–974, 2005.

[9] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features," in *proceedings of the 10:th European Conference on Machine Learning*, 1998, pp. 137–142.

[10] 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.

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

[12] D. D. Lewis, "Naive (bayes) at forty: The independence assumption in information retrieval," in *proceedings of the 10:th European Conference on Machine Learning*. Springer Verlag, 1998, pp. 4–15.

[13] 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.

[14] G. G. Roy, "A risk management framework for software engineering practice," in *proceedings of the Australian Software Engineering Conference*, 2004, pp. 60–67.

[15] 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.

[16] F. Sebastiani and C. N. Ricerche, "Machine learning in automated text categorization," *J. of ACM Computing Surveys*, vol. 34, pp. 1–47, 2002.

[17] 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.

[18] 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.

[19] M. Toman, R. Tesar, and K. Jezek, "Influence of word normalization on text classification," in *proceedings of Multidisciplinary Approaches to Global Information Systems*, 2006.

[20] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques, Third Edition (The Morgan Kaufmann Series in Data Management Systems)*, 3rd ed. Burlington, MA: Morgan Kaufmann, Jan. 2011.

[21] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, "Top 10 algorithms in data mining," *Internation Journal of Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, 2007.

[22] Y. Yang, "An evaluation of statistical approaches to text categorization," *Journal of Information Retrieval*, vol. 1, pp. 67–88, 1999.

[23] B. Yu, “An evaluation of text classification methods for literary study,” Ph.D. dissertation, University of Illinois at Urbana-Champaign, Champaign, IL, USA, 2006.

[24] 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.