Universitetet i Stavanger

DET SAMFUNNSVITENSKAPELIGE FAKULTET,
HANDELSHØGSKOLEN VED UIS

MASTEROPPGAVE

STUDIEPROGRAM:	OPPGAVEN ER SKREVET INNEN FØLGENDE SPESIALISERINGSRETNING:
Økonomisk-administrative-fag, masterstudium	**Risikostyring** ER OPPGAVEN KONFIDENSIELL? **NEI** (**NB**! Bruk rødt skjema ved konfidensiell oppgave)

TITTEL:

Bayesiansk Nettverksmodellering for Analyse av Datainnbrudd i Bank

ENGELSK TITTEL:

Bayesian Network Modeling for Analysis of Data Breach in a Bank

FORFATTER(E)		VEILEDER:
Studentnummer:	Navn:	**David Häger**
208224 ……	**Vasily Apukhtin** ……	
……	……	

OPPGAVEN ER MOTTATT I TO – 2 – INNBUNDNE EKSEMPLARER

Stavanger, ……/…… 2011 Underskrift administrasjon:……

Bayesian Network Modeling for Analysis of Data Breach in a Bank

Abstract

Nowadays it is impossible to imagine a modern financial institution which operation does not depend upon information technology (IT). There has been a huge emphasis on data security recently. Financial institutions are the primary targets for different kind of abuse because they possess a lot of sensitive information that can easily be converted into money. That is why banks and other financial institutions are trying to protect themselves and their clients from different kinds of malicious activity. Data security is one of the most important aspects of everyday banking and an important part of a sound operational risk management (ORM). In modern world ORM lends itself well to IT and while it is difficult for companies, and in some cases impossible, to control external events, it is feasible to manage people, systems and processes in order to prevent or reduce operational losses.

Most of information today is stored and transferred electronically that makes it more exposed to breaches. Leak of financial information about customers or classified business information about, for example, future investments might have negative effect on organization that was not able to protect it. That is why this type of exposure needs to be accounted for in the operational risk management system. Mitigating that risk through effective security controls can help in both lowering the probability of loss and decreasing the institution's capital requirements.

Financial institutions use huge amounts of money and other resources to protect sensitive information. But in spite of huge investment into security, data breaches continue to occur. Financial institutions experience data breaches caused either by their own employees or external attackers. The “insider threat” or “insider problem” has received considerable attention and is cited as the most serious security problem in many studies. It is also considered the most difficult problem to deal with, because an “insider” has information and capabilities not known to other, external attackers.

Companies do not have much internal information about data breaches. Information that is available from external sources is often not easy to analyze due to the variety of scenarios and/or incompleteness of cases. The companies face a challenge to develop an approach that draws upon information coming from different sources. The flexible modeling framework provided by Bayesian Networks (BN) makes it an appropriate candidate for modeling this challenging issue. In addition, BNs ability to represent complex interrelationships among entities and its mathematically sound interface can make it the best match to create a model for quantitative analysis of sensitive data breach.

Bayesian Network Modeling for Analysis of Data Breach in a Bank

Table of Contents

Tables

Figures

Abbreviations

Basel Capital Accord (BCA)

Basel Committee on Banking Supervision (BCBS)

Bayesian Network (BN)

Conditional Probability Distribution (CPD)

Conditional Probability Table (CPT)

Data Breach Investigation Report (DBIR)

Data Loss Prevention (DLP)

Directed Acyclic Graph (DAG)

Graphical model (GM)

Information Technology (IT)

Joint Probability Distribution (JPD)

Lincoln Financial Advisors Corporation (LFA)

Lincoln Financial Securities, Inc. (LFS)

Operational Risk Management (ORM)

Operational Risk (OR)

Probability of data breach (PDB)

Sony Online Entertainment (SOE)

Structured Query Language (SQL)

United States Secret Service (USSS)

Preface

This master thesis represents the final work of my Master's degree in Economic and Business Administration at the University of Stavanger. The process of completing this thesis has been interesting and educational, but also challenging and time-consuming. This assignment gave me the opportunity to expand my knowledge on several subjects of interest like, for example, operational risk, data security and practical use of Bayesian Network (BN) software.

AgenaRisk was used to develop the BN. This state-of-art software has built-in function for sensitivity analysis that was used to validate the developed model. The graphical representation of the designed BN and results of the sensitivity analysis can both be found further in the text. The CD included with this thesis contains the model file, as well as this report in PDF-format.

