jamk.fi

Implementing a Quantitative Risk Management Methodology in a Cyber Exercise

Petteri Nakamura

Master’s Thesis
October 2020
School of Technology, Communication, and Transport
Master’s Degree Programme in Information Technology
Cyber Security

Jyväskylän ammattikorkeakoulu
JAMK University of Applied Sciences

jamk.fi

Kuvailulehti

Tekijä(t) Nakamura, Petteri	Julkaisun laji Opinnäytetyö, Ylempi AMK	Päivämäärä 10 2020
		Julkaisun kieli: Englanti
	Sivumäärä 73	Verkkojulkaisulupa myönnetty: x

Työn nimi

Kvantitatiivisen riskinhallintamenetelmän implementointi kyberharjoituksessa

Tutkinto-ohjelma

Information Technology

Työn ohjaaja(t)

Hautamäki Jari, Saharinen Karo

Toimeksiantaja(t)

Nakamura Petteri

Tiivistelmä

Kvalitatiiviset riskinhallintamenetelmät ovat yleisesti käytössä kyberturvallisuusalalla, siitä huolimatta, että tutkimukset osoittavat useita lähestymistavasta johtuvia ongelmia. Lähtökohtaisesti kvantitatiiviset menetelmät perustuvat kvalitatiivisia menetelmiä vankemmin tieteelliseen tutkimukseen ja mahdollistavat suuremman joustavuuden riskianalyysissä. Ne myös mahdollistavat sekä yksittäisten riskien, että myös kokonaisten riskiportfolioiden rahallisen arvon arvostuksen. Riskien arvon ilmaiseminen rahayksiköissä taas mahdollistaa paremman viestinnän riskianalyytikoiden ja päättäjien välillä kuin kvalitatiiviset ilmaisut, kuten "matala", "keskitasoinen" tai "korkea" tai kvalitatiiviset asteikot esimerkiksi yhdestä viiteen.

Kirjallisuusanalyysi tehtiin vastaavien kvantitatiivisten menetelmien käyttöönottoon liittyvien tutkimusten löytämiseksi sekä nykyisin tarjolla olevien menetelmien löytämiseksi. Kaksi saman tyyppistä menetelmää erottui tästä massasta; Factor Analysis of Information Risk (FAIR) ja Hubbard and Seiersen -menetelmä. Jälkimmäinen valittiin tapaustutkimukseen, jossa tutkittiin kvantitatiivisen lähestymistavan toteuttamista kyberharjoituksen yhteydessä kevään 2019 aikana. Kyberharjoitusta käytettiin simuloimaan pientä organisaatiota, jotta voitiin selvittää, onko tällaiselle organisaatiolle mahdollista käyttää Kvantitatiivista lähestymistapaa kvalitatiivisen sijasta. Tapaustutkimuksen toisena tavoitteena oli myös saada kokemusta valitun kvantitatiivisen lähestymistavan toteuttamisesta ja käytöstä.

Hubbardin ja Siersenin menetelmä osoittautui käyttökelpoiseksi harjoituksessa. Käytännön työkalu kehitettiin kvantitatiivisen mallin jatkuvaksi päivittämiseksi harjoituksen edetessä kerätyn datan avulla. Kvantitatiivisen menetelmä todettiin käyttökelpoiseksi myös pienissä organisaatioissa ja ideoita mahdolliseen jatkotutkimukseen esitettiin.

Avainsanat (asiasanat)

Kvantitatiivinen riskianalyysi, Monte Carlo, Beta-distribuutio

Muut tiedot (salassa pidettävät liitteet)

jamk.fi

Description

Author(s) Nakamura, Petteri	Type of publication Master’s thesis	Date 10 2020
		Language of publication: English
	Number of pages 73	Permission for web publication: x

Title of publication

Implementing a Quantitative Risk Management Methodology in a Cyber Exercise

Degree programme
Information Technology

Supervisor(s)
Hautamäki Jari, Saharinen Karo

Assigned by
Nakamura Petteri

Abstract

Qualitative risk management methods are in common use in the field of cyber security despite research showing various inherent problems with the qualitative approach. The premise is that quantitative methods are more grounded in research than qualitative ones, allow for greater flexibility in the analysis of risk, and enable better expression of risk in terms of monetary valuations for individual risks and portfolios containing multiple risks. The expression of the value of risks in monetary units allows for better communication between risk analysts and the decision makers than qualitative expressions such as “low”, “medium”, or “high” or qualitative scales of for example from one to five.

A literature analysis was carried out to seek for similar works on the implementation of quantitative methods and the current prominent methodologies were inspected. Two similar methodologies stood out; Factor Analysis of Information Risk (FAIR) and Hubbard and Seiersen method. One of the two prominent methodologies, the Hubbard and Seiersen method, was selected for a case study of testing implementation of a quantitative approach to risk analysis in the context of a cyber exercise, carried out during spring 2019. The cyber exercise was used as a surrogate for a small organisation in order to find out if using a quantitative approach instead of a qualitative one is feasible for a small organisation. The second objective of the case study was also to gain experience in implementing and using the selected quantitative approach.

Hubbard and Siersen method proved feasible to use in the exercise and a practical tool for continuously updating the quantitative model with data gathered during the exercise was developed. Implementing quantitative risk analysis methodology was found to be feasible also in small organisations and ideas for further research were presented.

Keywords/tags (subjects)
Quantitative risk analysis, Monte Carlo, Beta Distribution

Miscellaneous (Confidential information)

Contents

Figures

Tables

1 Introduction

Risk Management is a discipline that endeavours to recognize and mitigate the likelihoods and effects of such events on an organization. Risk management itself is a very old discipline spanning different fields requiring investments on uncertain outcomes. According to Rhodes (2015), some historians believe the earliest concepts of managing risk being traceable to ancient civilizations playing games with dice and bones. Some evidence of gaming giving rise to probability theory, which is important for risk management, comes from writings by Dante and Galileo, and later Pascal and Fermat corresponded about games of chance in the 17th century. This is believed to have given rise to the modern probability theory. First professional measurers of risk and uncertainty were actuaries in England working as corporate risk managers in the 18th century (Rhodes 2015).

Corporate risk management recognizes and needs to address multiple types of risks in the present-day environments. For example, Wolke (2017) categorizes risk into financial risks, further divided into market price risks, default risks, and liquidity risks, and performance risks, further divided into operational risks and sales and procurement risks.

Compared to risk management, cyber security is a very new field, which emerged during the latter half of the 20th century after the advent of internet. According to Matthews (N.d.), mathematician John von Neumann predicted the idea of a computer virus before computer networks even existed, but the first virus was only created some 30 years later in 1971 during the age of ARPANET. ARPANET was the predecessor of the modern internet and its early users were researches who trusted each other. Therefore, setting up security measures in the network was not a top concern before the first virus (Matthews N.d.). According to Townsend (N.d.) the first denial-of-Service Attack took another 18 years to occur after the first virus, when a worm, written by Robert Morris, slowed down the early internet significantly.

Responses to the perceived threats began with the first patent for protecting communications in the network which was granted to MIT in 1983 for a "cryptographic communications system and method" describing the RSA algorithm (Matthews N.d.). Legislators began fighting against cybercrime with one of the first acts in the United Kingdom in 1990, when The Computer Misuse Act made illegal any unauthorized attempts to access computer systems (Townsend N.d.). The DEF CON cybersecurity technical conferences were established in June of 1993. (Townsend N.d.).

Over the last five decades, computers and software have permeated personal and professional lives in developed countries to the point where it is hard to imagine any profession that in today's post-industrialized information societies would not be in some way dependent on information technology. According to Pensworth (2020), As of January 2019 the number of active internet users totals almost 4.4 billion, equivalent to 57% of the global population of the planet. Asia has the most internet users with over 2 billion users in 2018, followed by Europe with 700 million internet users in 2018 (Pensworth 2020). This number is expected to grow to 5.3 billion by 2023, with the India and Africa catching up to the Asia and the west (*Cisco Annual Internet Report (2018-2023) White Paper* 2020). Even money and payments are experiencing digitalization, with governments and central banks already experimenting with digital currencies in addition to government and central bank independent crypto currencies, such as China planning to rollout e-yuan by 2022 and Sweden planning to test e-krona in 2021 (Hackett 2020).

With the continuous advancement of information technology, organizational entities and nation states form a deeply intertwined and often fragile environment whose normal operation can easily be disturbed by a cyber-attack by a nation state against another, corporate espionage and sabotage or even something as simple as a failed hard drive. The effects of these events can cause trouble to anywhere from a single company to a large number of organizations, and the effects will cascade to the general population in case vital infrastructure becomes unavailable. The first recorded cyber-attack on critical national infrastructure at the Trans-Siberian pipeline in 1982 resulted in an explosion visible from space (Cherdantseva, Burnap, Blyth, Eden, Jones, Soulsby, & Stoddart 2016). In a more recent example, many companies around the world fell collateral victims to the NotPetya ransomware in June 2017, which Western intelligence agencies deemed a creation of the GRU, Russia's military intelligence agency, and an apparent attack on Ukraine (Greenberg 2018). NotPetya forced shipping company Maersk to rebuild their entire IT infrastructure from scratch to recover from a complete IT-environment encryption and crippled Merck & Co.'s production facilities to the extent where the company could not meet the yearly demand for Gardasil9 human papillomavirus vaccine against cervical cancer (Greenberg 2018; Voreacos, D., Chiglinsky, K., & Griffin, R. 2020).

Any company has a risk of falling victim to a data breach in the modern world (IBM Security 2019). However, the magnitude of this risk appears not to always be well understood by decision makers, based on the results of a survey conducted by Helsinki Chamber of Commerce in 2019 on cyber threats facing Finnish companies. Helsinki Chamber of Commerce asked 600 business leaders about their investments in cyber security during the last four years, 36 percent of the respondents stated that they had not done any investments in cyber security during the last four years and 14 percent said that they did not know if they had or not (*Corporation Targeted Cyber Threats* 2019). Therefore, the total percentage of companies that had not accounted for cyber security in any way in their budgets was 50% (ibid.). Smaller companies were disproportionately represented in this number, which also presents risks of supply chain attacks to larger companies with higher level of cyber security investments, as smaller companies often have some levels of access to larger companies' networks through subcontracting and other arrangements. At the same time the global probability for a company to experience a data breach within two years has steadily increased by 31 percent from 22.6 percent in 2014 to 29.6 percent in 2019, shown in Figure 1 below (IBM Security 2019).

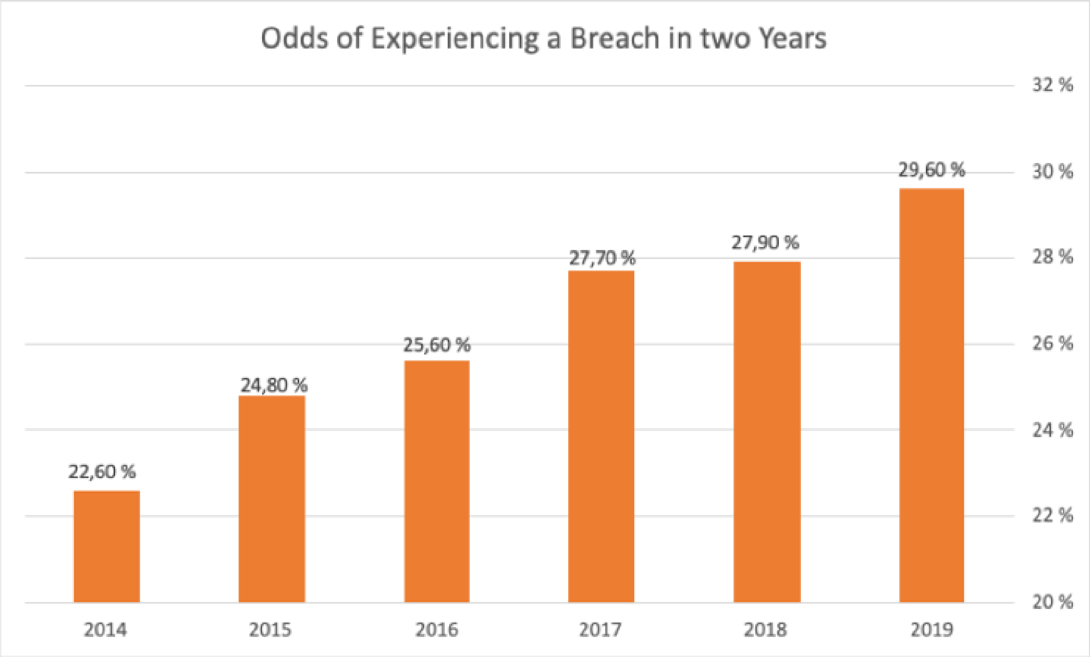


Figure 1: Probability of a company experiencing a data breach within two years (IBM Security 2019)

IBM Security (2019) also shows that the relative costs of a data breach are higher for smaller companies than for larger ones. The report also shows that being prepared is the greatest factor contributing to lower than average costs after a data breach, with effective preparation actions including the formation of an incident response team, use of encryption, testing of the incident response plan, implementation of business continuity management, training of employees and involvement of board members in cyber security aspects (IBM Security 2019). In November 2013, Federal Bureau of Investigation (FBI) director James B. Comey testified before the Senate Homeland Security and Governmental Affairs Committee that “we anticipate that in the future, resources devoted to cyber-based threats will equal or even eclipse the resources devoted to non-cyber-based terrorist threats” (Miller 2013).

Therefore, the current trend seems to be that while cyber security risks are increasing constantly, half of Finnish companies do not understand or recognize them. This could be remedied with research on the true level of risk for Finnish companies, but the main problem for this is scarceness of data. During his Q&A session after a presentation on risk management at Elisa ICT Day 2019, Tuomas Miettinen, Solution Consultant at F-Secure, noted that estimating the real risk of a data breach is not possible due to lack of statistics in Finland, and would require more openness from companies in reporting data breaches. He also noted that the European General Data Protection Regulation should also provide more data for this in the coming years. The Finnish National Cyber Security Centre publishes monthly reports of cyber threat situation in Finland; however, these Cyber Weather reports mostly give verbal descriptions, and their system for showing the relative severity of the situation uses a three-step ordinal scale that is impossible to use for any quantitative analysis or for making deductions about the risk for any specific company.

Refsdal, Solhaug and Stolen (2015) note that the cybersecurity strategies in the European Union and nations worldwide are pushing for organizations in various fields to ensure that cyber risk is managed appropriately. Cyber risks are no longer only an issue for IT professionals as incidents and financial impacts continue to soar (Refsdal, A. et.al. 2015).

Risk management lies in the very heart of Cyber Security, as according to SFS-ISO 27001 (2011), risk assessment should be the basis with which investment decisions for mitigations are made. Hubbard and Seiersen (2016) note that even though a vast body of research shows that no evidence exists that qualitative methods actually help in reducing risk, there is plenty of research showing the opposite for quantitative methods. The authorities in Cyber Risk Management, such as the National Institute of Standards and Technology (NIST), the International Standards Organization (ISO), MITRE, and the Open Web Application Security Programme (OVASP) all seem to promote some version of qualitative risk scores for assessing the magnitudes of such risks in their frameworks (Hubbard, D. W. & Seiersen, R. 2016).

Going back to the financial and actuarial world, the two have a long and somewhat parallel history in risk management. Whelan (2002) writes that many applied statistics and probability discoveries can be traced back to 19th century actuaries, but during the 20th century the actuaries developed their own jargon and failed to communicate their discoveries in practical application of probability theory with a wider audience. As a result, when other fields, such as finance, began to use probability theory, they often rediscovered the results themselves instead of finding them in the existing highly field-specific technical literature. (Whelan 2002.)

Currently the same rediscovering of the wheel seems to happen in Cyber Security, but instead of pursuing mathematically proven statistical methods in risk management, the field is strongly inclined to use unproven qualitative methods (Hubbard, D.W. & Seiersen, R. 2016). Quantitative methods are mentioned, for example, in SFS-ISO/IEC 27000 (2011), but these methods are not given much weight in the standard. Without referring to the unproven nature of the qualitative methods, SFS-ISO/IEC 27000 (2011) states that “qualitative analysis is often used first to obtain a general indication of the level of risk and to reveal the major risks” and continues to state that a quantitative analysis may be necessary to undertake for the major risks due to quantitative analysis being more complex and expensive to undertake (SFS-ISO/IEC 27000, 2011).

Freund and Jones (2015) states that risk analysis methods need to be useful, practical and the results need to be defendable. In order to improve risk management in the field of cyber security and to give business decision makers better useful actionable information and analysis results to support their decision making, cost-effective scientifically proven methods and tools need to be provided for cyber security and IT professionals working with risk management in companies of various sizes. These methods and tools must not be significantly more difficult to implement and use than the current de-facto methods, while simultaneously creating significantly better value. Risk assessment results should be provided in monetary units, that are easy for decision makers to understand, compare, and to discuss things such as returns on investments for risk mitigations or determining if a company’s insurance coverage is in line with the risk retained by the company. The utilized methods also need to be transparent so that they are defendable when challenged (Freund, J. & Jones, J. 2015).

2 Research Setting

2.1 Research Problem

As ever-increasing digitalization and interconnectivity and complexity of systems cause the cyber risks to increase every year in relation to the revenues of the companies, the required investments are similarly increasing in significance (IBM

Security 2019). Therefore, quantitative methods to determine whether investments in cyber security are in line with the severity of the risks also need to develop and be implemented. While quantitative risk analysis methods are argued to be more effective than qualitative methods by Hubbard and Seiersen (2016), searches into research articles and master's theses yield very few papers on implementing quantitative methods in small and medium size environments.

An empirical analysis of ten currently available cyber risk frameworks, methodologies, systems and models found that only three out of the ten analysed systems computed risk quantitatively, and called for use of quantitative methods in assessing the economic impact of cyber risks (Radanliev, P., De Roure, D. C., Nurse, J. R., Montalvo, R. M., Cannady, S., Santos, O., Maddox, L., Burnap, P. & Maple, C. 2019).

The research problem behind this thesis revolves around implementing quantitative methods in fields in which qualitative methods are considered the simple de facto methods, while quantitative methods are considered too difficult or impractical to use. This disposition is portrayed by Ammar, Berman, and Sataporn, commenting in their research paper that data for a quantitative approach is often hard to come by due to lack of historical data for statistical analysis, and therefore a qualitative approach is often appropriate for risk analysis (Ammar, A, Berman, K, & Sataporn A, 2017). Similarly, SFS-ISO/IEC 27000 (2011) introduces quantitative methods as an extension to a qualitative analysis, for deeper analysis of only a subset of risks. However, Hubbard and Seiersen (2016) maintain that a quantitative analysis should always be used in favour of qualitative alternatives due to strong evidence of their effectiveness and continue that the amount of information needed for informative quantitative risk analyses is often much less than expected. Hubbard and Seiersen (2016) also demonstrate a method to generate calibrated quantitative estimates of uncertainty using subjective expert estimation.

2.2 Research Question

The main objective of this work was to challenge the notion that performing quantitative risk analysis is more complex and expensive than qualitative risk analysis, and to show that using a quantitative approach from the start, while giving major advantages over the common qualitative methods, is not significantly more expensive to implement.

The research question therefore was, is it feasible for a small company, considering an IT Risk Management model, to follow a quantitative approach in favor of a qualitative one and what kinds of advantages, based on the experiment, would there be. A secondary objective was to gain experience in setting up and using a quantitative approach to risk analysis and management in the context of a small organization with limited resources to allocate to the task.

Feasibility of implementation in this scenario means that the quantitative approach must not be significantly more difficult or costly to implement than the common qualitative alternatives readily available to small and medium size enterprises. If the model is possible to set up in a three months long cyber exercise by a simulated organization using the same resources that other similar teams used to set up qualitative methods, then the method can be seen as feasible and comparable to the qualitative methods in regards to the required setup effort and costs.

2.3 Research Method

The chosen research question did not easily yield itself to quantitative research methods, as this would have required experimenting with different quantitative and qualitative methods in similar environments and creation of a feasibility score which could be compared between different methods. However, the cyber exercise used as a base for implementing and testing the quantitative implementation. had four blue teams that all were required to create their respective risk analyses. If different teams are able to create their respective qualitative and quantitative analyses with the same resources during the exercise, then by the subjective definition of feasibility given above, the described quantitative approach is feasible relative to the qualitative approach.

On the other hand, the aim of the work was also to demonstrate how to set up a quantitative methodology and use the methodology to gain knowledge. A development study as research methodology was first considered, as the research methodology for the thesis is indicative of a development study as described by Kananen (2014). However, the end result of the study is a better understanding of the phenomenon under study based on the experiences gained in a simulated environment, instead of an implemented system being used in a real environment, and as a result a case study was deemed to be a better option.

Therefore, the chosen research method is a qualitative case study, in which the objective is to produce deep understanding of the selected case (*Tapaustutkimus [Case Study]* 2015).

In this scenario the case under research consists of implementing a quantitative risk analysis model in a small organization during a simulated cyber exercise in order to gain experience in such an undertaking, and to answer the research questions using the results of the implementation.

2.4 Research Setting

In qualitative research, the research target is not understood, and therefore the objective is to gain understanding of the phenomenon under research (Kananen 2014). The target phenomenon in this case is the quantitative risk management model and its use, adapted from the example model of Hubbard and Seiersen (2016). The setting was a cyber security exercise during the first quarter of 2019, jointly performed between three classes of students from the JAMK University of Applied Science in Jyväskylä and the University of Jyväskylä. The aim of the exercise was to simulate cyber-attacks against mock environments in an enclosed internet environment, complete with all the services available in the normal internet, such as Domain Name Servers, E-mail services, search engines, internet banking services and social media. The students were divided into four blue teams, which simulated different types of companies and set up their respective corporate environments, including internal hierarchies, and operating and risk management procedures. In addition, one red team was assigned to attack against the blue team environments, and a white team was assigned to administer the game. The game itself was conducted after an intensive setup phase during two separate weekends in Jyväskylä, providing a good opportunity to test implementation and use of a quantitative risk management methodology in a sped up simulated environment, making it possible to quickly acquire the experience required by the case study.

3 Term Definitions

Beta Distribution

A distribution of probability distributions that can be used to evaluate population proportions.

Blue Team

A team in a cyber exercise tasked for maintaining and defending a system.

Cyber Exercise

An exercise for practicing and gaining experience in managing adverse situations in maintaining and defending an IT environment against attacks and infiltration attempts.

Key Performance Indicator

A performance measurement used to evaluate the success of an organization or a particular activity.

Monte Carlo simulation

A method to simulate uncertain outcomes by running thousands of simulations using random values.

Ordinal Scales

A type of scales of measurement that only specifies the order of the items within a set but not the distance between the items. For example, “first”, “second”, “third”, or “low”, “medium”, “high”.

Qualitative Risk Analysis

A family of risk management methodologies that are based on expressing risk in qualitative terms such as “low”, “medium”, or “high”, or using ordinal scales.

Quantitative Risk Analysis

A family of risk management methodologies that are based on expressing risk in quantitative terms such as percentages or ranges of monetary values.

Red Team

A team in a cyber exercise tasked for attacking or infiltrating a system.

Risk Portfolio

A portfolio describing the recognized risks within a project or facing a company.

Risk Tolerance

Also called Risk Appetite. An individual's willingness to bear risk expressed in either qualitative or quantitative terms.

White Team

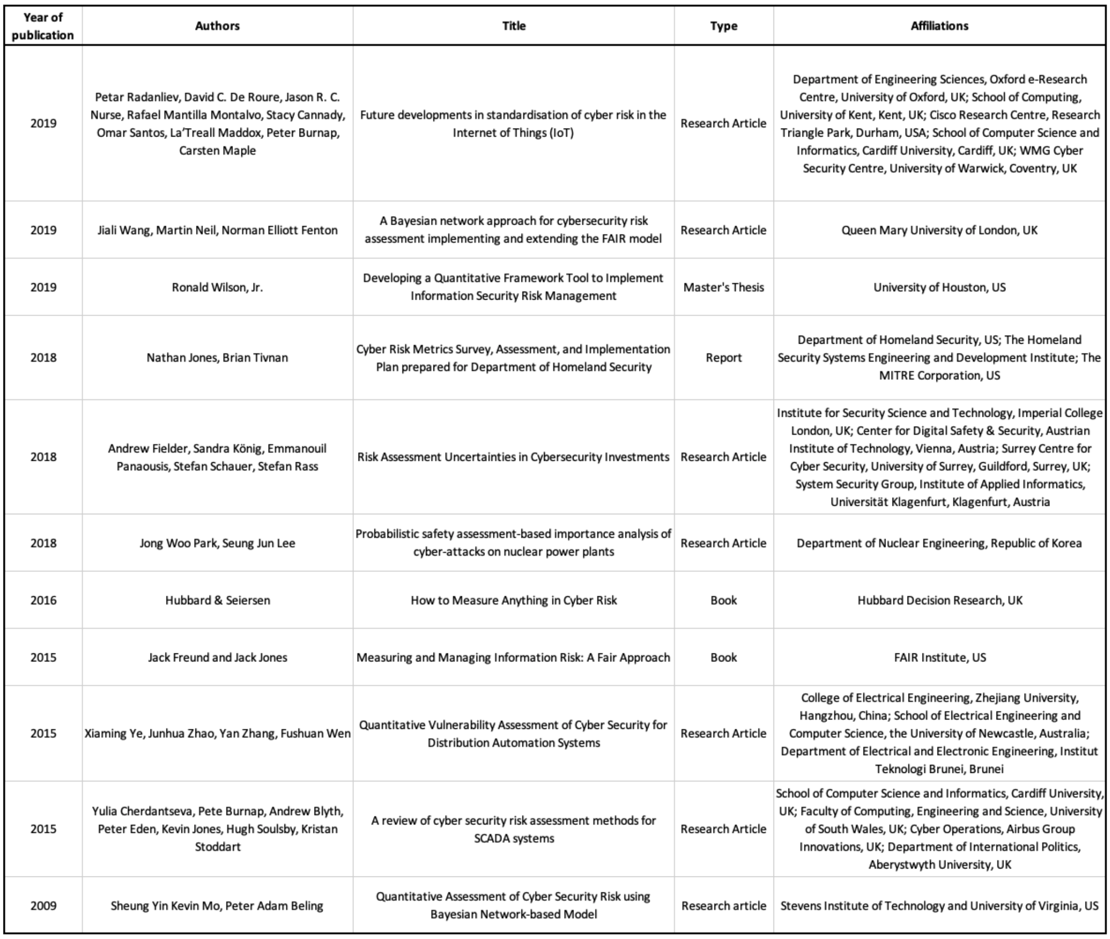
A team in a cyber exercise tasked for managing the exercise.

4 Theoretical Basis

4.1 Literature Review

Literature review was conducted by assessing several articles and books published on advancements in quantitative risk management methods during the past decade, the contents of these resources are briefly described here. Table 1 shows a summary of the literature reviewed in descending order based on the year of publication. This is not intended to be an exhaustive list of the publications in the field, but rather a subset of the publications and resources. Based on the literature review, an impression emerges of a growing trend in the recent years of interest in quantitative risk analysis and management methods in the field of cyber security from researchers, organizations from different fields, and regulators alike

Table 1: List of recent publications on quantitative risk management



Year of publication	Authors	Title	Type	Affiliations
2019	Petar Radanliev, David C. De Roure, Jason R. C. Nurse, Rafael Mantilla Montalvo, Stacy Cannady, Omar Santos, La'Treall Maddox, Peter Burnap, Carsten Maple	Future developments in standardisation of cyber risk in the Internet of Things (IoT)	Research Article	Department of Engineering Sciences, Oxford e-Research Centre, University of Oxford, UK; School of Computing, University of Kent, Kent, UK; Cisco Research Centre, Research Triangle Park, Durham, USA; School of Computer Science and Informatics, Cardiff University, Cardiff, UK; WMG Cyber Security Centre, University of Warwick, Coventry, UK
2019	Jiali Wang, Martin Neil, Norman Elliott Fenton	A Bayesian network approach for cybersecurity risk assessment implementing and extending the FAIR model	Research Article	Queen Mary University of London, UK
2019	Ronald Wilson, Jr.	Developing a Quantitative Framework Tool to Implement Information Security Risk Management	Master's Thesis	University of Houston, US
2018	Nathan Jones, Brian Tivnan	Cyber Risk Metrics Survey, Assessment, and Implementation Plan prepared for Department of Homeland Security	Report	Department of Homeland Security, US; The Homeland Security Systems Engineering and Development Institute; The MITRE Corporation, US
2018	Andrew Fielder, Sandra König, Emmanouil Panaousis, Stefan Schauer, Stefan Rass	Risk Assessment Uncertainties in Cybersecurity Investments	Research Article	Institute for Security Science and Technology, Imperial College London, UK; Center for Digital Safety & Security, Austrian Institute of Technology, Vienna, Austria; Surrey Centre for Cyber Security, University of Surrey, Guildford, Surrey, UK; System Security Group, Institute of Applied Informatics, Universität Klagenfurt, Klagenfurt, Austria
2018	Jong Woo Park, Seung Jun Lee	Probabilistic safety assessment-based importance analysis of cyber-attacks on nuclear power plants	Research Article	Department of Nuclear Engineering, Republic of Korea
2016	Hubbard & Seiersen	How to Measure Anything in Cyber Risk	Book	Hubbard Decision Research, UK
2015	Jack Freund and Jack Jones	Measuring and Managing Information Risk: A Fair Approach	Book	FAIR Institute, US
2015	Xiaming Ye, Junhua Zhao, Yan Zhang, Fushuan Wen	Quantitative Vulnerability Assessment of Cyber Security for Distribution Automation Systems	Research Article	College of Electrical Engineering, Zhejiang University, Hangzhou, China; School of Electrical Engineering and Computer Science, the University of Newcastle, Australia; Department of Electrical and Electronic Engineering, Institut Teknologi Brunei, Brunei
2015	Yulia Cherdantseva, Pete Burnap, Andrew Blyth, Peter Eden, Kevin Jones, Hugh Soulsby, Kristan Stoddart	A review of cyber security risk assessment methods for SCADA systems	Research Article	School of Computer Science and Informatics, Cardiff University, UK; Faculty of Computing, Engineering and Science, University of South Wales, UK; Cyber Operations, Airbus Group Innovations, UK; Department of International Politics, Aberystwyth University, UK
2009	Sheung Yin Kevin Mo, Peter Adam Beling	Quantitative Assessment of Cyber Security Risk using Bayesian Network-based Model	Research article	Stevens Institute of Technology and University of Virginia, US

4.1.1 Future Developments in Standardisation of Cyber Risk in the Internet of Things (IoT)

Radanliev and others' (2019) focus on quantifying economic impact of IoT cyber risk concludes that hybrid and interdisciplinary approaches are required to design cyber risk assessments, which include the economic impact of IoT related risks. Fast growth of IoT has led to a situation where finance and insurance markets lack empirical data to create actuarial tables, and models are lacking in relation to IoT risks. As a result, banks and insurance companies are unable to price IoT related cyber risks with the same precision as they price traditional insurance lines, and the current macroeconomic cost estimates of IoT related cyber risks are entirely speculative. To improve the situation, Radanliev et. al. (2019) propose a four-step epistemological framework for standardization of cyber risk impact assessment, with recommendations of tools to use in "Measurement", "Standardization", "Computing", and "Recovering". Radanliev et al's (2019) recommendation for computing the economic impact is to use quantitative risk analysis with Monte Carlo simulations and sensitivity analysis, but the implementation is out of the scope of their research. They note that there are currently two leading quantitative cyber risk models, RiskLens and Cyber VaR, which should be used in the estimation of economic impact of cyber risk from IoT devices. (Radanliev et. al. 2019).

4.1.2 A Bayesian Network Approach for Cybersecurity Risk Assessment Implementing and Extending the FAIR Model

Wang, Neil and Fenton (2019) found the FAIR model to have a number of limitations due to the types of distributions and the use of cached data in the calculations of the model. They therefore created two extensions of the FAIR model with Bayesian Networks and replaced the cached data and built-in distributions with Monte Carlo simulations, calling the extensions FAIR-BN and FAIR-MC respectively, and compared the performance of the three models in computing quantified risks. They conclude that they were able to improve the limitations of the FAIR model to overcome the found difficulties, and also give suggestions to the kinds of situations in which to use each of the three variants. (Wang et. al. 2019).

4.1.3 Developing a Quantitative Framework Tool to Implement Information Security Risk Management

In his Master's Thesis, Wilson (2019) presents a framework for estimating risk of a cyber-attack for use by small and medium sized businesses, defined as businesses with less than 1000 employees. Wilson (2019) used data from the Verizon Data Breach and Ponemon Institute Cost of Data Breach reports as inputs to Bow Tie risk analysis model in order to enable SME's to calculate expected annual losses from data breaches and create cost/benefit analyses (Wilson, 2019). Wilson’s (2019) solution avoids subjective expert estimates by using data from the two reports, but it only considers data breaches, leaving other types of ICT related risks out.

4.1.4 Cyber Risk Metrics Survey, Assessment, and Implementation Plan prepared for Department of Homeland Security

Jones (2018) surveyed the cyber security metrics in use in the Financial Service Sector in the United States in an effort to begin standardizing quantitative cyber risk metrics for the sector. The work was done in order to be able to aggregate risk reporting and information flow across sub-sectors to enable financial sector institutions to respond to cyber threats more effectively than possible at the moment. Jones (2018) also reviewed the available quantitative models against the requirements, presented in the study, and ended up recommending two existing approaches: “FAIR” and “Hubbard and Seiersen Approach”.

4.1.5 Risk Assessment Uncertainties in Cybersecurity Investments

Fielder, König, Panaousis, Schauer, and Rass (2018) used game-theoretical simulations to test the coverage of different levels of budgets for cyber security controls in relation to the uncertainty of the effectiveness of the controls. The purpose was to investigate different strategies to allocate investments to cyber security controls to receive the best possible coverage for the least amount of money (Fielder et al, 2018).

4.1.6 Probabilistic Safety Assessment-Based Importance Analysis of Cyber-Attacks on Nuclear Power Plants

Park and Lee (2018) analyzed various possible cyber-attacks on nuclear power plants in response to the current risk analyses in the field, including risks involving unintended events such as malfunctions or operator errors, but not intended events like cyber-attacks. Park and Lee (2018) extended the Event and Fault Tree based Probabilistic Safety Assessments (PSA) currently in use, with including intentional failures in different categories in the assessment model (Park, J. & Lee, S. 2018).

4.1.7 How to Measure Anything in Cyber Risk

In their book Hubbard and Seiersen (2016) examine the different problems of using qualitative risk analysis methods and suggest using quantitative methods instead, referring to the research in the field from the last hundred or so years. They also lay out methods for gathering data to use as inputs for analysis. Most notably they show a way to refine human subjective estimates using calibration exercises in order to quantify the amount of uncertainty and the error itself in the estimates, and therefore making them compatible with quantitative analyses. Hubbard and Seiersen (2016) give plenty of examples and step by step guidance on how to use the methods they propose, providing example excel sheets to get started. The book is a cyber security specific edition of an earlier book “How to Measure Anything: Finding the Value of Intangibles in Business” published in 2014 (Hubbard et al, 2016).

4.1.8 Measuring and Managing Information Risk: A Fair Approach

Freund and Jones (2015) discuss the limitations of qualitative methods in analyzing risk and present their quantitative framework and methodology “Factor Analysis for Information Risk” (FAIR) as a solution. They make note that subjective subject matter expert estimates are often necessary to be used as inputs to these analyses and offer Hubbard’s calibration methodology as a solution for increasing the accuracy, claiming that the average accuracy rate after their own training courses can reach 80-90%. They also provide examples of applying the concepts in the FAIR framework in the form of analyses of various types of threats and vulnerabilities like website denial of service attack and unencrypted internal network traffic (Freund et al, 2015).

4.1.9 Quantitative Vulnerability Assessment of Cyber Security for Distribution Automation Systems

Ye, Zhao, Zhang, and Wen (2015) created a game-theory based quantitative vulnerability assessment model for estimating potential losses in megawatt hours in case of cyber-attacks on distribution automation systems (DAS), used in smart electric grids. In their conclusions Ye et al (2015) warn that DAS is more vulnerable to cyber-attacks due to their lesser physical security compared to control systems in power plants or substations. However, the quantity of DAS units makes completely securing each DAS unit uneconomical, and therefore they propose their method to be used to assess the relationships of different vulnerabilities in order to better focus investments. (Ye et al, 2015).

4.1.10 A review of Cyber Security Risk Assessment Methods for SCADA Systems

Cherdantseva et al (2016) conducted a review of 24 risk assessment systems available for Supervisory Control and Data Acquisition (SCADA) systems used in critical infrastructure, such as delivery of electric power, water, and gas, and in telecommunication systems in addition to various other industry systems. They also developed a categorization scheme for the risk assessment methods and outlined the research challenges in the domain with suggestions of approaches to tackle the problems (Cherdantseva et al, 2016).

4.1.11 Quantitative Assessment of Cyber Security Risk using Bayesian Network-based Model

Mo et al (2009) developed a quantitative model using Bayesian networks to model the probabilities of cyber risks. The goal was to present the field an implementable model to use in the future to continuously gather situational data from different companies to produce accurate quantitative risk scores. The risk scores could then be translated into percentage values of probabilities for companies to use in their risk analyses as basis for improving the quality of their cyber security (Mo et al, 2009). Mo et al (2009) suggest further research by testing implementing the model in a real environment or running simulations on it.

4.1.12 Conclusions From the Literature Review

In the recent literature of quantified risk analysis, three recurring themes can be seen.

First, the lack of standardized tools to quantify cyber risk. A great number of frameworks, tools, and models are used, but majority of these are qualitative in nature, and either focus on compliance or qualitative risk scores instead of expressing risk in monetary units (Radanliev, 2019). Lack of standardized methods of reporting and sharing information about risk has also led the Department of Homeland Security in the US into taking steps to find such methods for the financial services sector (Jones, 2018).

Second, the difficulty of gathering empirical data to use in risk assessments often leads to reliance on subjective estimates by subject matter experts. Some reject this method and look for exclusively empirical sources for data, while others attempt to compensate for the problems in subject matter expert estimates and reduce the subjectivity in them. For example, Wilson (2019) looked into using the Verizon Data Breach Investigation Reports and Ponemon Institute Cost of a Data Breach Reports for inputs into his analysis (Wilson, 2019). Cherdantseva et al (2016) found that the available probabilistic methods rely either on historical system data or on subjective data, while the scarce availability of the former limits their usefulness and the subjective nature of the latter limits their applicability (Cherdantseva, 2016). On the other hand, Hubbard (2016) has devised a way to reduce the subjectivity in the subject matter expert's assessments through calibration training. Freund (2015) has reported good results with Hubbard's (2016) method and calls for focusing on the usefulness of risk analysis findings, in regard to investment decisions, over the degree of subjectivity versus objectivity behind the methods. Freund (2015) also reminds that while any decision is always going to entail some amount of subjectivity, any qualitative analysis is always going to be more subjective than a quantitative one where rigor has been expended into reducing the subjectivity contained in the analysis as much as possible (Freund et al, 2015).

Third, bridging the gap in communication between cyber security and IT staff and senior management. Hubbard (2016) and Freund's (2015) methods express risk in monetary terms in order to use the same language as the decision makers in the senior managements of the companies. Freund (2015) calls for the usefulness of the methods and describes the process of risk management as pulling together data of varying quality, applying models to it in order to interpret what the data means, and finally someone making decisions based on the presentation of those interpretations (Freund et al 2015). According to Jones (2018), the current state of affairs in the US financial services sector is that the strategic and board of directors level manages a qualitatively predefined risk appetite, investments are primarily focused on where controls are perceived to be inadequate, and root cause analysis, pen tests, and external audits are performed to assess the adequacy of the overall risk management process. However, there are no traditional return on investment or financial calculations and the budgets are generally built based on the CISO’s recommendations (Jones, 2018). According to Freund (2015), a methodology must be useful, practical, and defensible. Decision makers are usually under pressure to decrease costs, while investments in cyber security controls usually cost money without producing clear returns on investments. Risks expressed in monetary units are more useful in communicating required investments to decision makers than classifying something as “low”, “medium”, or “high” risk (Jones, 2015).

The current research aims to improve the situation by providing experience in setting up and using a simple set of quantitative risk management tools that can be used to provide decision makers actionable information on the IT and cyber risks in their environments while addressing all of these themes. The first theme is addressed by using a completely quantitative approach. The second theme is addressed by suggesting calibration method for improving the subjective evaluations of experts and by using Beta Distribution to mathematically update risk probabilities based on observed data. The third theme is addressed by expressing risk in monetary terms and by the use of various illustration techniques in the form of graphs showing the development of risk.

4.2 Risk

According to Krausse (2006), risk is considered when there are decisions to be made involving uncertain future consequences. Holton (2004) on the other hand defines risk as exposure to a proposition of which one is uncertain. A proposition is either true or false and the uncertainty of the future reflects not knowing which state will come to pass.

Holton (2004) gives an example. A person is playing a game of rolling a die. If the result is three, the person will lose 100 €. What is the probability of the person losing the money? When the roll comes, a three-sided die is revealed. This example illustrates that one can be uncertain without being aware of the uncertainty (Holton 2004).

If the state of a proposition ends up being true, there will be exposure to consequences. For example, a proposition may be that it is raining, and a person is outside without an umbrella. The person experiences exposure due to having a preference to this proposition being false. The degree of uncertainty of the proposition does not affect the degree of exposure to it. Therefore, even being completely unaware of the proposition the person would still be exposed to it as long as he or she cared about the consequences. Finally, if a person jumped out of an airplane without a parachute, facing certain death, he or she faces no risk as risk requires both exposure and uncertainty (Holton 2004).

According to Holton (2004), risk is also a condition of individuals that are self-aware. Organizations are not self-aware and therefore incapable of being at risk, but instead work as conduits through which individuals bear risk (Holton 2004).

According to Krause (2006), risk is mathematically defined as the product of the probability of an event occurring and its impact after the fact. However, historical data is often not available, and risk considerations must often be based on expert judgement subject to over- or underestimation depending on matters like individual personality, background, experience, or the formulation of the question. Depending on the various consequences, risks can be divided into different categories, for example environmental, economic, political, social and technical. (Krause, 2006).

4.3 Risk Tolerance

Knowing the risk level and in which areas the risks lie in the organization is important, but this information alone does not reveal if the level of risk is appropriate or acceptable for the organization. Therefore, an organization must define the acceptable level of risk, also called “risk tolerance” or “risk appetite”. In practice the level of risk tolerance must be defined by the senior leadership. According to (Timothy Virtue, Justin Rainey, in HCISPP Study Guide, 2015) risk tolerance is determined as a part of the organizational risk management strategy to ensure consistency across the organization, and also used in making investment and operational decisions. (Virtue & Rainey 2015). According to Evan Wheeler (2011) the security leaders need to gather and document the risk tolerance posture and priorities from the C-level executives such as Chief Executive Officer (CEO), Chief Information Officer (CIO), and Chief Financial Officer (CFO) of the organization. This is the only way to make sure the decisions made based on the risk analyses are in line with the management goals of the organization (Wheeler 2011).

Risk tolerance is a continuum between risk averse and risk seeking postures, with the middle of the both extremes being risk neutral. Risk tolerance is a highly personal trait in people, but according to Brownley (2013) people tend to be risk neutral when the risks involved in a decision are small relative to their total assets. A leader would be considered risk neutral towards a set of alternatives when he would sell the set for its expected value. Brownley et al. gives the following example. Suppose an executive believes there is a 40 percent chance the new product is worth $40 million in profit, a 40 percent chance it will cause $10 million in losses, and a 20 percent chance it will cause $30 million in. The expected value of the new product is therefore 0,4*60 + 0,4*-10 + 0,2*-30 = $14 million. A risk neutral executive would sell the product for the expected value of $14 million, while a risk averse executive would sell it for less to minimize a chance of losing $30 million, and a risk seeking executive would require more money for it to maximize the profits. According to Brownley et al, risk averse posture is much more common among businesspeople. (Brownley et al 2013)

4.4 Risk Management

Risk management means structuring a firm's portfolio of activities so that the composition of the returns and risks taken to produce them are optimal. Therefore, risk management is about maximizing value (Krause, 2006).

SFS-ISO 27005 (2011) defines a six-stage iterative process for managing information security risk. The stages are "context establishment", "risk assessment", "risk treatment", "risk acceptance", "risk communication and consultation", and "risk monitoring and review". The six stages are aligned to Information Security Management System (ISMS) following the classical Plan-Do-Check-Act (PDCA) cycle (SFS-ISO, 2011) developed by the Japanese manufacturers as a part of the Kaizen continuous improvement model (creative safety supply). The process of SFS-ISO 27005 (2011) is illustrated in Figure 2 (SFS-ISO, 2011).