I am very thankful to David Häger, for his assistance in writing of this thesis. He has contributed with both relevant information and valuable guidance throughout this semester. He has been very supportive and available during the entire process.

I would also like to thank Lasse Berg Andersen for taking special care of the students who have chosen specialization in Risk Management.

Bayesian Network Modeling for Analysis of Data Breach in a Bank

Introduction

For a long time credit risk and market risk had been considered the two largest contributors to banks' risks and operational risk (OR) had been regarded just as a part of other risks. Nevertheless operational losses are not new to banks. They occur every day and can affect soundness and operating efficiency of all banking activities and business units. Operational losses can be internally inflicted or can result from external sources.

Modern companies are focused on ensuring of business performance and at the same time protecting investors and corporate brands. Due to this fact the executives are being prompted to re-prioritize the importance of the ORM within their organizations. Conditions of the modern world make top management and boards of directors expect a deeper understanding of how OR is being managed. Globalization that led to increase in transaction volumes and stronger reliance on IT have introduced higher degrees of complexity and uncertainty to organizations. In order to be competitive and improve performance, many organizations are trying to understand and proactively manage the risks that can influence their business. Operational risks exist as soon as a company uses employees and can emerge long before credit or market risks come into light. Therefore it is vital to have a sound ORM in place. Of course risk management and particularly ORM is not a solution in itself, but one of the most important instruments that should be genially brought into all strategic and operational decisions.

Operational risk was brought to the light in January 2001 while banks were still focusing on credit risk. Financial institutions were able to understand how to set policies for credit risk management requirements, but not operational risk. Moreover, banking industry is still trying to understand how to deal with operational risk. Globalization and IT made data security one of the most important aspects of the sound ORM in the modern world. Companies use huge amounts of money and other resources to protect sensitive information. However data breaches continue to occur within all types of organizations and most of them happen in companies that provide financial services. There have been a lot of scandals in media regarding losses of sensitive information. These losses especially when covered in media, have negative impact on the breached companies and should be avoided.

The data breach is an operational issue that can lead to loosing reputation among customers and business partners and in extreme cases can even put organization existence at risk. Lacking precautions in information security constitute a significant operational risk. Nevertheless every kind

of risk, even so unpredictable and complicated like operational, can be significantly reduced by implementing and constant improving of solid risk management.

Operational risk is a very comprehensive issue. However different researchers claim that BNs are potentially powerful tools to deal with this kind of risk. This puts a question whether the BN approach is applicable to the problem of data breach. Thereafter the objectives of this master thesis are defined as following:

- Provide insight into the problem of data security in the ORM perspective
- Get deeper understanding of the nature of data breaches
- Define a specific issue for BN modeling
- Find out whether BN is an appropriate candidate for quantitative analysis of the chosen issue
- Map most relevant factors that can be used in the developing of BN
- Develop, validate and test the model
- Draw conclusions

Bayesian Network Modeling for Analysis of Data Breach in a Bank

1. Globalization and Information technology

There has been a great deal of discussion in the recent years about globalization which can be defined as a process of interaction and integration among the people, companies, and governments of different nations. This process is driven by international trade and aided by information technology. Globalization has effects on the environment, on culture, on political systems and on economic development and prosperity.

Financial globalization has caused higher degree of competition between financial institutions and has given customers and investors choices and opportunities they did not have before. This has resulted in the development of new financial products, instruments and services. They has coincided or maybe even triggered a number of technological innovations including the development of the Internet, leading to revolutionized banking activities such as online banking, growth of e-commerce, and e-mail services. On one hand, these innovations helped to increase considerably the speed of information flow and the amount of information itself. But on the other hand they resulted in an elevated exposure of the financial institutions to various sources of risk. Increased use of computer based-banking services leads to several vulnerabilities like for example viruses, computer failures and credit card frauds. Furthermore, previously nonexistent or insignificant risk factors have become a large (or larger) part of the complex risk profiles of financial institutions.

The technology is perhaps the most visible aspect of globalization and in many ways its driving force. Communication technology has revolutionized information systems all around the world. Emergence of information technology (IT) made globalization perceptible and observable in nearly every aspect of our lives. IT integrates people all over the world through the common platform that helps people to communicate and share information despite the distances. This new technology named Internet is available to nearly everyone and its impact both positive and negative sides. One of the positive effects of IT on globalization is the modernization and improvement of the business sector throughout the world. IT made business more competitive and productive by allowing instant access to information and by providing efficient electronic transaction. The modern market has due to IT become more competitive and as the result consumers and business itself have greater choices. The internet has impact on the services and products that are bought, sold and delivered, altering relations between customers, companies, and employees and therefore speeding the globalization.