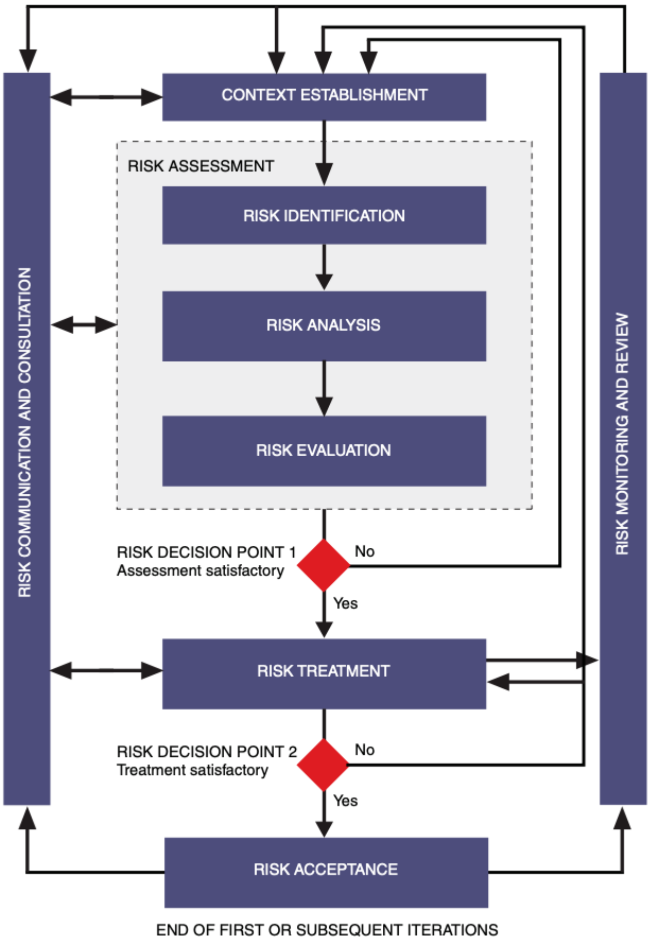
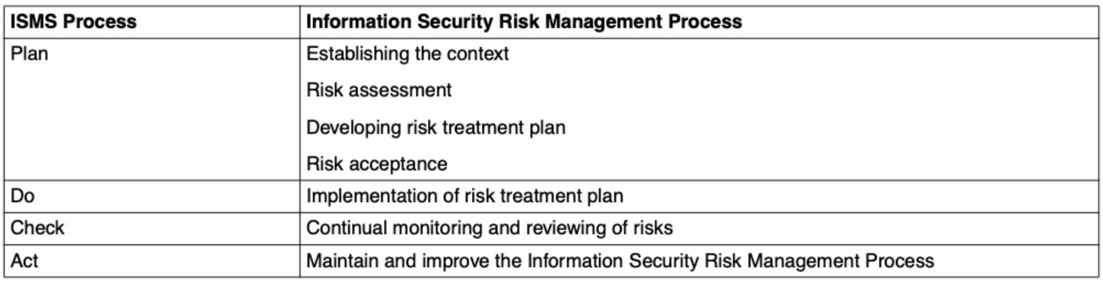


Figure 2: The information risk management process (SFS-ISO, 2011)

The risk management process begins with establishing the context for risk management, after which a risk assessment is performed, including the identification of risks, analyzing their impacts and likelihoods and evaluating their relative magnitudes. The risk assessment phase can also be iterative to increase the depth and detail of the assessment with each iteration. The assessment is used to plan the actions required to modify the risks to an acceptable level in relation to the Risk Tolerance of the management, and the management accepts the plan at decision point 1. Risk treatment follows after decision point 1, where the plan is put into action. Then the results of the treatment and the remaining residual risk to be retained is presented to the management in decision point 2 for explicit acceptance. If the management accepts the new level of retained residual risk, the next step is to continuously monitor and review the risk portfolio and revisit periodically revisit the risks. At each state communication of risk is performed to different stakeholders and stakeholders are consulted as required. (SFS-ISO, 2011). Table 2 shows the different tasks mapped to the PDCA cycle as described by SFS-ISO (2011).

Table 2: Information security risk management tasks mapped to the PDCA cycle (SFS-ISO 27005, 2011)

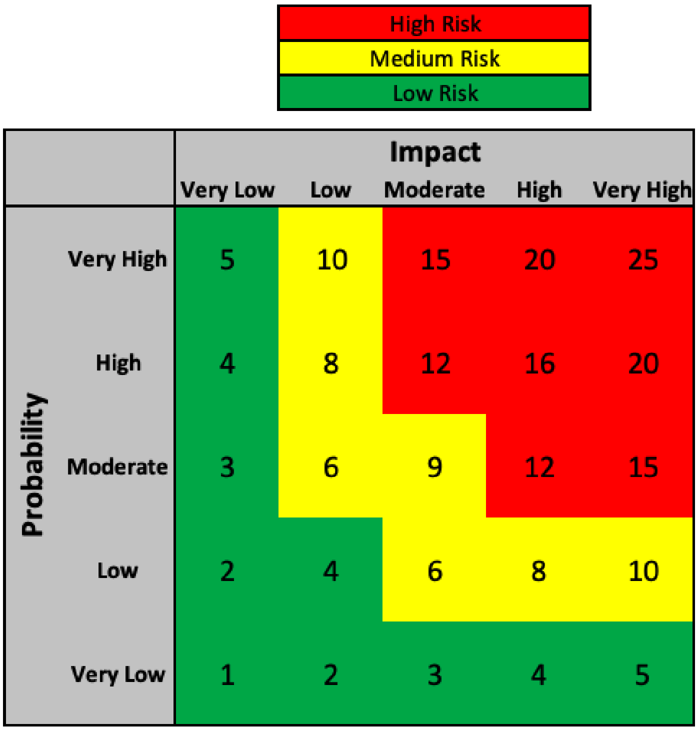


ISMS Process	Information Security Risk Management Process
Plan	Establishing the context Risk assessment Developing risk treatment plan Risk acceptance
Do	Implementation of risk treatment plan
Check	Continual monitoring and reviewing of risks
Act	Maintain and improve the Information Security Risk Management Process

4.5 Qualitative Risk Analysis with Risk Matrices

In the qualitative risk matrix methods, the probability and impact are defined with verbal descriptions, such as low, medium or high, or with numerical representations corresponding to these verbal descriptions. These numerical representations are then multiplied together in order to come up with individual risk scores for various identified risks (Hubbard, 2016). The advantage of this approach is that these risk scores can then be compared and ranked in order of priority (SFS-ISO, 2011).

Figure 3 shows an example of a qualitative risk matrix, in which the probability and impact are respectively described in five steps: “Very Low (1)”, “Low (2)”, “Moderate (3)”, “High (4)”, and ”Very High (5)”. Each cell in the model represents a risk score, and individual risks are mapped in the matrix so that if, for example, “Risk A” has a “High” probability and “Moderate” impact, the risk score for “Risk A” would therefore be 12 which translates to “High Risk”. This mapping is performed for each risk in a risk portfolio and the risks can then be ranked in order of priority based on their risk scores.



🟥 High Risk
🟨 Medium Risk
🟩 Low Risk

		Impact				
		Very Low	Low	Moderate	High	Very High
Probability	Very High	🟩 5	🟨 10	🟥 15	🟥 20	🟥 25
	High	🟩 4	🟨 8	🟥 12	🟥 16	🟥 20
	Moderate	🟩 3	🟨 6	🟨 9	🟥 12	🟥 15
	Low	🟩 2	🟩 4	🟨 6	🟨 8	🟨 10
	Very Low	🟩 1	🟩 2	🟩 3	🟩 4	🟩 5

Figure 3: An example of a qualitative risk matrix

4.6 Accuracy of Measurement Using Qualitative and Quantitative Methods

Most notable publications about whether or not quantitative methods are more accurate than qualitative ones have been written by Paul E Meehl (1954) in his book “Clinical versus Statistical Prediction; A Theoretical Analysis and a Review of the Evidence” and Philip E. Tetlock (2005) in “Expert Political Judgment: How Good Is It? How Can We Know?”. Meehl’s collection of 136 studies show that mechanical methods are “almost invariably” better at predicting things than expert intuition (Meehl 1954). The main finding was that human judgement is prone to various types of biases and errors in thinking, and is, in general, inconsistent to the extent that the mere consistency of “mechanical methods” is often enough to outperform human judgement (Meehl, 1954). This is elaborated by Tetlock who commented based on his 20-year experiment, that “it is impossible to find any domain in which humans clearly outperformed crude extrapolation algorithms, less still sophisticated statistical ones” (Tetlock, 2005).

4.7 Human Factor in Risk Management Assessments

The literature review revealed that human assessments are often used in the risk analysis process. However, in addition to the qualitative versus quantitative question delved into by Meehl (1954) and Tetlock (2005), research into the assessment capabilities of humans have been carried out by Stuart Oskamp et al (1965), William M. Grove et al (1996), and C. Tsai et al (2008). In their research, Oskamp et al (1965), and C. Tsai et al (2008) demonstrate how increasing information on a matter leads to increasing confidence in estimates without simultaneous increase in accuracy of estimates, leading to overconfidence among the individuals making the estimations. People make decisions based on very little fragmentary information, and as they receive more information, they tend to become more and more convinced of their understanding of the subject. Based on this increasing confidence, these individuals also become increasingly unwilling to change their original decisions. (Oskamp et al, 1965) This finding is mirrored by C. Tsai et al (2008); people who received less information in the beginning of their experiments ended up more overconfident than those who received more information from early on, corroborating Oskamp et al’s (1965) findings. C. Tsai et al (2008) interpret that the test subjects did recognize the value of the additional information that they received but were poor at recognizing their own limitations incorporating that information to their decisions (C. Tsai et al, 2008). Grove et al (1996) explain that when making estimates, a person’s brain is working as a substitute for an “explicit regression equation or actuarial table”, but unlike a computer, human brain cannot process all the information and cannot stay consistent in the calculations (Grove et al, 1996).

When using qualitative expressions of ranges of values compressed into ordinal point values, accounting for inherently inaccurate nature of human judgment is impossible. Research into the calibration of experts providing subjective quantitative estimates provides tools to improve and account for this error. Research has been performed, for example, by Walker et al (2003) in their two-part study of the calibration of expert judgements about personal exposures to benzene. The study gathered a group of experts to discuss research results on the subject, after which they were interviewed individually and asked to create subjective estimates of probability distributions about personal exposure levels in various situations. The results were later compared to measured values to find out the level of calibration of the experts. Walker et al (2003) cite methods of quantifying the calibration of the experts by comparing their provided answers to the measured values and calculating calibration factors based on the difference (Walker et al, 2003). Koehler et al (2002) found out different varying degrees of miscalibration among collection of similar studies from different domains, including calibration in medical settings, weather forecasting, legal judgements, business settings and sports.

Hubbard, (2016), proposes a method for increasing the calibration of experts before using them as inputs in statistical models. In the literature review, Hubbard (2016) was the only one to propose such methods in the context of risk assessments, and his methods have also been incorporated into the FAIR model presented by Freund (2015). The calibration exercises can be used to teach a person to express estimates with a consistent amount of uncertainty (Hubbard, 2016).

The method described by Hubbard (2016) is as follows. The expert is asked to express estimates in 90 % confidence intervals to a large number of prepared questions, for which the correct answers are known, such as "I am 90% confident that Isaac Newton was born between years 1600 and 1700". After this the expert is given a choice to either spin a wheel of fortune, which gives him a 90 percent chance to win an amount of money, or to check if the year of Newtons birth is inside the given range. In either case the expert will win a prize money, but if the expert truly is confident that the answer has a 90% chance of being inside the range, the expert should be indifferent to the options. If the wheel of fortune is a preferred option, then the range needs to be adjusted until the expert is indifferent to the two options. After ten or more questions, it becomes possible to count the correct answers and the questions are continued until the ratio of 9/10 correct answers is reached (Hubbard, 2016).

According to Hubbard (2016), this ability to express the amount of uncertainty in the estimate consistently is an acquired skill and therefore transferable to estimation of anything, as the only thing the person is learning is to assess and express his or her personal uncertainty of the matter at hand. The method was demonstrated in 1997 when 16 Giga Information Group analysts underwent this calibration training and were asked to make true/false predictions stated as events occurring or not occurring by June 1, 1997, for example, “True or False: Intel will release its 300 MHz Pentium by June 1”. The same questions were presented to 20 uncalibrated Giga Information Group client CIOs, and the results of the two groups predictions were presented in Giga Information Group symposium 1997, after the true results had become known. As a result, the calibrated analysts had been able to estimate their chances of being correct very close to the “ideal confidence”, meaning that for questions for which they had claimed their chances of being correct to be for example 90%, about 90% of the answers actually were correct. Figure 4 shows the results presented in the 1997 symposium, with the dotted line representing the ideal confidence. The un-calibrated client CIOs fared significantly worse than the calibrated experts. The results were very similar to the findings in Oskamp et al (1965) and C. Tsai et al’s (2008) experiments (Hubbard, 2016).

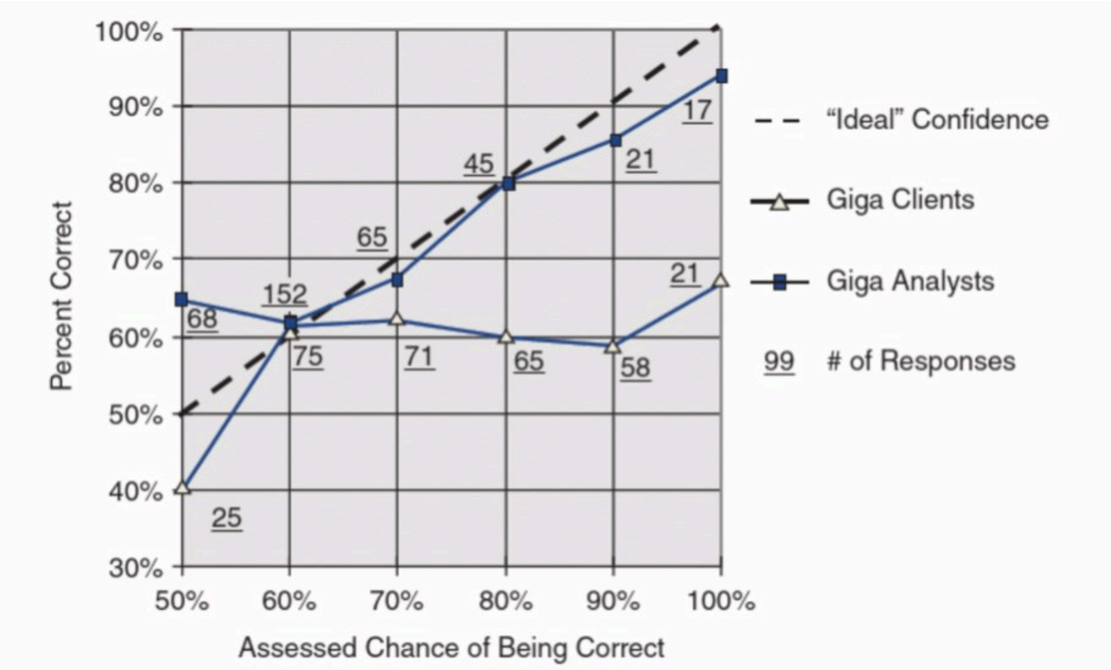


Figure 4: Calibrated vs. uncalibrated estimations for results presented in Giga World 1997 (Hubbard 2016).

Expert estimate data gathered in this manner can further be refined by presenting the same questions to a group of calibrated experts, and the averages used to reduce outlier errors. This data can then be used as input data in, for example, stochastic simulations such as Monte Carlo simulations, and refined with further by using the

Beta Distribution or other statistical methods as more data becomes available (Hubbard, 2016).

In addition to Hubbard (2016), Clemen et al (2002) have also suggested debiasing techniques based on Bayesian statistics that could be used to further refine the expert estimates.

4.8 Problems with Risk Matrices

4.8.1 Ordinal Scales

The problems with the risk matrix approach stem from the use of ordinal scales, which only define the order of the objects within a list but not the type or distance between them (Hubbard, 2016). Qualitative risk matrices may imply a fixed distance between the values of probability and impact, however, ordinal scales are clearly used, such as from 1 to 5, or "low", "medium", and "high". Ordinal scales are described by Stanley Smith Stevens in his paper On the Theory of Scales of Measurement (1946), where he describes the different types of scales used in mathematics: nominal, ordinal, interval, and ratio scales. As shown by Stevens (1946), ordinal scales do not adequately support the multiplication operations that are commonly performed within risk matrices, namely when using numerals to describe the probability and impact, the risk score is calculated as the product of the two. Stevens (1946) agrees that "for this 'illegal' statisticing there can be invoked a kind of pragmatic sanetion: In numerous instances it leads to fruitful results" (Stevens 1946) but right after he warns that "when only the rank-order of data is known, we should proceed cautiously with our statistics, and especially with the conclusions we draw from them" (Stevens 1946). Stevens (1946) also points out that even for the statistics that are normally appropriate for ordinal scales, "rigor is sometimes found compromised" (Stevens 1946).

The use of ordinal scales is a fundamental problem in qualitative risk analysis methods, and after a review of the models in use in American financial service sector, Jones et al (2018) state that due to lack of meaningful quantitative data, the models are "based on subjective probability assessments along with an acknowledged decision to violate basic math by aggregating and trending ordinal scales" (Jones et al, 2018).

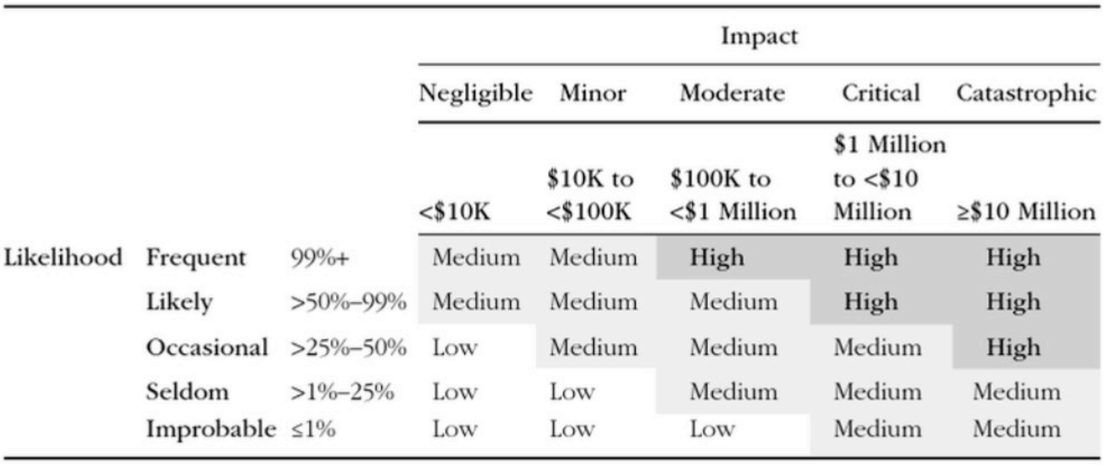
Risk data often needs to be aggregated and analysed using higher math functions than those which are supported by ordinal scales, which is evident for example based on Jones et al (2018) observation of aggregation routinely being done in the American Financial Service Sector. Stevens (1946) states that conclusions based on calculations on ordinal scales are without solid basis and therefore, methods based on interval or ratio scales ought to be used instead. The following chapters describe causes and effects of compromised rigor in qualitative risk matrices.

4.8.2 Range Compression

The qualitative risk matrix approach compresses ranges of values into single points, resulting in a kind of extreme rounding error called by Hubbard (2016) "range compression". The problem is further exacerbated by multiplying two range compressed values together, resulting in situations where a decision maker would be required to treat greatly varying risks equally (Hubbard, 2016).

Sometimes attempts are made to more strictly define what the different values in a risk matrix mean, while the underlying scales stay ordinal, easily leading to range compression which can yield unexpected results. In Hubbard's (2016) example the risks are defined in such a manner, that likelihoods are given in percentile ranges and the impacts are given in monetary ranges, shown in Table 3 (Hubbard et al, 2016).

Table 3: Example of a risk matrix with defined probability and impact ranges (Hubbard 2016)



			Impact				
			Negligible	Minor	Moderate	Critical	Catastrophic
			<$10K	$10K to <$100K	$100K to <$1 Million	$1 Million to <$10 Million	≥$10 Million
Likelihood	Frequent	99%+	Medium	Medium	High	High	High
	Likely	>50%–99%	Medium	Medium	Medium	High	High
	Occasional	>25%–50%	Low	Medium	Medium	Medium	High
	Seldom	>1%–25%	Low	Low	Medium	Medium	Medium
	Improbable	≤1%	Low	Low	Low	Medium	Medium

For risk A with likelihood of 2% and an impact of $10 million, the resulting risk is defined as Medium. For risk B with likelihood of 20% and an impact of $100 million, the resulting risk is also defined as Medium. The decision maker would therefore be required to treat these two risks equally, while risk B is in reality a hundred times greater than risk A. Risk C with likelihood of 26% and impact of $10.000.001, on the other hand, would be considered High, while the monetary impact is only about one tenth of the risk B (Hubbard, 2016).

The problem of range compression can be avoided completely by using the ranges directly, by expressing them as confidence intervals, for which an approach is described in detail later.

4.8.3 Interpretation Problems

Budescu et al (2007) studied how people interpret verbal descriptions of likelihoods and found out that people interpret words like “unlikely”, “probable”, “likely” or “highly likely” differently, even when told exactly how to interpret the words. One instance of such research was conducted by Budescu et al (2007) on how well people followed guidelines when assigning numerical values to probability terms in the context of Intergovernmental Panel on Climate Change reports. One of the findings was that between 43 and 67 percent of the responses violated the guidelines to the extent that for “Very Unlikely”, which was defined as “less than 10%”, the maximum of all responses was 75%. Only 7.5% of the respondents followed the guidelines completely. Figure 5 shows the distribution of the estimates of the subjects of four probability terms used in the study. Translation group was asked to give their answers adhering to a translation table, which defined the terms as “Very likely (>90%)”, “Likely (>66%)”, “More likely than not (>50%)”, “Unlikely (<33%)” and “Very unlikely (<10%)”, while the control group did not have a translation table. Each box on the figure represents the central 50% of the answers, with the solid line inside the box representing the median. The dotted horizontal lines represent the category boundaries in the boxes. The numbers in the boxes show the percentage of responses inconsistent with the guidelines with the placement below or over the median line, indicating the direction of the misinterpretation (Budescu et al, 2007).

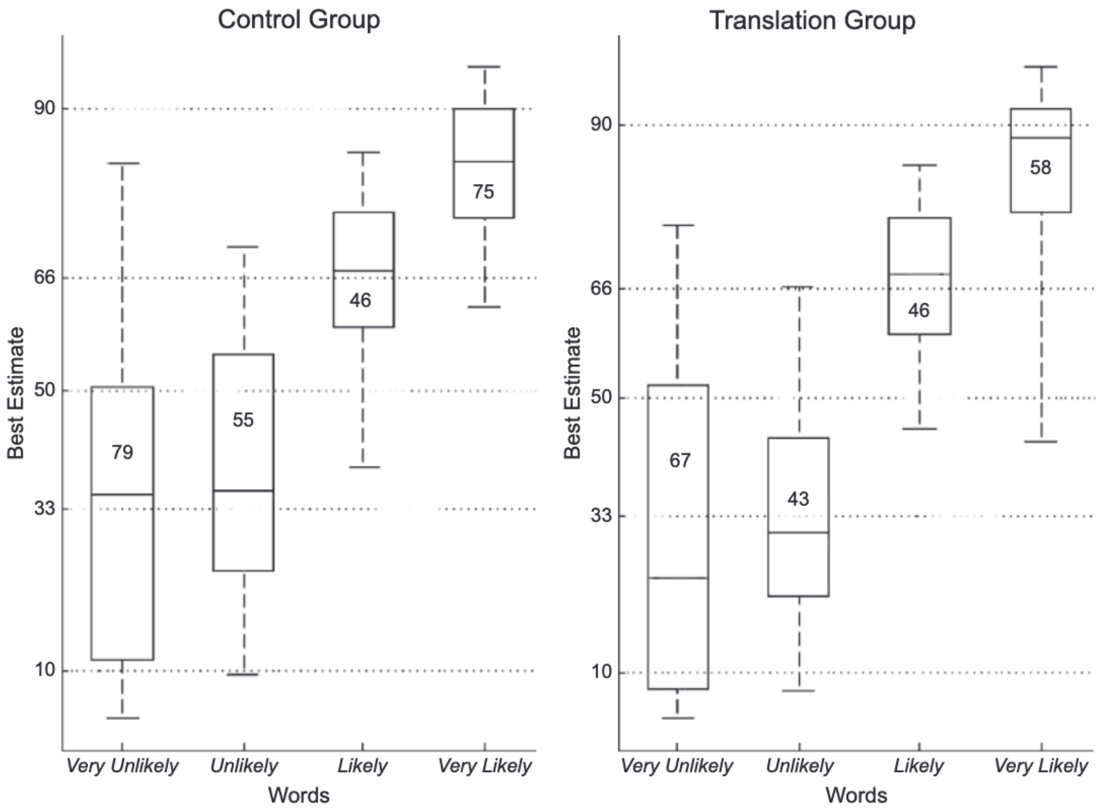


Figure 5: Distribution of estimates of the meanings of four probability terms Budescu et al (2007)

Based on Budescu et al’s (2007) findings, people’s interpretation of the meaning of the same descriptive words can vary to the extent that one person’s “very likely” can easily be the same as another person’s “very unlikely”. Using these descriptive words is convenient and makes people think they understand what is being said, but the interpretation of them is shown to be highly variable (Politi et al, 2007). Therefore, using words such as “likely” or “unlikely” when collecting expert estimates for analyses or communicating risk comes with inherent error based on the precariousness of verbal expressions of uncertainty and the differing ideas people have of them. The solution would then be to always use numeric expressions and avoid verbal ordinal scales altogether (Hubbard et al, 2016).

4.8.4 Problems Aggregating Portfolios

In addition to the variation in interpretations, creating and aggregating portfolios of risks to show the combined risk rating of combined risk portfolios is difficult due to the inability to add risk portfolios and risk scores up when using ordinal scales (Stevens, 1946). Jones (2018) notes that while aggregating and trending ordinal scales violates basic mathematics, this seems to be routinely done anyway (Jones et al, 2018).

4.8.5 Problems Comparing Risks and Calculating Return on Investment

Comparison of different risks and measuring the effects of controls afterwards is difficult due to the tendency of ordinal scales to compress large ranges of values into ordinal point values. Questions like how many “low” risks equal one “medium” or “high” risk, and how can it be known that a “high” risk actually dropped to “medium” are difficult to answer. Also, different people can interpret the words used completely differently. The inherent inaccuracy of ordinal scales and the precarious nature of human perception of the scales renders ordinal qualitative risk analyses arbitrary. (Hubbard 2016; Budescu 2007).

4.9 Advantages of a Quantitative Approach

Numerous studies have shown that even crude algorithms consistently match or out-perform human judgement. Meehl (1954) gathered a significant body of studies in different fields, ranging from football games, to business failures and success in military training. After reviewing 136 studies on comparing the accuracy of human predictions to algorithmically created ones, Grove and Meehl (1996) state that algorithms are almost invariably equal or superior to the human predictions. Based on Grove and Meehl’s (1996) findings, quantitative risk analysis methods can be expected to outperform qualitative methods in various ways and as such provide better value in decision making, rendering the development and implementation of quantitative methods over qualitative ones worthwhile.

4.10 Quantitative Risk Management Model for The Cyber Exercise

4.10.1 Available Methods

Jones et al (2018) found an almost uniform agreement that no generally accepted model exists for organizations to guide cyber security investments, and the lack of credible data on the number and impact of cyber-attacks increases the difficulty of quantifying risks for organizations, which drives organizations to fall back to their own metrics, used as the basis for investment decisions (Jones et al, 2018). Freund et al (2015) call for risk analysis methods which yield results that are “useful”, “practical”, and “defensible” when making investment decisions. Instead of stating that a risk is “high” the results need to be in units that are useful when comparing investment options. For example, “The annualized exposure is between 60 000 € and 330 000 €”. The method needs to be practical, simple to use and understand, and the analysis needs to be transparent in order to be defensible when challenged by an executive making investment decisions (Jones et al, 2018).

In their empirical analysis of the ten currently available cyber risk frameworks, methodologies, systems, and models, Radanliev et al (2019) found only three quantitative risk assessment options: FAIR, RiskLens and Cyber VaR (Radanliev et al, 2019).

Value-at-Risk (VaR) is a category of probabilistic risk measures used in the financial sector. A financial portfolio consists of several assets, each of which has a current market value, the market value of the portfolio is the sum of the respective market values of the assets in the portfolio. The future market value of the assets and the portfolio, on the other hand, is uncertain. A VaR metric is a function of the current value and a probability distribution describing the possible future values of an asset. In recent years efforts have been made to modify VaR into a “Cyber VaR” version to specifically quantify cybersecurity risks, with the goal of helping risk and information security professionals to articulate cyber risk in financial terms (Jones et al, 2018).

Factor Analysis of Information Risk (FAIR) is a Cyber VaR method released originally in 2005 under the Creative Commons Attribution-Noncommercial-Share Alike 2.5 license. The private company RiskLens has produced a number of software applications to quantify cybersecurity risk in a manner consistent with FAIR (Jones et al, 2018).

Jones et al (2018) also note the “Hubbard and Seiersen Approach” described by Hubbard (2016) and state that currently two methods are available which can potentially fulfil the needs of the NGCI program and quantify cybersecurity risks for the US Financial Services Sector: FAIR and Hubbard and Seiersen Approach (Jones et al, 2018).

Of the two available methods for quantifying cybersecurity risks, Hubbard and Seiersen Approach was selected for implementation in the cyber exercise. Hubbard and Seiersen (2016) describe a set of methods and tools to perform quantitative risk analysis using calibrated human estimates as inputs to a risk portfolio. The risks in the portfolio are analysed using Monte Carlo simulations to calculate the values of the respective risks in it, as well as loss exceedance curves to visualize the probabilities of different levels of losses in the portfolio. The portfolio can be divided into sub-portfolios based on, for example, different types of risks, and the different sub-portfolios can be aggregated together to create summarized views to all retained risk (Hubbard, 2016).

4.10.2 Monte Carlo Simulation

Monte Carlo was originally a term used by Von Neumann and Ulam during World War II as a code word for secret work at Los Alamos on the first atomic bomb, involving simulation of random neutron diffusion in nuclear materials. The term has later been used to describe stochastic computer simulations which make use of randomness in the underlying model. (Rubinstein, Reuven Y., and Dirk P. Kroese. Simulation and the Monte Carlo Method, John Wiley & Sons, Incorporated, 2016.)

According to Rubenstein et al, simulations are useful when the system to be analyzed is so complex that formulating simple mathematical equations to describe it is impossible. On the other hand, the results of a simulation are relevant and accurate to the real-world problems only to the extent that the model used in the simulation is a valid representation of the system under study (Rubinstein et al 2016).

The type of Monte Carlo Analysis described by Hubbard et al (2016) uses three inputs to create the loss exceedance curves used in the risk analysis: a range of expected impact in case a risk realizes given as a 90% confidence interval, meaning that the impact is expected to be within the interval 90% of the time, less than the lower bound 5% of the time, and greater than the upper bound 5% of the time. The probability of the events is given as percentage values and a probability distribution is used to approximate the distribution of the events. Log normal distribution is selected due to it emphasis on the less extreme impacts more than the extreme impacts (Hubbard, 2016).

Modern computers allow for creation of Monte Carlo simulations with hundreds of thousands of samples, and Hubbard (2016) offers excel examples on a website supplementary to his book. The following section provides an illustration of how the calculation works using an analysis from the cyber exercise, with the actual exercise and the results described later in more detail.

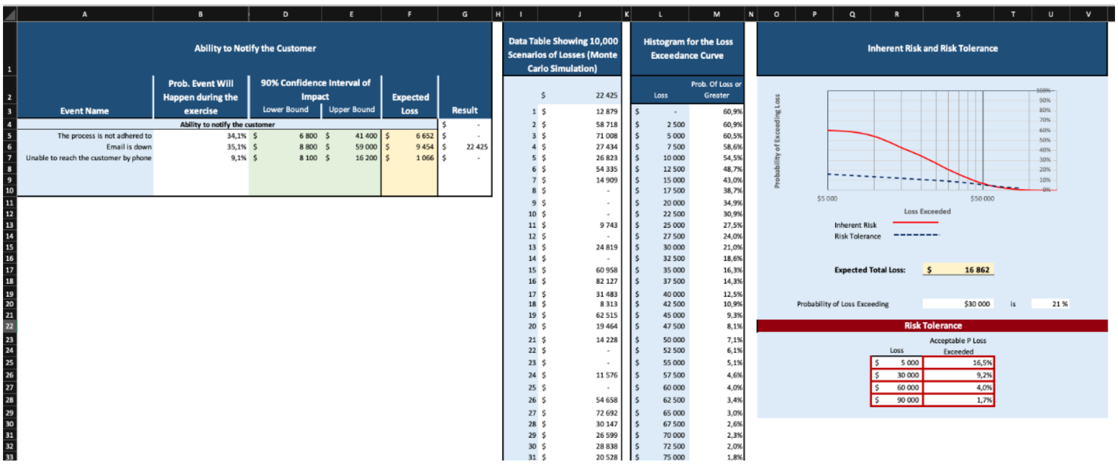
In the initial analysis a risk described as “Email is down”, referred to as “PROB” in subsequent excel formulas, was given a 35,1 % probability of realizing during the exercise. The Lower Bound (LB) of the impact was set to $8,800 and the Upper Bound (UB) was set to $59,000. The excel formula for calculating an instance of losses with these inputs, as provided by Hubbard (2016), is:

```
=IF(RAND()<PROB;LOGNORM.INV(RAND();(LN(UB)+LN(LB))/2; (LN(UB)-
LN(LB))/3,29);0)
```

The excel formula first generates a random number between 0 and 1 and checks if the number is smaller than the required probability percentage. If the number is larger, which will happen 64,9 % of the time, the result will be 0. If the random number is smaller, which will happen 35,1 % of the time, the formula will proceed to create a log normal distribution with the given Upper Bound and Lower Bound values and pick a random point on that distribution. The chosen random value will then be the result of the calculation.

A portfolio is created by adding the risks in scope to the excel sheet, performing the same calculation for each risk and summing up the results.

In order to create a loss exceedance curve for the portfolio, the calculation is repeated 10,000 times in an Excel data table to create 10,000 simulated exercise runs. 10,000 is a round number to use in the simulation to provide enough resolution but if needed, the number of instances can be increased. A histogram for the loss exceedance curve is created by counting the instances within the 10,000 samples where the loss was greater than a given number. Figure 6 shows the simulation being performed in excel moving from left to right. The same illustration can be found in the appendix 1 in greater size.



Ability to Notify the Customer

Event Name	Prob. Event Will Happen during the exercise	90% Confidence Interval of Impact		Expected Loss	Result
		Lower Bound	Upper Bound		
Ability to notify the customer					$ -
The process is not adhered to	34,1%	$ 6 800	$ 41 400	$ 6 652	$ -
Email is down	35,1%	$ 8 800	$ 59 000	$ 9 454	$ 22 425
Unable to reach the customer by phone	9,1%	$ 8 100	$ 16 200	$ 1 066	$ -

Data Table Showing 10,000 Scenarios of Losses (Monte Carlo Simulation)		Histogram for the Loss Exceedance Curve	
	$ 22 425	Loss	Prob. Of Loss or Greater
1	$ 12 879	$ -	60,9%
2	$ 58 718	$ 2 500	60,9%
3	$ 71 008	$ 5 000	60,5%
4	$ 27 434	$ 7 500	58,6%
5	$ 26 823	$ 10 000	54,5%
6	$ 54 335	$ 12 500	48,7%
7	$ 14 909	$ 15 000	43,0%
8	$ -	$ 17 500	38,7%
9	$ -	$ 20 000	34,9%
10	$ -	$ 22 500	30,9%
11	$ 9 743	$ 25 000	27,5%
12	$ -	$ 27 500	24,0%
13	$ 24 819	$ 30 000	21,0%
14	$ -	$ 32 500	18,6%
15	$ 60 958	$ 35 000	16,3%
16	$ 82 127	$ 37 500	14,3%
17	$ 31 483	$ 40 000	12,5%
18	$ 8 313	$ 42 500	10,9%
19	$ 62 515	$ 45 000	9,3%
20	$ 19 464	$ 47 500	8,1%
21	$ 14 228	$ 50 000	7,1%
22	$ -	$ 52 500	6,1%
23	$ -	$ 55 000	5,1%
24	$ 11 576	$ 57 500	4,6%
25	$ -	$ 60 000	4,0%
26	$ 54 658	$ 62 500	3,4%
27	$ 72 692	$ 65 000	3,0%
28	$ 30 147	$ 67 500	2,6%
29	$ 26 599	$ 70 000	2,3%
30	$ 28 838	$ 72 500	2,0%
31	$ 20 528	$ 75 000	1,8%

Inherent Risk and Risk Tolerance

Expected Total Loss: $ 16 862

Probability of Loss Exceeding $30 000 is 21 %

Risk Tolerance

Loss	Acceptable P Loss Exceeded
$ 5 000	16,5%
$ 30 000	9,2%
$ 60 000	4,0%
$ 90 000	1,7%

Figure 6: Example of a sub portfolio

Figure 6 shows the risks listed in column A, the probability of the events in column B and 90% confidence interval in columns D and E. Column F shows the expected loss for each individual risk as the probability weighted average using log normal distribution, while column G shows the results of a single random pass of the simulation. The data table in columns I and J continues down for 10,000 rows, in which each row contains the previously described calculations for each risk, representing the 10,000 individually simulated exercise runs. The sum of the results is displayed in column G. Column M shows a percentage of how many values in the data table are over the given loss on the corresponding line in the column L. The histogram is used to create the red line in the graph to the right. The dotted Risk Tolerance line is created by using the Risk Tolerance table below the graph.

4.10.3 Log Normal Distribution

Probability distributions are used to approximate natural phenomena and need to be selected based on the phenomenon under study in analyzes. Log normal distribution is well known in many fields, including ecology, economics, and risk analysis (Cokhale & Mullen, 2008) and is a variation of the bell-shaped normal distribution. Figure 7 illustrates the difference between log normal and normal distributions. Where normal distribution is a symmetric curve centered on the median value of the distribution, a log normal curve peaks at much lower values, but also contains a wider distribution of larger values compared to normal distribution. According to Cockhale and Mullen (2008) the theoretical foundations underlying the lognormal distribution allow for use in software reliability engineering; a field close to analyzing IT risks, as faults in software systems can be seen as amounting to realizing IT risks. According to Cockhale and Mullen (2008) events and failure rates in IT systems follow log normal distribution. Faults are subsets of events, and therefore faults have failure rates that are a sample from the rates of all events. Therefore, if event rates are log normally distributed, then failure rates of faults are also log normally distributed. (Cokhale & Mullen, 2009). Hubbard and Seiersen (2016) suggest using the lognormal distribution in Monte Carlo simulations when modelling positive values that are primarily moderate in scope, but have potential for rare extreme events (Hubbard, 2016). Hubbard and Seiersen also give losses incurred by a cyberattack or costs of a project as examples. Therefore, all risks analysis in the scope of the cyber exercise were assumed to follow the lognormal distribution.

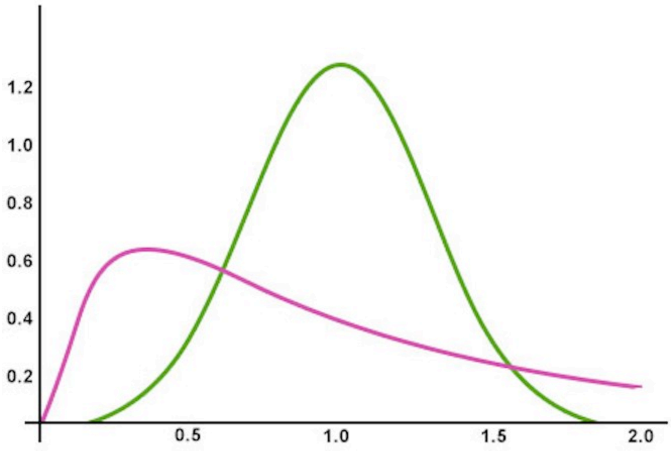


Figure 7: log normal distribution (pink) compared to a normal distribution (green)

4.10.4 Beta Distribution

According to Chattamvelli et al, beta distribution is widely used in statistics, Bayesian models with unknown probabilities, and in order statistics and reliability analysis. (Chattamvelli et al, 2015). In essence, beta distribution is a probability distribution of probability distributions. For a simple illustration, consider a marksman whose proficiency with his gun can be expressed as a probability of him hitting his mark. If nothing was known of his proficiency, an analysis could begin with expecting him to miss 100% of the time. Therefore, his probability of hitting the mark would be 0%. Next the marksman is given ten tries to hit the mark. He hits eight times and misses twice. As a result, his probability of hitting the mark based on these ten tries would be 80%. Next, the marksman is given ten more tries, and this time he hits six times while missing four times. Therefore, his probability of hitting the mark based on these twenty tries would be 70%. Setting the initial expectation to a perceivable value in the beginning of an analysis is however possible, instead of beginning with either 0 % or 100 % and Beta Distribution can be used to update the expected probability using a cumulative probability, as information becomes available. According to Hubbard (2016) Beta Distribution is useful in estimating population proportions with very few data points and they illustrate the procedure with an example of an urn with red and green marbles. The proportion of red marbles could be anywhere between 0% and 100%, and to estimate the population proportion of the red marbles, samples are retrieved from the urn. Each retrieved marble is either a hit (red) or a miss (green), and each hit and miss will update the range in which the population proportion is going to be. (Hubbard and Seiersen, 2016). The more hits and misses are recorded, the smaller the uncertainty of the actual population is going to be making this approach a practical way to describe and continuously upgrade distributions of uncertain outcomes such as the probabilities of risks realizing during the cyber exercise. Figure 8 shows an example of a beta distribution from the cyber exercise risk portfolio, signifying the probability distribution of probabilities of the risks in the portfolio realizing (hits) during the exercise after three exercise runs. With all updated information, the distribution got thinner and taller as the population proportion of realized risks as opposed to not realized risks became clearer.

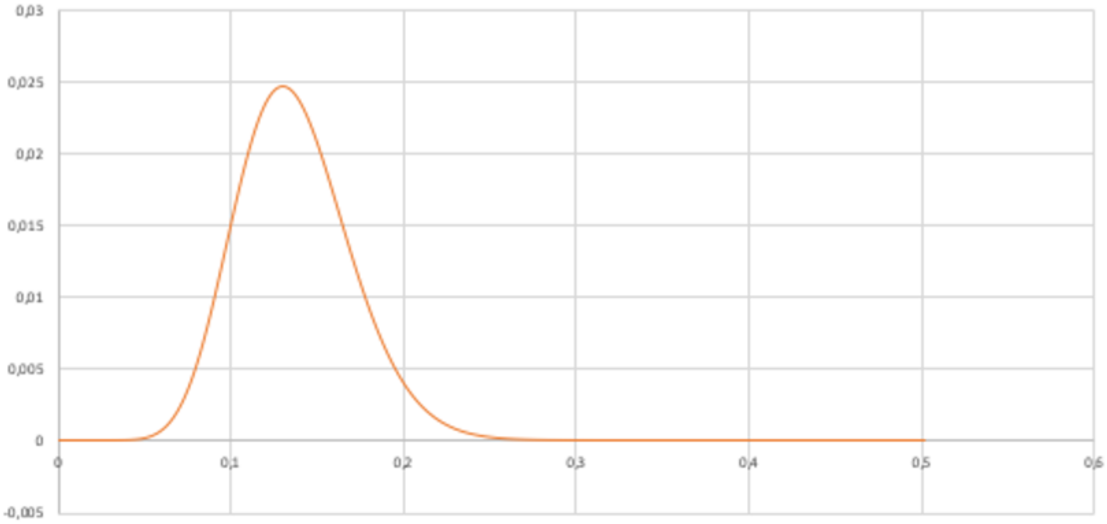


Figure 8: Example of a Beta Distribution from the Cyber Exercise after three exercise runs

4.11 Strategies to Deal with Risk

In the information risk management process by SFS-ISO 27005 (2011), risk treatment follows risk assessment. Sweeting (2011) lists four categories of how to treat risks: reduce, remove, transfer, or accept. (Sweeting, 2011). Virtue and Rainley (2014) recognize four similar categories: mitigate, avoid, transfer, and accept. (Virtue et al, 2014) Selecting the appropriate action for risk treatment depends on the company, its approach to dealing with risk, and its risk tolerance. Depending on the available information about the risks the chosen approach may change. For example, the company may decide to retain risks that are seen as insignificant without any actions, or the company may decide to invest on new hardware or safeguards, or set up redundancies to mitigate risks that are seen as worth the effort. On the other hand, if a risk is too great to retain and mitigation efforts are deemed too expensive, the company may seek to transfer the risk to someone else by outsourcing a system or buying insurance.