Nowadays it is impossible to imagine a modern financial institution which operation does not depend upon IT. Most of the banks provide services like online banking and some banks do not even

have traditional offices because they operate through internet and telephone. Banking today does not have geographical restrictions like some decades ago. People can live in Australia and have accounts in USA or nearly any other country in the world. All this sounds very optimistic and in an ideal world it could have been a perfect arena for further globalization and global prosperity. But, unfortunately, we do not live in a perfect world and IT is used not only for the benefit of the mankind. There are a lot of people worldwide who abuse these new opportunities.

There has been a huge emphasis on data security recently. Financial institutions are the primary targets for different kind of abuse because they possess a lot of sensitive information that easily can be converted into money. That is why banks and other financial institutions are trying to protect themselves and their clients from different kinds of malicious activity. Data security is one of the most important aspects of everyday banking. So what is data security? Data security, in simple terms, is the practice of keeping data protected from corruption and unauthorized access. The focus behind data security is to ensure privacy while protecting personal and corporate data (SpamLaws, 2009). Data can be defined as raw form of information (from personal files and intellectual property to market analytics and details intended to top secret) stored in databases, network servers and personal computers.

Although everybody understands the importance of data loss prevention (DLP), it is not paid sufficient attention in some organizations to it. There is a lot of DLP software from different vendors available on the marked but nevertheless some IT managers are not comfortable with deploying DLP, because it requires admitting to an internal weakness and confessing to not doing their job (Israeli Software, 2011). Even some CEO's are against implementation of DLP solutions as it implies employee monitoring not to mention some countries, like Germany, where it is forbidden by law to monitor employees (Israeli Software, 2011). However most organizations have a number of information controls. But without proper management, the controls can be somewhat disorganized and disjointed, having been implemented often as point solutions to specific situations or simply as a matter of convention (Altal Security, 2005). The security controls in operation usually cover certain aspects of IT and data security, specifically, leaving non-IT information assets (such as paperwork and proprietary knowledge) less well protected on the whole (Altal Security, 2005).

Both security and network operations are an important part of operational risk management (ORM) where operational risk is defined by Basel Committee on Banking Supervision (BCBS) as "the risk of loss resulting from inadequate or failed internal processes, people and systems, or from external

events". In modern world ORM lends itself well to IT and while it is difficult and in some cases impossible for companies to control external events, it is feasible to manage people, systems and processes in order to prevent or reduce operational losses. The identification and management of operational risk is a real and live issue for modern day banking, especially since the decision by BCBS to introduce a capital charge for this type of risk as a part of the capital adequacy framework known as Basel II.

2. Operational Risk in BCBS's perspective

2.1 The Basel Committee on Banking Supervision

The BCBS plays the leading role in the financial risk regulation network, setting risk management regulations to financial institutions worldwide. BCBS is the key player in establishing risk assessment and management guidelines for banks. The Committee does not have any supranational authority with respect to banking supervision. Its recommendations do not have legal force, because it is up to the national authorities to decide whether to implement them or not (Moosa, 2008).

2.2 Basel I

The first Basel Capital Accord (BCA) usually referred to as Basel I, was released in 1988. Basel I was revolutionary due to fact that it wanted to develop a single risk-adjusted capital standard that would be applied throughout the major banking countries in the world (Allen, 2003). The goal was the adaptation of best practices by banks all over the world, thereby enhancing the efficiency, productivity and soundness of the global financial system (Allen, 2003). The idea of Basel I was generally positive but it has also been criticized as having significant shortcomings. One of the most significant shortcomings of Basel I is that it does not take operational risk into consideration. This sounds odd nowadays because operational risk is considered to be an important part of sound risk management in modern financial institutions.

2.3 Basel II

In response to the criticism of the Basel I Accord and in order to address changes in the banking environment that the 1988 Accord could not deal with effectively, the BCBS decided to create a new capital Accord, Basel II (Moosa, 2008). The accord was intended to deal with market innovations and a fundamental shift towards more complexity in the banking industry. Another objective was to narrow the gap between regulatory and economic capital.

One of the major differences between Basel I and Basel II is that the latter has more focus on operational risk. Operational risk was brought to the light in January 2001 while banks were still focusing on credit risk. Financial institutions were able to understand how to set policies for credit risk management requirements, but not operational risk. Moreover, banking industry is still trying to understand how to deal with operational risk.