The company will need methods to quantify and calculate the value of the risks in its risk portfolio in order to assess returns on mitigation investments, adequacy of its insurance coverage, and the amount of risk retained.

4.11.1 Reducing or Mitigating a Risk

Reducing or mitigating a risk means taking measures to lower its probability or limiting its impact by setting up safeguards or redundancies (Sweeting, 2011; Virtue et al, 2014). These redundancies or safeguards could be, for example, setting up load balancing clusters to reduce the risk of a service outage due to overload or a hardware failure, or investing in systems to better detect unauthorized breaches into the environment in order to reduce the risk of a data breach. Knowing the value of the risk being reduced before and after actions is important in order to understand if the mitigation succeeded in reducing the risk by the required amount.

4.11.2 Removing or Avoiding a Risk

Removing or avoiding a risk means eliminating a risk completely by not taking up a project that would create the risk, or by decommissioning a system to remove a known vulnerability (Sweeting, 2011; Virtue et al, 2014). An example of removing a risk could be removing an old server version from the environment after the software vendor has ceased support. The risk of a breach increases constantly while the server is in use, and the removal from the environment will effectively remove the risk entirely. A tradeoff must be made if some part of the company operations requires the server and updating is not possible.

4.11.3 Transferring a Risk

Transferring a risk means outsourcing the consequences of a risk to someone else (Sweeting, 2011; Virtue et al, 2014). Information Technology Infrastructure Library (ITIL) 4 Edition (2019) includes the transfer of risk from the service consumer to the service provider in the definition of “Service” (ITIL 4 Edition, 2019). Outsourcing the maintenance of a system to a third party or buying a service is effectively a risk transfer. Conversely, the service consumer will also incur a different risk from, for example, the service provider going out of business; a risk which must be analyzed separately.

Another example is transferring the financial consequences of an incident to an insurance company through an insurance. In the case of insurance, knowing the value of the individual risks in order to assess the insurance coverage is important.

4.11.4 Retaining a Risk

Retaining or accepting a risk means taking or retaining the risk without doing anything, for example, because the expected severity is very small or because trying to avoid, transfer, or mitigate the risk would be more expensive than to suffer the consequences upon its realization (Sweeting, 2011; Virtue et al, 2014). An example of a company being forced to retain a risk could be an old server version that cannot be updated due to a function in the company depending on it. In this case the development of the risk should be followed closely and be continuously weighed against the benefits of retaining the risk, and the risk should be revisited at latest when the negative value of the risk exceeds the positive value gained from using the system.

5 Cyber Exercise

The research setting was selected to be the JAMK Cybersecurity exercise course during the spring of 2019. The exercise was a five ECTS credit joint course with about 90 Bachelor’s and Master’s level students from JAMK University of Applied Sciences and from the University of Jyväskylä. The participants were divided into six teams: one white team to plan and manage the game, one red team to attack the different organizations in the game, and four blue teams with their respective in-game organizations and environments to defend against the attacks by the red team. The target environment for this research was a Security Operations Center (later YSOC) operated by one of the blue teams, responsible for offering cyber security monitoring services to two other blue team environments: a bank (YBANK) and an online retailer (YSHOP). The game setup required setting up the technical environment for YSOC, as well as various YSOC business procedures required for the operation of the in-game organization, including a risk management system. A quantitative risk management system was implemented alongside the environment setup for the first time in the history of cyber security exercises at JAMK.

5.1 Phases of the Exercise

For YSOC, the exercise proceeded in roughly two partly overlapping phases; preparation phase and exercise phases. In the preparation phase, the eleven members of YSOC team organized so that a management team consisted of six Master's level students responsible for designing the technical and organizational structure, as well as risk management and internal procedures for YSOC organization. The remaining five bachelor's level students were mainly responsible for building and implementing the environment, while also encouraged to take part in management meetings. Similarly, the management team also heavily participated in technical set up of the environment. In the preparation phase, each member of the team also created their own in-game character with a background story and personality to use during the exercise phase.

The exercise phase was played out during two different weekends. The first weekend consisted of one test run in which the technical environments, internal procedures, and the ability to defend against the red team's attacks were tested out. The rest of the time was reserved for discussing the test run and planning modifications based on the experience and was followed by a month of time to improve the environment and defences. The second weekend consisted of two exercise runs which were identical to the test run, in that during each exercise run YSOC worked to identify attacks in its own and its customers' environments, inform the customers of the potential attacks and to defend its own environment. The in-game YSOC organization was divided into four different teams: a company management team responsible for overseeing the big picture during the exercise runs while taking care of social media and communications with the public and other stakeholders, a technical administration team responsible for taking care of YSOC's internal systems, and two customer specific teams tasked with monitoring and responding to threats observed within the customer environments.

5.2 The Risk Management Model

A risk management model for YSOC was developed during the last week of February in 2019 by first brainstorming among the management team and defining the risks that could be foreseen in the beginning of the exercise. The SOC had to be implemented within two months, and only limited time could be allocated for the risk management model. Decision was made to approach risk management from the point of view of the organization's core business offering: to offer threat modelling, detection, identification and troubleshooting, defence tactics, and threat information to the organization's customers. In practice, the organization offered Security Operation Centre services to its customers according to a Service Level Agreement (SLA) and all risks were identified through their potential to affect the fulfilment of these Service Level Agreements.

Six different categories of risks were identified.

- Ability to receive log data
- Ability to monitor customer environment
- Ability to assess and categorize the incidents
- Ability to analyse the data effectively
- Ability to notify the customer
- Ability to keep the organization's own environment securely up and running

Various risks were identified for each category to form YSOC risk portfolios, and the management team assigned probabilities and ranges of expected losses for the risks resembling a qualitative risk matrix, but instead of a low-average-high type of assignment, the probabilities were assigned as percentage values indicating how sure the management team was that the specific risk would realize during an exercise run.

The expected ranges of losses were assigned as two monetary values with expectation that per exercise run, the organization's profit would be $100.000 and when realizing, the respective risks would reduce that profit by an amount inside the specified range. The lower number would be the minimum expected loss incurred by the event and the higher number would be the maximum expected loss with 10 % change that the actual loss would be outside the specified range, arriving to 90 % confidence intervals.

The Management Team also defined the risk tolerance for the organization by each member providing estimates in percentage values of how comfortable they would be with losses in four different scenarios. The procedure for expressing risk tolerance is described in detail later.

The identified risks in the Ability to Receive Log Data portfolio included time inconsistencies within the network, which could prevent communication protocols from working properly or introduce inconsistency in log data, missing agents or configurations on hosts resulting in logs not being sent to the SOC, network connectivity between the SOC and its clients being unavailable, and the log aggregation and analysis system being unavailable.

In the Ability to Monitor Customer Environment portfolio, the major risks were the monitoring systems or their parts being unavailable in YSOC or the customer environments, and emails from the monitoring systems not being generated in the YSOC ticketing system.

In the Ability to Assess and Categorize the Incidents portfolio, the recognized risks included technical analysis tools being unavailable, operators not following the prescribed procedures, and the procedures not being efficient, with each risk leading to information not being passed along or waste of time doing so.

The Ability to Analyse the Data Effectively included the analysis tools going down, corruption of data, and the presentation of data in the systems not being efficient with the consequence of the operators not being able to view or analyse the data. Risks also related to the operator distraction or otherwise being incapacitated to perform.

The Ability to Notify the Customer risk portfolio included the process not being adhered to, unavailability of email system and the inability to reach the customers by phone.

The Ability to Keep YSOC Environment Securely up and Running portfolio included data leaking to outside through technical vulnerabilities or misconfiguration, data loss due to an attack on the system, due to a misconfiguration or a mistake, data integrity compromise through a configuration error, an admin error causing downtime, and, related to the personality of a particular in-game character, the CEO leaking password to twitter.

Due to the limited time in the preparation phase of the exercise, the management team did not go through any calibration training, and therefore the estimates were expected to contain errors. Calculation of calibration scores was also impossible without calibration testing. However, the members gave their assessments on the probabilities, ranges of expected losses, and their risk tolerance individually without consulting each other, and averages of their estimates were computed to arrive to the team's common perception of the risks and the risk tolerance.

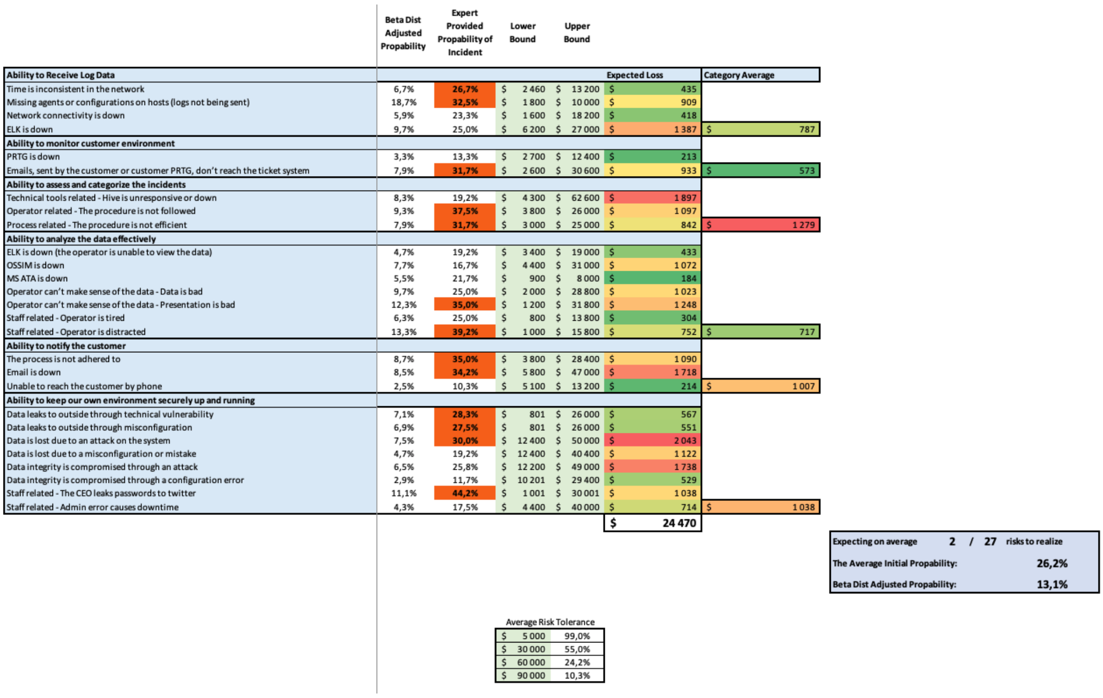
The risk management model also included updating the risk estimates with beta distribution using data about the realized risks gathered during the test run and the exercise runs. How much and to which direction this data was going to correct the estimates was interesting to see and the results were quite surprising.

Two distinct tools were used in the assessment and measurement of risk: Monte Carlo analysis for running thousands of simulations of all of the given risk probability and impact combinations, and Beta Distribution to update the original assessments over time with the actual data collected after each exercise run in the form of "yes/no" answera to the question "did this risk realize during the exercise run" to each risk in the risk portfolio respectively. All data was recorded and manipulated in an excel tool built to house the model.

5.3 The Excel Tool

The Excel tool used to calculate the risks during the cyber exercise was adapted from the example provided by Hubbard (2016) in excel format. In the beginning of the exercise each management team member gave their individual assessments of each risk item identified in the foundation of the risk management model, and averages of these estimates were then calculated to be used as the baseline throughout the exercise. After the test run and each exercise run, the risks that had realized during that run were noted and recorded in the tool, the tool then used Beta Distributions to adjust the previous values using this data. In calculating the expected loss figures the Beta Distribution adjusted probabilities and the lower bound and the upper bound of the expected losses were used.

Figure 9 shows the dashboard tab of the Excel tool showing all the risks and their respective current risk levels after the last run of the exercise. Figure 9 also shows the original averages of the evaluations by the management team and the beta distribution adjusted probabilities after the last exercise run. The figure also illustrates that with these data adjusted assessments, 2 out of the 27 recognized risks could be expected to realize during the next exercise run. The average risk probability after analyzing each test run dropped from the original average 26.2% assessment by the Management Team to an adjusted 13.1%, showing that the original assessments were far bleaker than what was actually observed during the exercise. The same illustration can be found in appendix 2 in greater size.



	Beta Dist Adjusted Propability	Expert Provided Propability of Incident	Lower Bound	Upper Bound		
Ability to Receive Log Data					Expected Loss	Category Average
Time is inconsistent in the network	6,7%	**26,7%**	$ 2 460	$ 13 200	$ 435	
Missing agents or configurations on hosts (logs not being sent)	18,7%	**32,5%**	$ 1 800	$ 10 000	$ 909	
Network connectivity is down	5,9%	23,3%	$ 1 600	$ 18 200	$ 418	
ELK is down	9,7%	25,0%	$ 6 200	$ 27 000	$ 1 387	$ 787
Ability to monitor customer environment						
PRTG is down	3,3%	13,3%	$ 2 700	$ 12 400	$ 213	
Emails, sent by the customer or customer PRTG, don’t reach the ticket system	7,9%	**31,7%**	$ 2 600	$ 30 600	$ 933	$ 573
Ability to assess and categorize the incidents						
Technical tools related - Hive is unresponsive or down	8,3%	19,2%	$ 4 300	$ 62 600	$ 1 897	
Operator related - The procedure is not followed	9,3%	**37,5%**	$ 3 800	$ 26 000	$ 1 097	
Process related - The procedure is not efficient	7,9%	**31,7%**	$ 3 000	$ 25 000	$ 842	$ 1 279
Ability to analyze the data effectively						
ELK is down (the operator is unable to view the data)	4,7%	19,2%	$ 3 400	$ 19 000	$ 433	
OSSIM is down	7,7%	16,7%	$ 4 400	$ 31 000	$ 1 072	
MS ATA is down	5,5%	21,7%	$ 900	$ 8 000	$ 184	
Operator can’t make sense of the data - Data is bad	9,7%	25,0%	$ 2 000	$ 28 800	$ 1 023	
Operator can’t make sense of the data - Presentation is bad	12,3%	**35,0%**	$ 1 200	$ 31 800	$ 1 248	
Staff related - Operator is tired	6,3%	25,0%	$ 800	$ 13 800	$ 304	
Staff related - Operator is distracted	13,3%	**39,2%**	$ 1 000	$ 15 800	$ 752	$ 717
Ability to notify the customer						
The process is not adhered to	8,7%	**35,0%**	$ 3 800	$ 28 400	$ 1 090	
Email is down	8,5%	**34,2%**	$ 5 800	$ 47 000	$ 1 718	
Unable to reach the customer by phone	2,5%	10,3%	$ 5 100	$ 13 200	$ 214	$ 1 007
Ability to keep our own environment securely up and running						
Data leaks to outside through technical vulnerability	7,1%	**28,3%**	$ 801	$ 26 000	$ 567	
Data leaks to outside through misconfiguration	6,9%	**27,5%**	$ 801	$ 26 000	$ 551	
Data is lost due to an attack on the system	7,5%	**30,0%**	$ 12 400	$ 50 000	$ 2 043	
Data is lost due to a misconfiguration or mistake	4,7%	19,2%	$ 12 400	$ 40 400	$ 1 122	
Data integrity is compromised through an attack	6,5%	25,8%	$ 12 200	$ 49 000	$ 1 738	
Data integrity is compromised through a configuration error	2,9%	11,7%	$ 10 201	$ 29 400	$ 529	
Staff related - The CEO leaks passwords to twitter	11,1%	**44,2%**	$ 1 001	$ 30 001	$ 1 038	
Staff related - Admin error causes downtime	4,3%	17,5%	$ 4 400	$ 40 000	$ 714	$ 1 038
					$ 24 470	

Expecting on average	**2 / 27** risks to realize
The Average Initial Propability:	**26,2%**
Beta Dist Adjusted Propability:	**13,1%**

Average Risk Tolerance

$ 5 000	99,0%
$ 30 000	55,0%
$ 60 000	24,2%
$ 90 000	10,3%

Figure 9: The summary view to the combined risk portfolio after the last exercise run

5.3.1 Risk Tolerance

In order to find out the level of tolerable risk for the YSOC organization, each member of the Management Team gave their answers to four questions about their tolerance of YSOC losses. The members were asked to define in percentage values, how acceptable for them it was for the losses to exceed $5000, $30,000, $60,000, and $90,000 respectably during each exercise run. The averages of these answers were then used to define the total YSOC Risk Tolerance. According to the Management Team, 99 % chance to lose up to $5000, 55% chance to lose up to $30,000, 24.2% chance to lose up to $60,000, and finally 10.3% chance to lose up to $90,000 was deemed acceptable.

Figure 10 shows YSOC risk tolerance plotted in a graph on a linear scale along with a blue line showing risk neutrality. The graph shows that the YSOC management team took a somewhat conservative and risk averse stance, as for most of the time the risk tolerance curve runs below the risk neutrality line. In contrast, risk tolerance curve dominating the risk neutrality line would have shown a risk seeking stance.

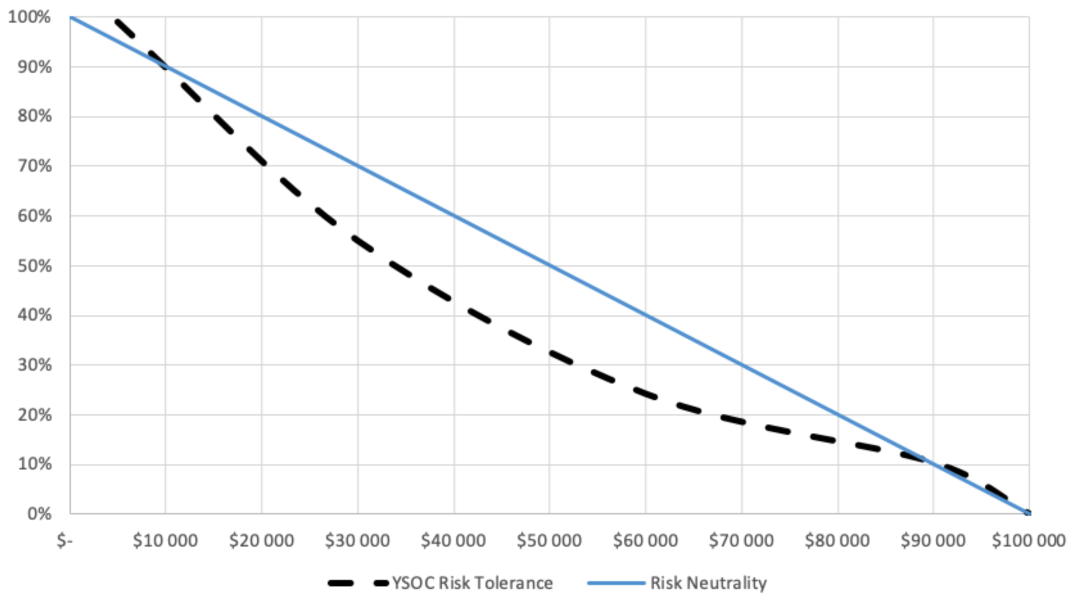
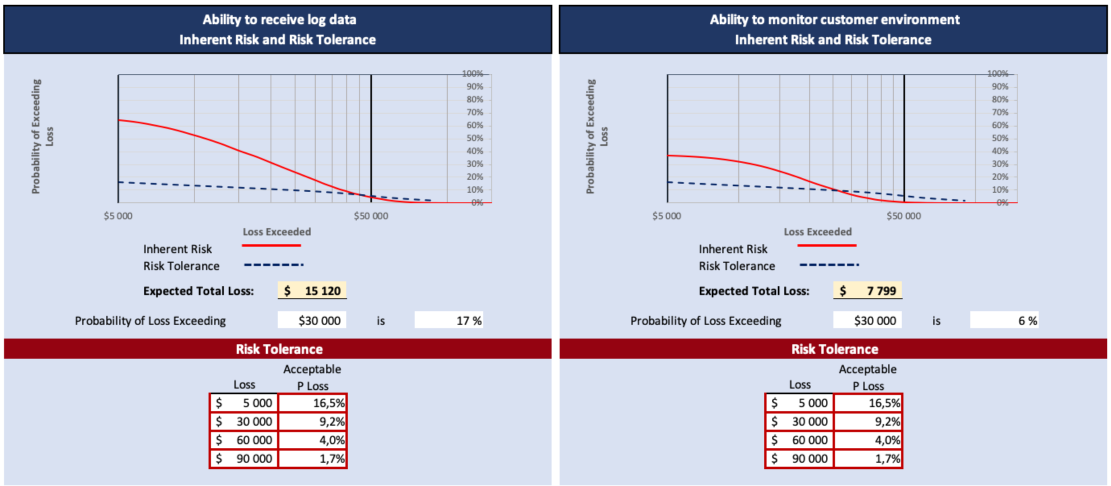


Figure 10: YSOC Risk Tolerance in Relation to Risk Neutrality

5.3.2 Individual Risk Portfolios

The combined YSOC risk portfolio included all risk portfolios described in chapter 5.2; in other words, the combined YSOC risk portfolio consisted of six sub-portfolios, such as “the ability to receive log data” and “the ability to monitor customer environment”. Under each sub-portfolio, the probabilities and ranges of impact of individual risks were used to calculate the expected total loss for that specific sub-portfolio, as well as the probability distribution of the calculated losses using Monte Carlo simulation. The total risk tolerance was also divided equally between each sub portfolio in order to distribute the tolerance consistently across the whole combined portfolio. The distributions could then be used to plot histograms for loss exceedance curves for each sub-portfolio and compare them against the risk tolerance curves on sub-portfolio and combined portfolio levels.

Figure 11 shows an example of the loss exceedance and risk tolerance curves in the initial analysis before building the environment had started on a logarithmic scale for two sub portfolios: “ability to receive log data” and “ability to monitor customer environment”. The expected loss curves dominating the risk tolerance curves showed high initial expectancy of failure by the management team.



Ability to receive log data
Inherent Risk and Risk Tolerance

Expected Total Loss: **$ 15 120**

Probability of Loss Exceeding $30 000 is 17 %

Risk Tolerance

Loss	Acceptable P Loss
$ 5 000	16,5%
$ 30 000	9,2%
$ 60 000	4,0%
$ 90 000	1,7%

Ability to monitor customer environment
Inherent Risk and Risk Tolerance

Expected Total Loss: **$ 7 799**

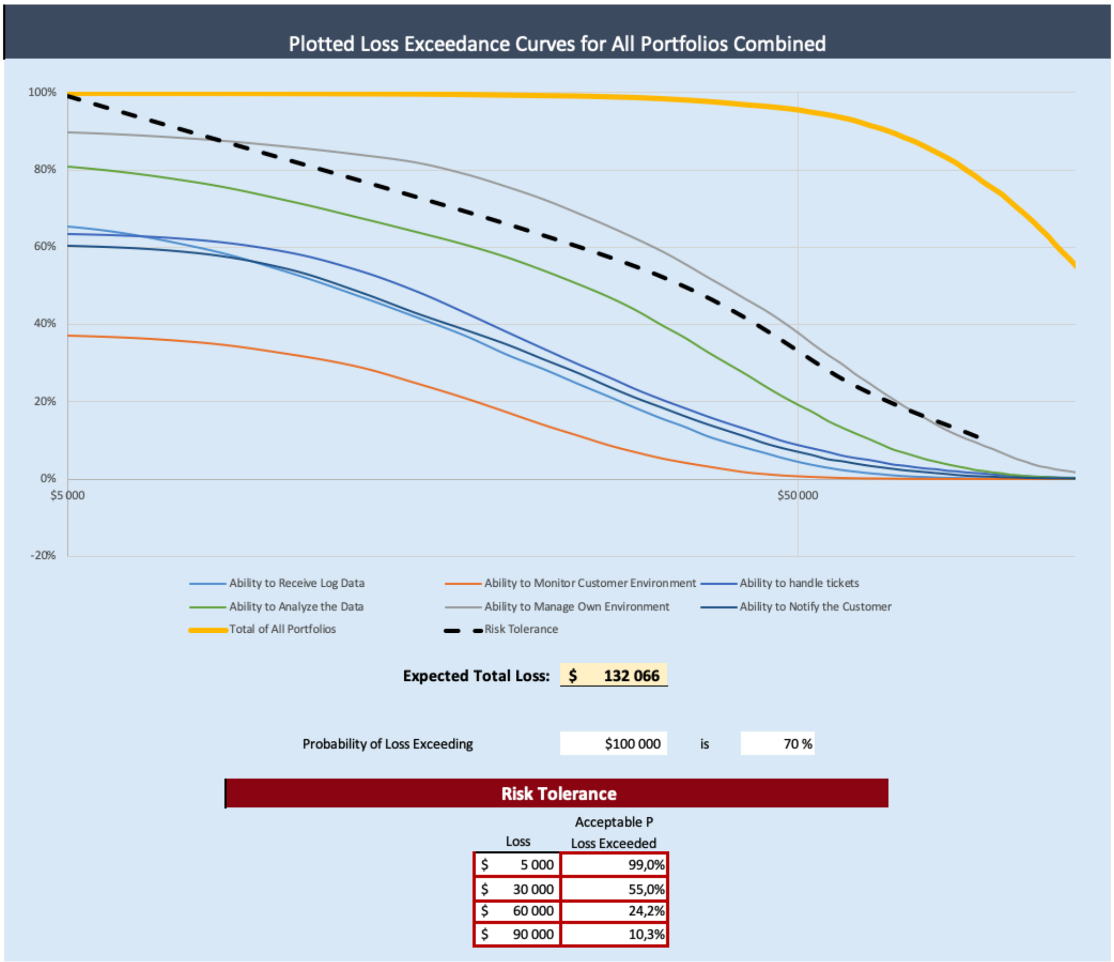
Probability of Loss Exceeding $30 000 is 6 %

Risk Tolerance

Loss	Acceptable P Loss
$ 5 000	16,5%
$ 30 000	9,2%
$ 60 000	4,0%
$ 90 000	1,7%

Figure 11: Loss exceedance and risk tolerance curves for two sub portfolios in the initial analysis

Figure 12, on the other hand, shows a combined view of all loss exceedance curves on a logarithmic scale compared against the total risk tolerance curve in the initial analysis. The yellow total risk line was far above the risk tolerance level and the expected total loss, defined as weighted average of the expected losses, and was clearly over the maximum profit of $100,000.


Plotted Loss Exceedance Curves for All Portfolios Combined
100%
80%
60%
40%
20%
0%
-20%
$5 000
$50 000
Ability to Receive Log Data
Ability to Monitor Customer Environment
Ability to handle tickets
Ability to Analyze the Data
Ability to Manage Own Environment
Ability to Notify the Customer
Total of All Portfolios
Risk Tolerance

Expected Total Loss: **$ 132 066**

Probability of Loss Exceeding $100 000 is 70 %

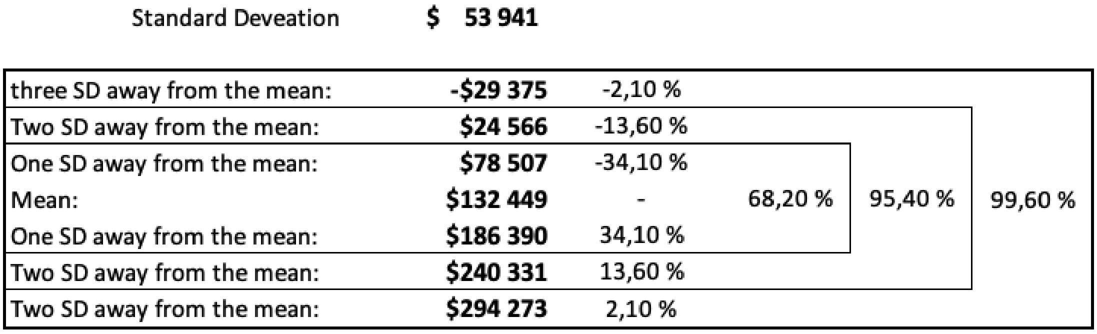
Risk Tolerance

Loss	Acceptable P Loss Exceeded
$ 5 000	99,0%
$ 30 000	55,0%
$ 60 000	24,2%
$ 90 000	10,3%

Figure 12: Combined view of all loss exceedance curves compared against the total risk tolerance in the initial analysis

In fact, the calculations showed an expected 70 % chance of losses exceeding $100 000. In addition, the standard deviation for the distribution of the expected loss exceedance was $53 941. Therefore, the expected losses were expected to be between $78 507 and $186 309 which are one standard deviation away from the mean on each side of the distribution with 68,2 % probability of the losses falling between this range. Two standard deviations away from the mean were $24 566 and $240 331 with 95,4 % probability of the losses falling between this range. Table 4 shows the distribution of the expected loss exceedance values in the initial analysis.

Table 4: Distribution of the expected loss exceedance with standard deviations in the initial analysis



Standard Deveation **$ 53 941**

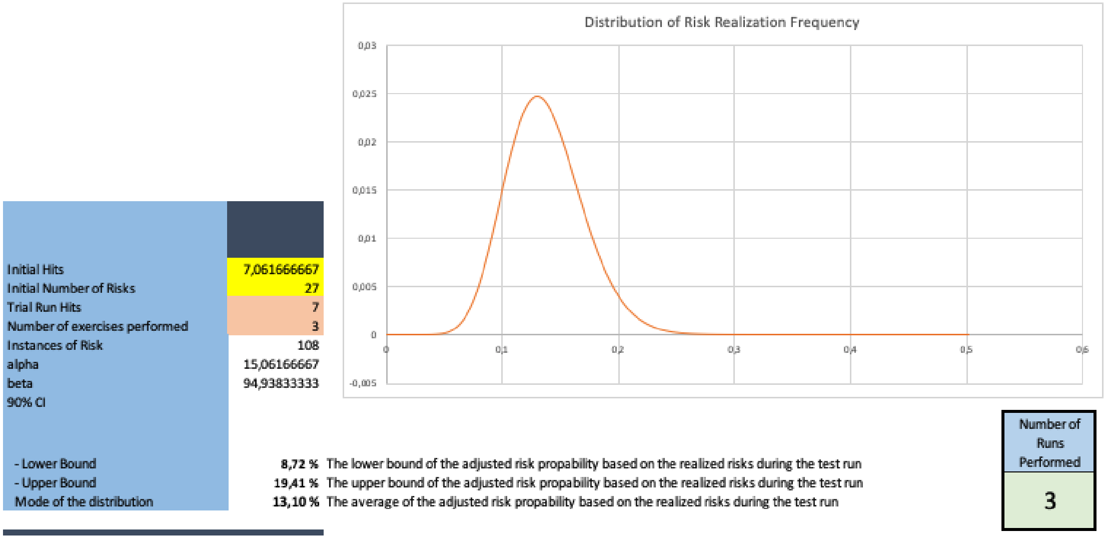
three SD away from the mean:	**-$29 375**	-2,10 %			
Two SD away from the mean:	**$24 566**	-13,60 %			
One SD away from the mean:	**$78 507**	-34,10 %			
Mean:	**$132 449**	-	68,20 %	95,40 %	99,60 %
One SD away from the mean:	**$186 390**	34,10 %			
Two SD away from the mean:	**$240 331**	13,60 %			
Two SD away from the mean:	**$294 273**	2,10 %			

The goal for the management team was therefore to lower the total risk curve to match the risk tolerance curve. Pushing the total risk curve as far below risk tolerance curve as possible may seem better since this indicates lower amount of risk, however, in financial risk management this scenario is not optimal. Paul Sweeting (2011) describes how lines can be drawn for investment portfolios to represent combinations of risk and return to create indifference curves, also known as risk tolerance curves, and how the "point at which an indifference curve is tangential to the efficient frontier defines the optimal portfolio" (Sweeting, 2011). Similarly, in figure 12, considering the risk tolerance of the management team, the optimal balance between investment and risk is reached when the loss exceedance curve matches the risk tolerance curve as closely as possible, indicating that the management team is comfortable with the amount of retained risk in the environment and any further investments would be sub-optimal.

5.3.3 Adding Beta Distribution to the Monte Carlo Model

In the cyber exercise the hits in the Beta Distribution are the realized risks during the test run and the exercise runs. The Beta Distribution allows for estimating the proportion of the hits and misses within the whole population of realized and unrealized risks. As stated in chapter 4.10.4, expecting all risks or none at all to realize in the beginning of the analysis is not necessary when using Beta Distribution. The initial expectation could be set to the percentage values for expectation expressed by the management team. The Beta Distribution function takes the expectation values as one input along with the elapsed time, or the number of performed exercise runs in this case, and updating the model could then be simply conducted by recording the number of times each recognized risk was observed as realized during the exercise runs. The frequency distributions were then adjusted for each of the risks, and the overall distribution of risk realization frequency for the whole portfolio was re-calculated and fed back into the Monte Carlo simulations to arrive to the Beta Distribution adjusted probability for observing any of the risks in the portfolio realizing in the next run. In a risk model for a real company, intervals such as company years or quarters would be used instead of runs performed. For the initial estimates on probabilities, expert estimates could be used on their own or they could be based on records of how often the same or similar risks have realized for comparable companies.

Figure 13 shows the Beta Distribution for the combined risk portfolio after three exercise runs. The mode of the distribution was used as the expected probability of risk realization in the risk model and all individual weighted probabilities were adjusted accordingly.



Initial Hits	7,061666667	
Initial Number of Risks	27	
Trial Run Hits	7	
Number of exercises performed	3	
Instances of Risk	108	
alpha	15,06166667	
beta	94,93833333	
90% CI		
- Lower Bound	8,72 %	The lower bound of the adjusted risk propability based on the realized risks during the test run
- Upper Bound	19,41 %	The upper bound of the adjusted risk propability based on the realized risks during the test run
Mode of the distribution	13,10 %	The average of the adjusted risk propability based on the realized risks during the test run

Number of Runs Performed
3

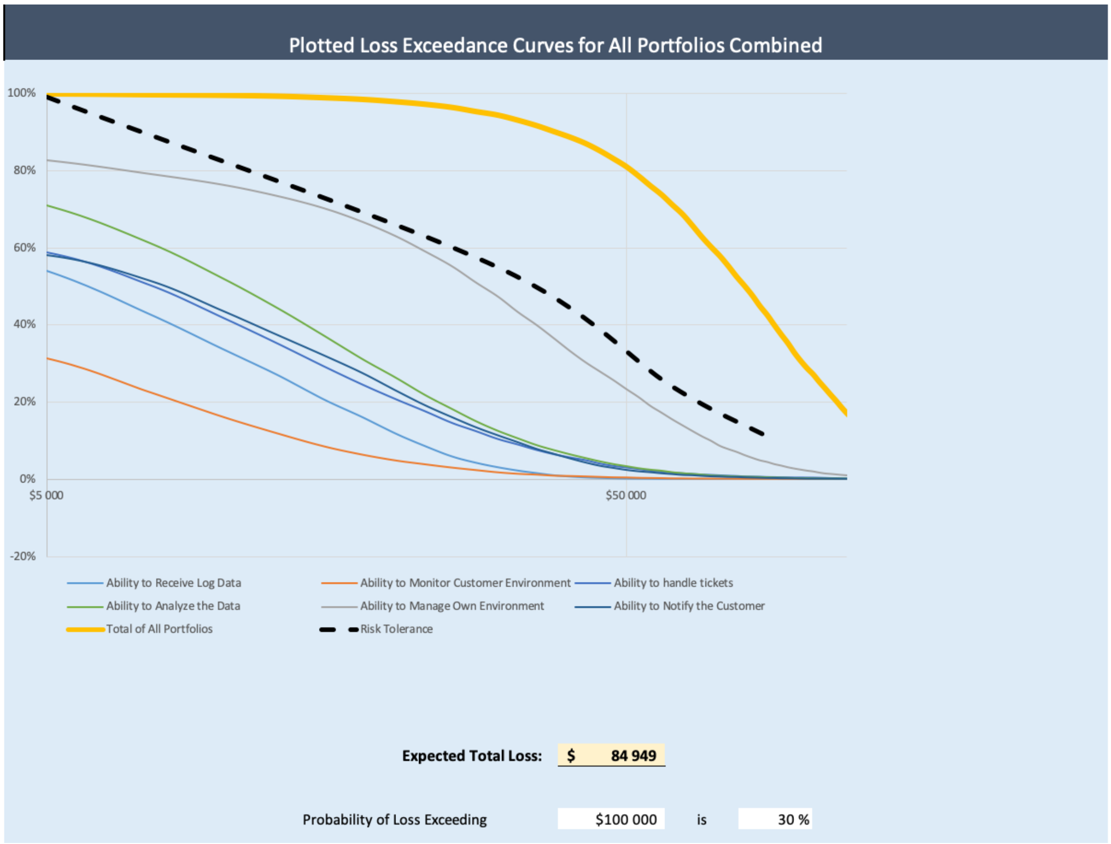
Figure 13: Beta Distribution for the Risk Portfolio After Three Exercise Runs

In the graph, the more data points there are to feed into the Beta Distribution the thinner and taller the distribution gets signifying the amount of uncertainty in the model decreasing. This mathematically shows how even if the original expert estimates of the probability of the events were unreliable, the model reduces the amount of uncertainty automatically as more and more data become available.

6 Results of the Exercise

Even though the initial values used in the exercise were almost arbitrary in nature, the fact that the yellow loss exceedance curve in Figure 12 clearly dominates the risk tolerance curve shows that the risks perceived by the management team were much higher than the declared tolerance. This was also highly understandable, as the nature of the task was to build the whole organization along with a functioning environment from scratch within limited time.

More interesting, however, is how quickly the loss exceedance curves moved towards the risk tolerance curve. The risk assessment was revised just before the first test run, in which the risks were seen to be considerably lower than in the initial review. Only a month of time was provided for setting up the environment, and the odds of getting the environment ready in time caused concern during the preparation phase. As a result, the drop in the perceived risk between the first and the second analysis can mostly be explained by the increasing confidence of the management team in the team’s ability to get the environment ready to perform during the exercise rather than the actual risks going down. The loss exceedance curves before the test run shown in Figure 14 illustrates the increasing confidence.



Expected Total Loss: $ 84 949

Probability of Loss Exceeding $100 000 is 30 %

Figure 14: Combined view of all loss exceedance curves compared against the total risk tolerance in the analysis before the test run

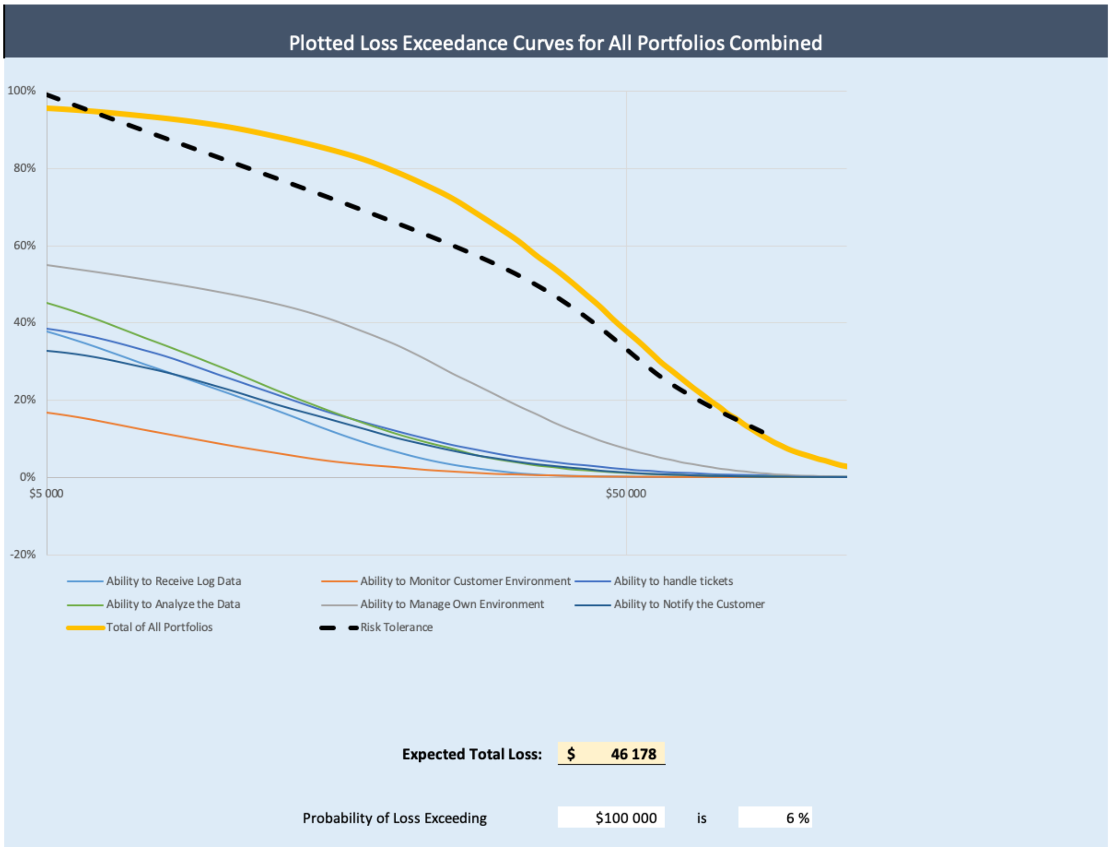
The total expected loss had dropped from about $132 000 in the initial analysis a month earlier to about $85 000. The probability of the loss exceeding $100 000 also dropped from 70 % to 30 %, and the standard deviation in the distribution of the probable loss exceedance fell from $53 941 to $41 027. Table 5 illustrates the distribution of the expected loss exceedance values in the analysis before the test run.

Table 5: Distribution of the Expected Loss Exceedance with Standard Deviations in the Analysis Before the Test Run

Standard Deveation $ 41 027

three SD away from the mean:	-$38 131	-2,10 %			99,60 %
Two SD away from the mean:	$2 896	-13,60 %		95,40 %	
One SD away from the mean:	$43 922	-34,10 %	68,20 %		
Mean:	$84 949	-			
One SD away from the mean:	$125 976	34,10 %			
Two SD away from the mean:	$167 003	13,60 %			
Two SD away from the mean:	$208 030	2,10 %			

After the test run was completed, the second tool, the Beta Distribution, was used on the risk model to adjust the risk levels based on which risks had been observed realizing during the run. This adjustment was the first time a mathematical tool was applied to incorporate empirical data to the risk analysis, and the results showed a considerable drop in the risk levels. In fact, the drop was considerable enough to settle the total loss exceedance curve just above the declared Risk Tolerance curve. The updated loss exceedance curves after the test run can be seen in Figure 15.



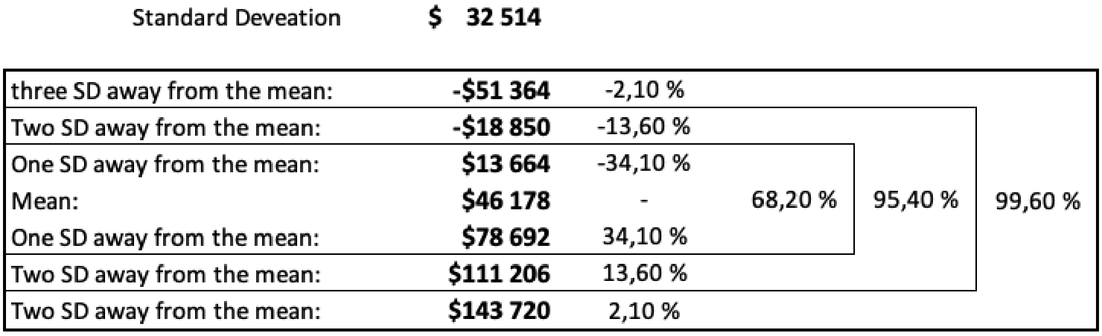
Expected Total Loss: $ 46 178

Probability of Loss Exceeding $100 000 is 6 %

Figure 15: Loss Exceedance curves after the test exercise run

The total expected loss was adjusted from the management team's expectation of about $85 000 to about $46 000. The probability of the loss exceeding $100 000 dropped from 30 % to just 6 %, and the standard deviation dropped from $41 027 to $32 514. Table 6 illustrates the distribution of the expected loss exceedance values in the analysis after the test run.

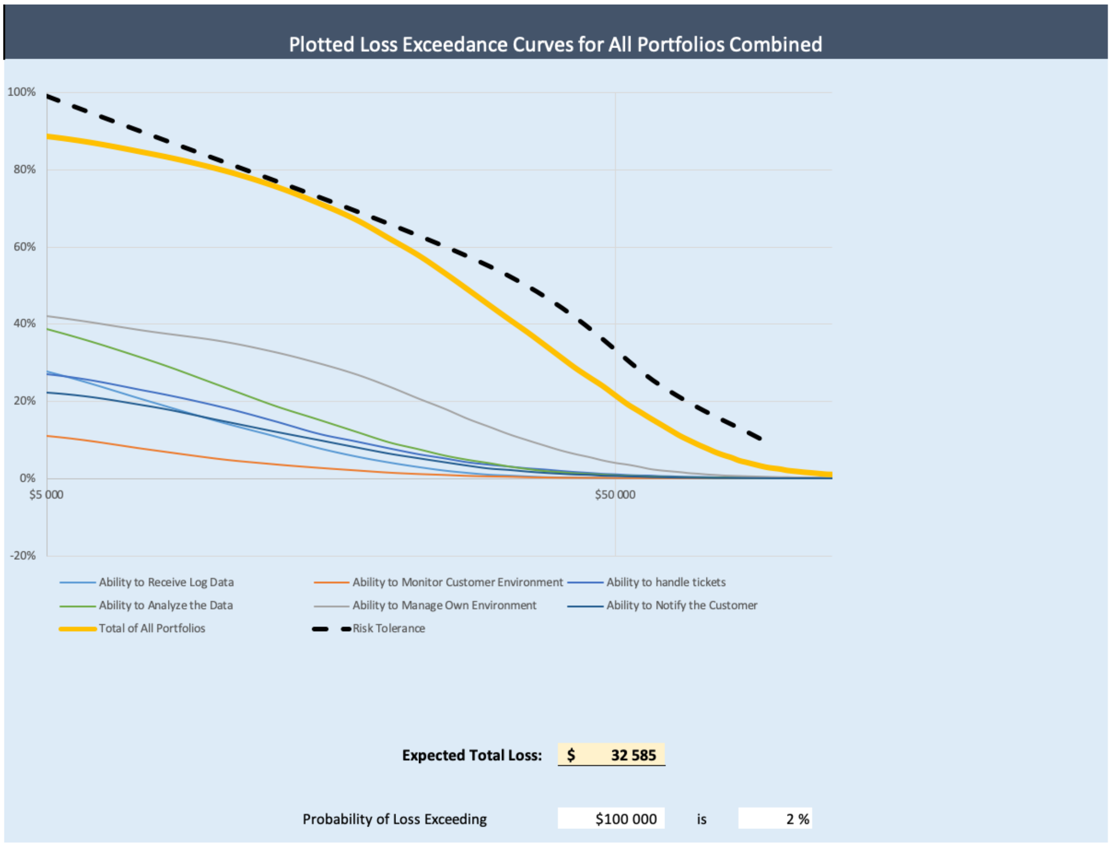
Table 6: Distribution of the expected loss exceedance with standard deviations in the analysis after the test run



Standard Deveation **$ 32 514**

three SD away from the mean:	**-$51 364**	-2,10 %			99,60 %
Two SD away from the mean:	**-$18 850**	-13,60 %		95,40 %	
One SD away from the mean:	**$13 664**	-34,10 %	68,20 %		
Mean:	**$46 178**	-			
One SD away from the mean:	**$78 692**	34,10 %			
Two SD away from the mean:	**$111 206**	13,60 %			
Two SD away from the mean:	**$143 720**	2,10 %			

After the test run, the loss exceedance curve settled a little above the risk tolerance curve indicating that the risk, the probabilities adjusted once with data from the test run using Beta Distribution and the impacts estimated by the management team, was much closer to the management team's tolerance as initially expected. The team identified a number of improvements during the test run, which were addressed within the month before the two exercise runs. The analysis was next updated after the first exercise run by incrementing the counters of observed times of realization for the risks that realized during the exercise run. The update resulted in the probabilities being adjusted further down, which pushed the loss exceedance curves below the risk tolerance level. The updated loss exceedance curves after the first exercise run can be seen in Figure 16.



Expected Total Loss: $ 32 585

Probability of Loss Exceeding $100 000 is 2 %

Figure 16: Loss exceedance curves after the first exercise run

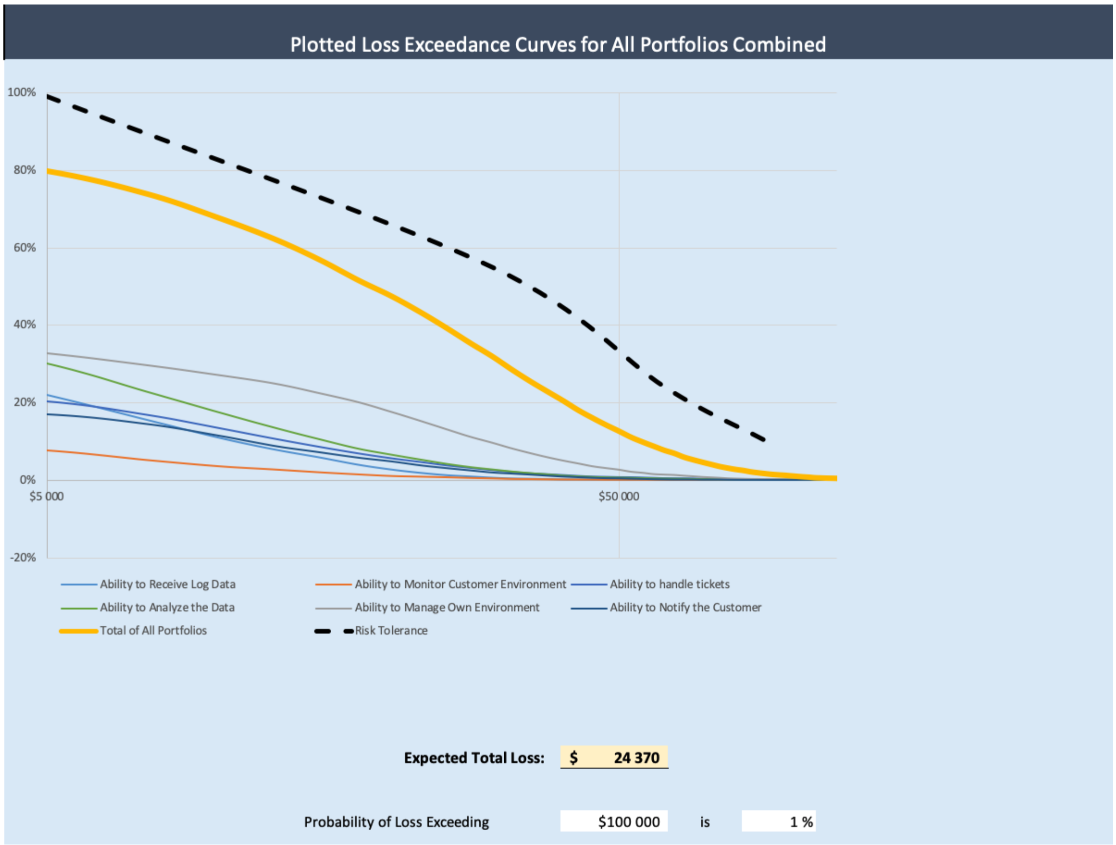
The total expected loss was adjusted from about $46 000 after the test run to about $33 000. The probability of the loss exceeding $100 000 dropped further from 6 % to 2 % and the standard deviation dropped from $32 514 to $26 646. Table 7 illustrates the distribution of the expected loss exceedance values in the analysis after the first exercise run.

Table 7: Distribution of the expected loss exceedance with standard deviations in the analysis after the first exercise run

Standard Deveation **$ 26 646**

three SD away from the mean:	**-$47 353**	-2,10 %			99,60 %
Two SD away from the mean:	**-$20 707**	-13,60 %		95,40 %	
One SD away from the mean:	**$5 939**	-34,10 %	68,20 %		
Mean:	**$32 585**	-			
One SD away from the mean:	**$59 231**	34,10 %			
Two SD away from the mean:	**$85 877**	13,60 %			
Two SD away from the mean:	**$112 523**	2,10 %			

The Beta Distributions were updated for the last time after the second exercise run, after which the probabilities and the loss exceedance curves were pushed further down below the risk tolerance curve. The relative rate of the changes slowed down, indicating that the Beta Distribution was settling at the correct level of event probabilities. The updated loss exceedance curves after the second exercise run are displayed in Figure 17.



Expected Total Loss: **$ 24 370**

Probability of Loss Exceeding $100 000 is 1 %

Figure 17: Loss exceedance curves after the second exercise run

The total expected loss was adjusted from about $33 000 after the first exercise run to about $24 000. The probability of the loss exceeding $100 000 also dropped from 2 % to 1 % and the standard deviation also went down a little from $26 646 to $23 324. Table 8 illustrates the distribution of the expected loss exceedance values in the analysis after the second exercise run.

Table 8: Distribution of the expected loss exceedance with standard deviations in the analysis after the second exercise run

Standard Deveation **$ 23 324**

three SD away from the mean:	**-$45 602**	-2,10 %			
Two SD away from the mean:	**-$22 278**	-13,60 %			
One SD away from the mean:	**$1 046**	-34,10 %			
Mean:	**$24 370**	-	68,20 %	95,40 %	99,60 %
One SD away from the mean:	**$47 694**	34,10 %			
Two SD away from the mean:	**$71 018**	13,60 %			
Two SD away from the mean:	**$94 342**	2,10 %			

As seen in the results, the Beta Distribution is a tool that can make large corrections when data is scarce, and the large drop of the risk in the mathematically adjusted risk portfolios showed clearly how much the management team had overestimated the risks. The test run and the two exercise runs showed similar drops in the risks in all sub-portfolios. All of the measurement points (test and exercise runs) for each sub-portfolio can be seen in Figure 18.

	Average Probability of an Incident	**Expected number of risks to realize**	**Expected Average Loss**	**Ability to Receive Log Data**	**Ability to monitor customer environment**	**Ability to assess and categorize the incidents**	**Ability to analyze the data effectively**	**Ability to notify the customer**	**Ability to keep our own environment securely up and running**
Initial review	25,00 %	7	$ 132 056	$ 3 854	$ 3 953	$ 6 318	$ 4 156	$ 5 724	$ 5 440
Review before test run	26,20 %	7	$ 85 237	$ 2 139	$ 2 301	$ 4 055	$ 2 063	$ 4 056	$ 4 163
Adjusted after test run	20,50 %	4	$ 46 268	$ 1 403	$ 1 155	$ 2 572	$ 1 129	$ 2 028	$ 2 080
Adjusted after exercise day 1	17,30 %	3	$ 32 521	$ 996	$ 765	$ 1 710	$ 957	$ 1 350	$ 1 392
Adjusted after exercise day 2	13,10 %	2	$ 24 470	$ 787	$ 573	$ 1 279	$ 717	$ 1 007	$ 1 038

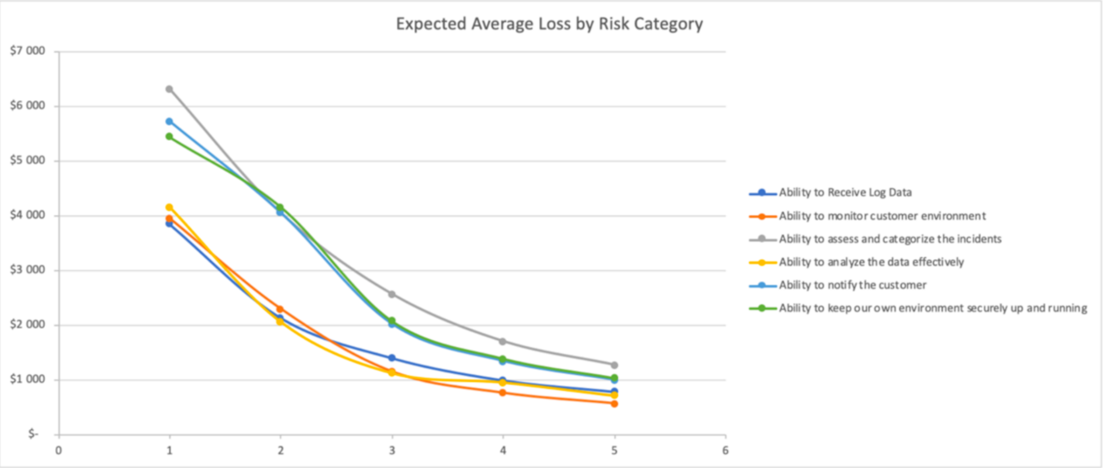


Figure 18: Development of risk in each sub-portfolio during the exercise

The graph above shows how the risk level falls after each measurement, but the rate of decrease also decreases each time as the true level of risk is reached. Due to the quantitative nature of the model, the data can also be displayed in an aggregated form. Figure 19 shows the total expected loss figures of the combined portfolio at each measurement point. The figure also shows a power trendline, displaying where the risk level was expected to settle had there been more exercise runs. It is noteworthy that this settling happened just after a couple of exercise runs, illustrating how quickly the Beta Distribution reacted to the new data, although only very few measurement points were available.

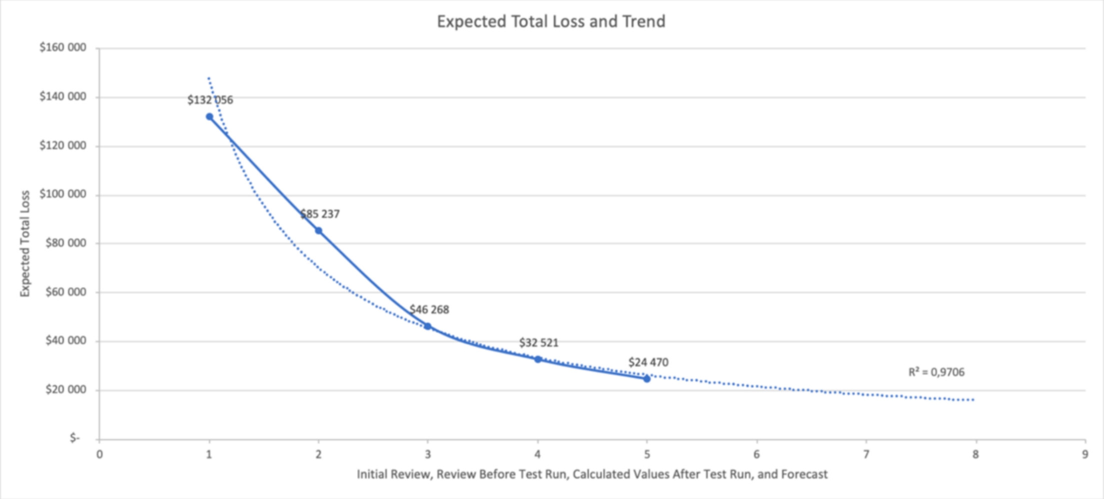


Figure 19: Development of risk in at the combined portfolio level with trendline

After just six exercise runs the power trendline places the total expected loss figure at around $17.000 or $18.000, or around 7 % to 8 % of the initial risk assessment figure by the management team, and well below the original risk tolerance.

7 Conclusions

7.1 Answers to the Research Questions

The answer to the first part of the research question “is it feasible for a small company, considering an IT Risk Management model, to go for a quantitative approach in favor of a qualitative one" is affirmative based on the results of the experiment. The model and the initial risk estimations were created by the YSOC management team within a week in the beginning of the exercise using the excel sheets accompanying the How to Measure Anything in Cyber Security Risk book, so reading the book to get an understanding of the tools and subsequently using the ready-made templates cannot be seen as an overwhelming investment for a company when initiating a risk management function. During the exercise, the calibration methods were not used to train the management team to give calibrated estimates, but for a company implementing the model suggested in this work, assembling a team for providing the estimates and providing them calibration training is highly advisable. According to Hubbard (2016) this training can be provided in, for example, a half-day workshop. Updating probabilities using the Beta Distribution and reviewing the impact ranges using the calibrated expert estimates is also easy to incorporate into for example in an annual risk management review workshop.

A risk management model is also only valuable to a company to the extent it is useful in supporting decisions (Freund, 2015). The decisions about whether to invest into control or mitigation and how much will be made regardless of the analysis deployed. If no analysis or information is available to the decision makers, the decisions will be based on the decision makers' personal views on how needed the controls or mitigations are. Also, not doing anything about the risks due to lack of information or understanding is a decision that will impact the environment. The purpose of risk management is to provide information and analysis to support decision making in a format that is useful and understandable to the decision makers. The analysis should address for example the current level of risk, the individual values of those risks, which way the risks are trending, the level of expected risk in the future, recommendations on the amounts of investments in relation to the level of risks and the risk tolerance of the company, targets for risk-mitigating investments, and expected returns on investments for those investments.

The quantitative model with its calculable value of each individual risk and portfolio, as well as graphs and projections on the level of risk, provide more valuable and actionable information to decision makers than qualitative risk matrices. The model can also be used to calculate returns on investments for mitigations or to estimate if insurance policies adequately transfer the risks. As the model automatically updates simply by incrementing the numbers of risks that realized, the outputs can also be used as Key Performance Indicators by the management to follow the development of the risks and to adjust their risk tolerance based on this situational awareness.

8 Discussion

8.1 Trustworthiness

The subject of quantitative risk management from the perspective of providing management actionable information through analysing and expressing risk in monetary units was explored in the theory part of this work. The theory part provided justification for the selected tools and their application in creating the quantitative risk management model for YSOC in the cyber exercise. This theoretical basis speaks for the trustworthiness of the selected tools, namely Monte Carlo analysis and Beta Distribution. The results of the case study showed rapid adjustment of the expressed probabilities of the risks as the exercise progressed, showing the utility of the Hubbard and Seiersen method in estimating risk probabilities in the context of IT risks. However, while review of the literature and research in to the calibration techniques make a compelling case for the utility of calibrating and using subjective subject matter experts in estimating ranges of expected losses, measuring the calibration of the management team was not conducted during the case study, and therefore the amount of error in the ranges of the impact estimates remains unknown. Still, this does not undermine the validity of the method or the results, as the same amount of unknown systematic error was present in all analyses throughout the exercise and the systematic error did not affect the probability side of the risk assessment calculations, which were responsible for the lowering risk throughout the exercise.

Another problem pertaining to the impact side of the equation is that the impacts were completely imaginary. The game did not originally contain any kind of notion of monetary loss and this had to be artificially created for the purpose of the risk analysis case study. According to Holton's (2004) definition of risk being a condition of self-aware individuals requiring both exposure and uncertainty, there was no risk included in the exercise as even if all events in the risk portfolios had realized, there would not have been any financial consequences to the individuals, and the exercise would still have been an educational experience and thus attaining the goals of the event. However, for the purpose of this study the artificial definition of impact was useful, and allowing for experience to be gained of implementing and using the methodology, so the secondary objective of the study to gain experience using a quantitative risk management methodology was also attained.

On a final note in the spirit of Freund (2015), as usefulness, practicality, and defendability should be the cornerstones of a risk management methodology, even with imperfect inputs, using the quantitative method can be more practical and yield more useful results than a qualitative five-by-five matrix. Use of a sound quantitative method also comes with clear understanding that the source of error is in the input data instead of the tools, providing defendability for the approach. When the source of error lies in the quality of the input data, if needed, investments can be made into improving the quality of the input data.

8.2 Future Research

The future work in this area should concentrate on providing Finnish companies with data and visibility into the true risks of data breaches. The reports published by foreign institutions showing global trends might easily be disregarded as not concerning Finnish companies or not giving accurate picture for the Finnish business environment. The current information available is also usually vague and can easily be brushed off as attempts to scare people into buying something of which usefulness or effectiveness there is very little to show for.

Research into the true probabilities and impacts of data breaches could be done, for example, by analysing public information such as news reports on data breaches, stock exchange releases, and annual reports of companies. Research could also be conducted by private cyber security companies or the National Cyber Security Centre by analysing their data and providing anonymized reports.

Another area of research could be finding out new ways to create awareness of the risk of data breaches in the 50% of the business leaders who, according to Helsinki Chamber of Commerce (2019), are not aware of the level of cyber risks facing their companies.

For a company adopting a quantitative risk modelling framework, suitable future research would include setting up a training system and training material for decision makers and analysts in expressing risk in terms of quantitative confidence intervals, and perhaps incorporating the Bayesian statistics based debiasing techniques suggested by Clemen et al (2002). The research by Hubbard and Seiersen shows that expressing risk in quantified manner is a learnable skill that can be acquired by almost anyone. Expression of risk also appears to be a skill that can be calibrated and measured with Hubbard (2016) providing descriptions of the methods to accomplish this. Setting up such a training system with monitoring the development of the experts' ability would be a good foundation for using expert opinions in the quantitative risk analyses.

Other future work could be to investigate using actuarial methods in measuring cybersecurity risks. To this end co-operation between institutions teaching actuarial methods or financial risk management and institutions teaching risk management in cyber security context would be needed. The rewards of this type of cooperation would be improved risk management methods for all kinds of organizations, as well as new types of synergies between educational institutions.

References

Ammar, A, Berman, K, & Sataporn A, 2007. *A review of techniques for risk management in projects*, Benchmarking: An International Journal 14 (1), 22-36

Brownley, C. 2013, *Multi-objective Decision Analysis: Managing Trade-offs and Uncertainty*, Business Expert Press.

Budescu, D., Broomell, S. & Por, H. 2007. *Uncertainty in the Reports of the Intergovernmental Panel on Climate Change*. Psychological Science. 2.

Chattamvelli, Rajan, et al. 2015. *Statistics for Scientists and Engineers*, John Wiley & Sons, Incorporated.

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H. & Stoddart, K. 2016. *A review of cyber security risk assessment methods for SCADA systems*. Computers & Security, 56, 1-27.

Cisco Annual Internet Report (2018-2023) White Paper. 2020. Article on Cisco's website. Accessed 23 August 2020. Retrieved from https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html.

Clemen, R. & Lichtendahl, K. 2002. *Debiasing Expert Overconfidence:A Bayesian Calibration Model*, Accessed 30 August 2020. Retrieved from https://faculty.fuqua.duke.edu/~clemen/bio/BayesianCalibration.pdf.

Creative Safety Supply. 2020. Article on creativesafetysupply website. Accessed 29 August 2020. Retrieved from https://www.creativesafetysupply.com/articles/history-of-the-kaizen-pdca-cycle/.

Fielder, A., König, S., Panaousis, E., Schauer, S., & Rass, S. 2018. *Risk Assessment Uncertainties in Cybersecurity Investments*, Games 9 (34)

Freund, J., Jones, J. 2015. *Measuring and Managing Information Risk*: A Fair Approach. Butterworth-Heinemann.

Gokhale, S. & Mullen, R. 2008. *Application of the Lognormal Distribution to Software Reliability Engineering*, chapter in ”Handbook of Performability Engineering” by Krishna B. Misra, Springer, London. Copyright Springer-Verlag London Limited 2008, Accessed 15 August 2020. Retrieved from https://link.springer.com/content/pdf/10.1007%2F978-1-84800-131-2.pdf.

Greenberg, A. 2018. *The Untold Story of NotPetya, the Most Devastating Cyberattack in History*. Article on Wired website. Accessed 23 August 2020. Retrieved from https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/.

Hackett, R. 2020. *China's Drive for Digital Currency Dominance*. Fortune, Aug/Sep, 76-79.

Holton, G. 2004. *Defining risk*. Financial Analysts Journal, 60(6), 19-25.

Hubbard, D. & Seiersen, R. 2016. *How to Measure Anything in Cybersecurity Risk*. 3rd. ed. Hoboken: Wiley.

IBM Security. 2019. *Cost of a data breach report 2019*. Michigan: Ponemon Institute.

ITIL 4 Edition, 2019. ITIL Foundation, Axelos limited.

Jones N., & Tivnan, B. 2018. *Cyber Risk Metrics Survey, Assessment, and Implementation Plan prepared for Department of Homeland Security*, The Homeland Security Systems Engineering and Development Institute Operated by The MITRE Corporation.

Jones, J. 2005. *An Introduction to Factor Analysis of Information Risk (FAIR)*. Risk Management Insight.

Kananen, J. 2014. *Laadullinen tutkimus opinnäytetyönä, Miten kirjoitan kvalitatiivisen opinnäytetyön vaihe vaiheelta [Qualitative research as a thesis, How do I write a qualitative thesis step by step]*. Jyväskylä: Jyväskylän ammattikorkeakoulu [University of Jyväskylä].

Koehler, D., Brenner, L., & Griffin, D. 2002. *The Calibration of Expert Judgment: Heuristics and Biases beyond the laboratory*, Accessed 30 August 2020. Retrieved from http://bear.warrington.ufl.edu/brenner/mar7588/Papers/koehlerbrennergriffin2002.pdf..

Krause, A. 2006. *Risk Management*, 32 (9). Emerald Publishing Limited Accessed 23 August 2020. Retrieved from http://ebookcentral.proquest.com/lib/jypoly-ebooks/detail.action?docID=275509

Matthews, T. N.d. *A Brief History of Cybersecurity*. Article on Cybersecurity Insiders website. Accessed 23 August 2020. Retrieved from https://www.cybersecurity-insiders.com/a-brief-history-of-cybersecurity/.

Miller, G. 2013. *FBI director warns of cyberattacks. Other security chiefs say terrorism threat has altered*. News article on The Washington Post website. Accessed 20 October 2019. Retrieved from https://www.washingtonpost.com/world/national-security/fbi-director-warns-of-cyberattacks-other-security-chiefs-say-terrorism-threat-has-altered/2013/11/14/24f1b27a-4d53-11e3-9890-a1e0997fb0c0_story.html.

Oskamp, S. 1965. *Overconfidence in Case-Study Judgments*. Journal of Consulting Psychology. 29 (3), 261– 265

Park, J., & Lee, S. 2018. *Probabilistic safety assessment-based importance analysis ofcyber-attacks on nuclear power plants*. Nuclear Engineering and Technology, 51 (1), 138-145.

Paul E. M. 1954. *Clinical versus Statistical Prediction. A Theoretical Analysis and a Review of the Evidence*. Minneapolis. University of Minnesota Press.

Pensworth, L. 2020. *2019 Internet Statistics, Trend & Data*. Article on Daily Wireless website. Accessed 23 August 2020. Retrieved from https://dailywireless.org/internet/usage-statistics/#Internet_Usage_Worldwide.

Politi M., Han P. & Col N. 2007. *Communicating the Uncertainty of Harms and Benefits of Medical Interventions*, Medical Decision Making, 27 (681).

Radanliev, P., De Roure, D. C., Nurse, J. R. C., Montalvo, R. M., Cannady, S., Santos, O., Maddox, L. T., Burnap, P., & Maple, C. 2019. *Future developments in standardisation of cyber risk in the Internet of Things (IoT)*, SN Applied Sciences, 2 (196).

Refsdal, A., Solhaug, B. & Stolen, K. 2015. *Cyber-Risk Management*. Cham: Springer.

Rhodes, A. 2015. *A Brief Summary of the Long History of Risk Management*. Article on Ventiv Website. Accessed 29 August 2020. Retrieved from https://blog.ventivtech.com/blog/a-brief-summary-of-the-long-history-of-risk-management.

SFS-ISO/IEC27000:2011. 2011. *Tietoturvallisuuden hallintajärjestelmä [Information system management system]*. Helsinki: Suomen Standardoimisliitto SFS [Helsinki: Finnish Standartisation Association].

SFS-ISO/IEC27005:2011. *Informaatioteknologia. Turvallisuus. Tietoturvariskien hallinta [Information technology. Security techniques. Information security risk management]*. 2011. Helsinki: Suomen Standardoimisliitto SFS [Helsinki: Finnish Standartisation Association].

Stevens, S. 1946. *On the Theory of Scales of Measurement*. Science, New Series, 103 (2684), 677-680.

Sweeting, P. 2011. *Financial Enterprise Risk Management*, Cambridge University Press.

Tapaustutkimus [Case study]. 2015. Article on University of Jyväskylä website. Accessed 24 August 2020. Retrieved from https://koppa.jyu.fi/avoimet/hum/menetelmapolkuja/menetelmapolku/tutkimusstrategiat/tapaustutkimus.

Tetlock, P. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton university Press.

Townsend, C. N.d. *A Brief and Incomplete History of Cybersecurity*. Article on United States Cybersecurity Magazine website. Accessed 23 August 2020. Retrieved from https://www.uscybersecurity.net/history/.

Tsai, C., Klayman, J., & Hastie, R. 2008. *Effects of Amount of Information on Judgment Accuracy and Confidence*. Organizational Behaviour and Human Decision Processes. 107(2), 97– 105

Virtue, T., & Rainey, J. 2014. *HCISPP Study Guide*, Elsevier Science & Technology Books.

Voreacos, D., Chiglinsky, K. & Griffin, R. 2020. *Asymmetric Warfare*. Bloomberg Markets, 28(6), 74-77.

Walker, K., Evans, J. & Macintosh, D. 2003. *Use of expert judgment in exposure assessment: Part 2. Calibration of expert judgments about personal exposures to benzene*, Journal of Exposure Science & environmental Epidemiology, 13, 1-16

Walker, K., Evans, J. & Macintosh, D. 2003. *Use of expert judgment in exposure assessment. Part I. Characterization of personal exposure to benzene*, Journal of Exposure Science & environmental Epidemiology, 13, 1-16

Wang, J., Neil, M., Fenton, & N. E. 2019. *A Bayesian network approach for cybersecurity risk assessment implementing and extending the FAIR model*. Computers & Security, 89.

Wheeler, E. & Swick, K. 2011. *Security risk management: building an information security risk management program from the ground up*. Syngress.

Whelan, S. 2002. *Actuaries' contributions to financial economics*. The Actuary, 12, 34-35.

William M., & Meehl, P. 1996. *Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy*. Psychology, Public Policy, and Law. 2, 293– 323

Wilson, R. 2019. *Developing a Quantitative Framework Tool to Implement Information Security Risk Management*. Master's thesis. College of Technology, University of Houston, Information and Logistics Technology, Cybersecurity. Accessed 25 August 2020. Retrieved from https://uh-ir.tdl.org/bitstream/handle/10657/5618/WILSON-THESIS-2019.pdf.

Wolke, T. 2017. *Risk Management*, Walter de Gruyter GmbH.

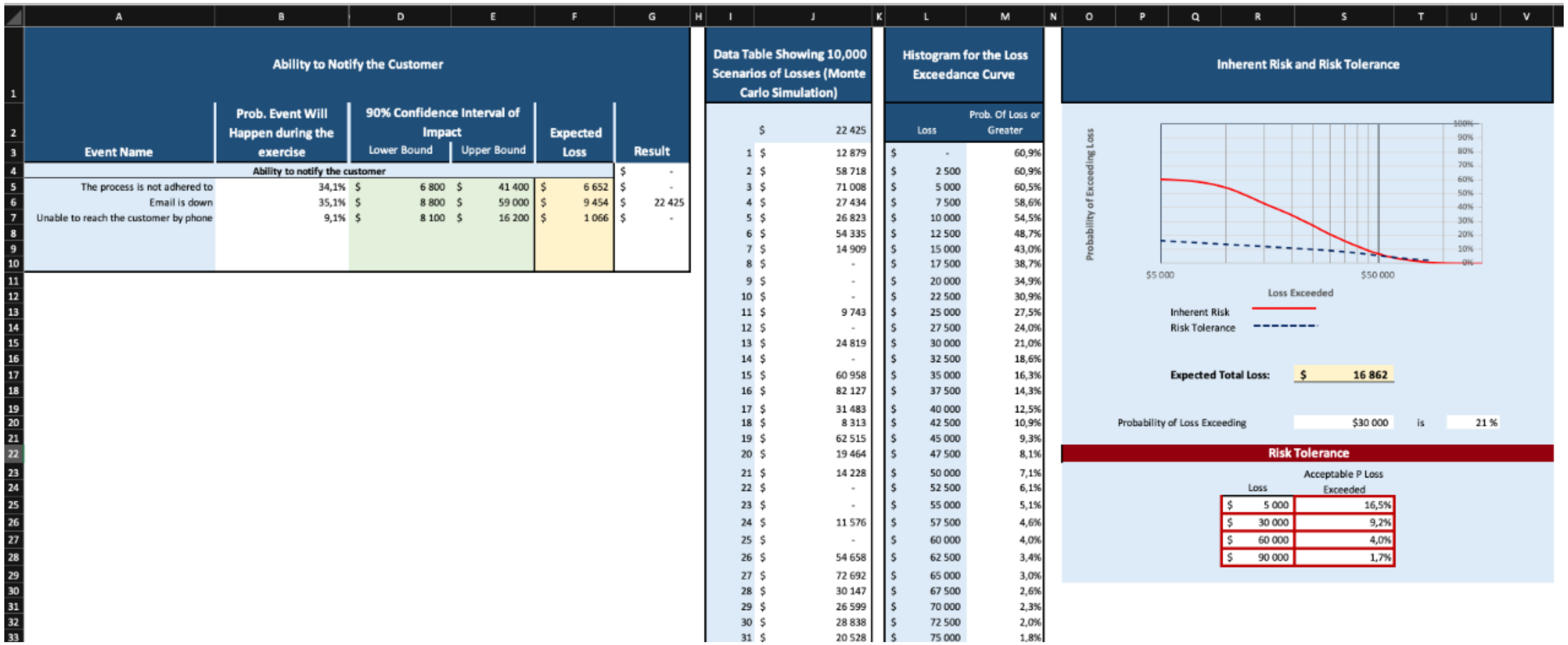
Ye, X., Zhao, J., Zhang, Y., & Wen, F. 2015. *Quantitative Vulnerability Assessment of Cyber Security for Distribution Automation Systems*, Energies. 8 (6) 5266-5286.

Mo, S. & Beling, P N.d. *Quantitative Assessment of Cyber Security Risk using Bayesian Network-based Model*. Accessed 25 August 2020. Retrieved from https://www.researchgate.net/publication/251890108_Quantitative_Assessment_of_Cyber_Security_Risk_using_Bayesian_Network-based_model.

Yrityksiin Kohdistuvat Kyberuhat 2019 [Corporation-Targeted Cyber Threats 2019]. 2020. Helsinki: Kauppakamari [Helsinki: Chamber of Commerce].

Appendices

Appendix 1. Example of a sub portfolio



Event Name	Prob. Event Will Happen during the exercise	90% Confidence Interval of Impact		Expected Loss	Result
		Lower Bound	Upper Bound		
Ability to notify the customer					$ -
The process is not adhered to	34,1%	$ 6 800	$ 41 400	$ 6 652	$ -
Email is down	35,1%	$ 8 800	$ 59 000	$ 9 454	$ 22 425
Unable to reach the customer by phone	9,1%	$ 8 100	$ 16 200	$ 1 066	$ -

	Data Table Showing 10,000 Scenarios of Losses (Monte Carlo Simulation)	Histogram for the Loss Exceedance Curve	
	$ 22 425	Loss	Prob. Of Loss or Greater
1	$ 12 879	$ -	60,9%
2	$ 58 718	$ 2 500	60,9%
3	$ 71 008	$ 5 000	60,5%
4	$ 27 434	$ 7 500	58,6%
5	$ 26 823	$ 10 000	54,5%
6	$ 54 335	$ 12 500	48,7%
7	$ 14 909	$ 15 000	43,0%
8	$ -	$ 17 500	38,7%
9	$ -	$ 20 000	34,9%
10	$ -	$ 22 500	30,9%
11	$ 9 743	$ 25 000	27,5%
12	$ -	$ 27 500	24,0%
13	$ 24 819	$ 30 000	21,0%
14	$ -	$ 32 500	18,6%
15	$ 60 958	$ 35 000	16,3%
16	$ 82 127	$ 37 500	14,3%
17	$ 31 483	$ 40 000	12,5%
18	$ 8 313	$ 42 500	10,9%
19	$ 62 515	$ 45 000	9,3%
20	$ 19 464	$ 47 500	8,1%
21	$ 14 228	$ 50 000	7,1%
22	$ -	$ 52 500	6,1%
23	$ -	$ 55 000	5,1%
24	$ 11 576	$ 57 500	4,6%
25	$ -	$ 60 000	4,0%
26	$ 54 658	$ 62 500	3,4%
27	$ 72 692	$ 65 000	3,0%
28	$ 30 147	$ 67 500	2,6%
29	$ 26 599	$ 70 000	2,3%
30	$ 28 838	$ 72 500	2,0%
31	$ 20 528	$ 75 000	1,8%

Inherent Risk and Risk Tolerance

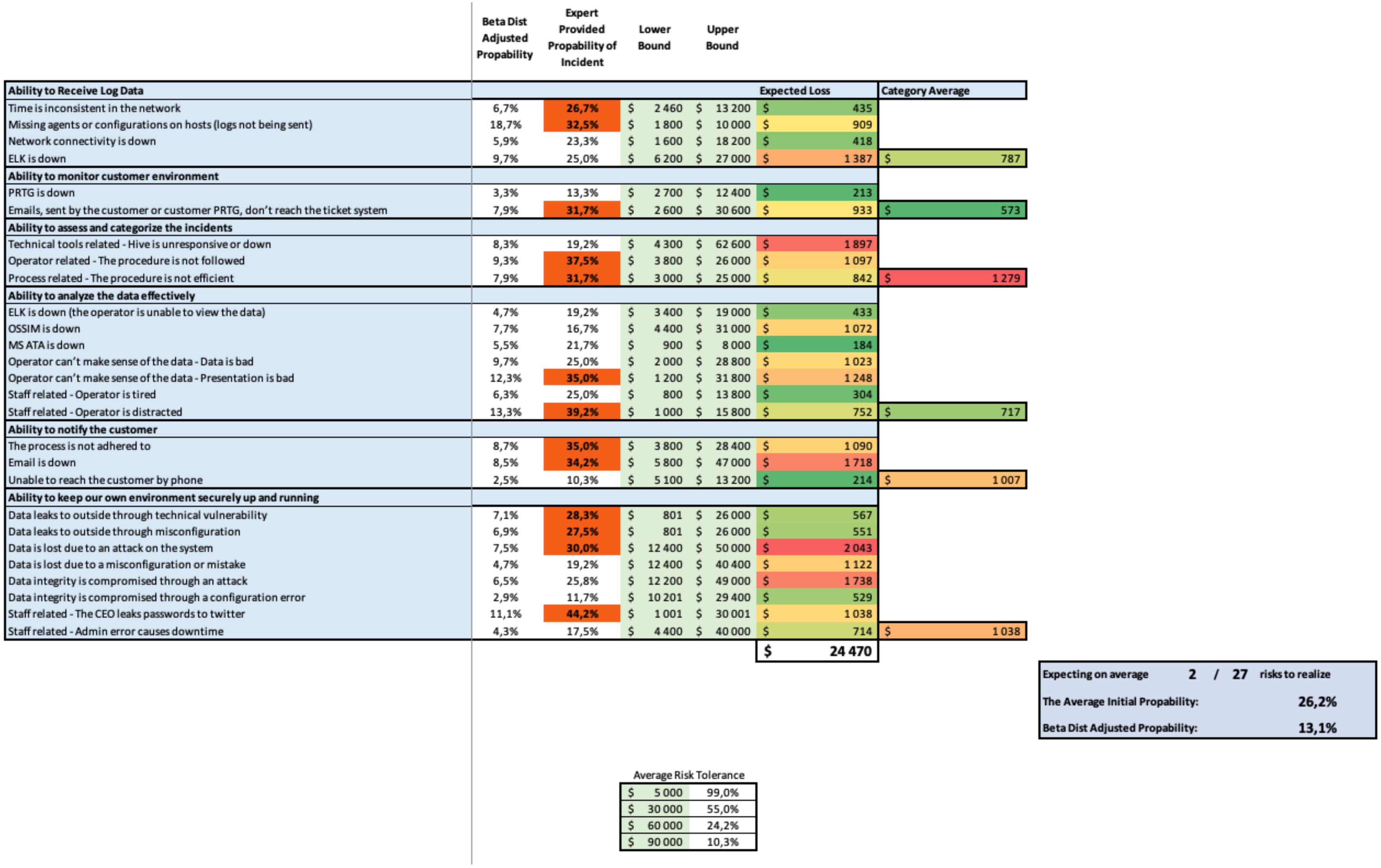
Expected Total Loss: **$ 16 862**

Probability of Loss Exceeding $30 000 is 21 %

Risk Tolerance

Loss	Acceptable P Loss Exceeded
$ 5 000	16,5%
$ 30 000	9,2%
$ 60 000	4,0%
$ 90 000	1,7%

Appendix 2. The summary view to the combined risk portfolio after the last exercise run



	Beta Dist Adjusted Propability	Expert Provided Propability of Incident	Lower Bound	Upper Bound	Expected Loss	Category Average
Ability to Receive Log Data						
Time is inconsistent in the network	6,7%	**26,7%**	$ 2 460	$ 13 200	$ 435	
Missing agents or configurations on hosts (logs not being sent)	18,7%	**32,5%**	$ 1 800	$ 10 000	$ 909	
Network connectivity is down	5,9%	23,3%	$ 1 600	$ 18 200	$ 418	
ELK is down	9,7%	25,0%	$ 6 200	$ 27 000	$ 1 387	$ 787
Ability to monitor customer environment						
PRTG is down	3,3%	13,3%	$ 2 700	$ 12 400	$ 213	
Emails, sent by the customer or customer PRTG, don't reach the ticket system	7,9%	**31,7%**	$ 2 600	$ 30 600	$ 933	$ 573
Ability to assess and categorize the incidents						
Technical tools related - Hive is unresponsive or down	8,3%	19,2%	$ 4 300	$ 62 600	$ 1 897	
Operator related - The procedure is not followed	9,3%	**37,5%**	$ 3 800	$ 26 000	$ 1 097	
Process related - The procedure is not efficient	7,9%	**31,7%**	$ 3 000	$ 25 000	$ 842	$ 1 279
Ability to analyze the data effectively						
ELK is down (the operator is unable to view the data)	4,7%	19,2%	$ 3 400	$ 19 000	$ 433	
OSSIM is down	7,7%	16,7%	$ 4 400	$ 31 000	$ 1 072	
MS ATA is down	5,5%	21,7%	$ 900	$ 8 000	$ 184	
Operator can't make sense of the data - Data is bad	9,7%	25,0%	$ 2 000	$ 28 800	$ 1 023	
Operator can't make sense of the data - Presentation is bad	12,3%	**35,0%**	$ 1 200	$ 31 800	$ 1 248	
Staff related - Operator is tired	6,3%	25,0%	$ 800	$ 13 800	$ 304	
Staff related - Operator is distracted	13,3%	**39,2%**	$ 1 000	$ 15 800	$ 752	$ 717
Ability to notify the customer						
The process is not adhered to	8,7%	**35,0%**	$ 3 800	$ 28 400	$ 1 090	
Email is down	8,5%	**34,2%**	$ 5 800	$ 47 000	$ 1 718	
Unable to reach the customer by phone	2,5%	10,3%	$ 5 100	$ 13 200	$ 214	$ 1 007
Ability to keep our own environment securely up and running						
Data leaks to outside through technical vulnerability	7,1%	**28,3%**	$ 801	$ 26 000	$ 567	
Data leaks to outside through misconfiguration	6,9%	**27,5%**	$ 801	$ 26 000	$ 551	
Data is lost due to an attack on the system	7,5%	**30,0%**	$ 12 400	$ 50 000	$ 2 043	
Data is lost due to a misconfiguration or mistake	4,7%	19,2%	$ 12 400	$ 40 400	$ 1 122	
Data integrity is compromised through an attack	6,5%	25,8%	$ 12 200	$ 49 000	$ 1 738	
Data integrity is compromised through a configuration error	2,9%	11,7%	$ 10 201	$ 29 400	$ 529	
Staff related - The CEO leaks passwords to twitter	11,1%	**44,2%**	$ 1 001	$ 30 001	$ 1 038	
Staff related - Admin error causes downtime	4,3%	17,5%	$ 4 400	$ 40 000	$ 714	$ 1 038
					$ 24 470	

Expecting on average	**2**	**/**	**27**	**risks to realize**
The Average Initial Propability:			**26,2%**	
Beta Dist Adjusted Propability:			**13,1%**	

Average Risk Tolerance

$ 5 000	99,0%
$ 30 000	55,0%
$ 60 000	24,2%
$ 90 000	10,3%