Bayesian Network Modeling for Analysis of Data Breach in a Bank

The original Basel accords stated financial institutions could maintain a fixed percentage of what they loaned to parties as capital for a certain amount of failures and mitigate both credit and operational risk (Mackey , 2008). However, development of financial sector resulted in higher complexity of financial instruments. In addition, the operational losses that had catastrophic consequences for some banks made the Basel Committee think about more sophisticated methods for measuring and managing risk. All this resulted in recommendation of advanced measurement models that deal with credit and operational risk separately.

Separated treatment of operational risk makes information security an important part of Basel II. Financial institutions need to establish a risk measurement, management, and reporting system that demonstrate to regulators the effectiveness of their risk management approach. Banks adopting more sophisticated approaches to risk modeling can benefit financially by reducing the amount of capital that needs to be set aside to mitigate risk (Mackey , 2008). This can lead to real business benefits in the global and technologically advanced financial sector.

The potential financial benefit comes from the risk management system itself. In this new model, the very systems used in managing risk become critical IT resources (Mackey , 2008). An attack on them, or even a failure due to human error, could blemish bank's reputation, lead to financial losses or in extreme cases even bring the organization existence into question. Consequently, financial organizations need to have adequate policies and mechanisms to ensure that these systems and the processes surrounding them are well under control. Security components as identity management, access control, application administration and monitoring are vital in order to protect these systems from malicious activity

As an institution's IT resources are a tidbit for many malefactors it is very important to analyze their contribution to operational risk. Most of information today is stored and transferred electronically that makes it more exposed to breaches. Leak of financial information about customers or classified business information about, for example, future investments might have a negative effect on an organization that was not able to protect it. That is why this type of exposure needs to be accounted for in the risk management system. Mitigating that risk through effective security controls can help a bank in both lowering the probability of loss and decreasing the institution's capital requirements (Mackey , 2008).

2.4 Basel III

The goal of the Basel III is to strengthen the resilience of banks and the global banking system. The BCBS's reforms seek to improve the banking sector's ability to absorb shocks arising from financial and economic stress, whatever the source, thus reducing the risk of spill over from the financial sector to the real economy (Basel Committee on Banking Supervision, 2010).

By introducing Basel III, BCBS is trying to strengthen the banking sector in order to make sure that such a disaster as financial crisis of 2008 will never happen again. While putting a lot of effort into regulation of credit risk in Basel III, the Committee does not pay much attention to operational risk. Implementing of stricter regulatory capital requirements for banks will not prevent a new crisis but can only help to absorb the shocks arising from it. Operational risk, in contrast to other risks, envelops every activity and every employee in a company and therefore can be regarded as a precondition for the financial crisis (Thirlwell, 2010).

3. Operational Risk Management

3.1 Overview

For a long time credit risk and market risk have been considered the two largest contributors to banks' risks and operational risk has been regarded just as a residual risk. Nevertheless operational losses are not new to banks. They occur every day and can affect soundness and operating efficiency of all banking activities and business units. For example, abandonment of sensible credit risk management before the disaster of 2008 can be seen as an operational issue (Thirlwell, 2010).

Fundamentally, the crisis was a "failure of risk management" or rather a "failure to apply risk management at all levels" (Thirlwell, 2010). Risk management failure is mainly about people risk, which in turn lies at the heart of the operational risk. Experience made in the 15 years before the recent financial crisis show that operational risks are major source of losses in the banking sector (Oesterreichische Nationalbank, 2006).

As it was mentioned before, BIS defines operational risk as "the risk of loss resulting from inadequate or failed internal processes, people or systems, or from external events". Most of these losses are relatively small in magnitude. Examples of such operational losses include losses resulting from accidental accounting errors, minor credit card fraud, or equipment failures. The fact that these losses are frequent makes them predictable and often preventable (Chernobai & Rachev, 2007).

Operational losses can be internally inflicted or can result from external sources. Internally inflicted sources include most of the losses caused by human, process, and technology failures, such as those due to human errors, internal fraud, unauthorized trading, injuries, business delays due to computer failures or telecommunication problems (Chernobai & Rachev, 2007). External sources include man-made incidents such as external fraud, theft, computer hacking, terrorist activities, and natural disasters (Chernobai & Rachev, 2007). Some of the internal losses can be prevented by, for example, appropriate control techniques and/or management of personnel. It is possible to reduce the number of external losses by implementing of complex computer driven security systems. But a complex security system is not a solution per se because there is always a possibility to bypass the system either from outside or in many cases more easily from inside of the organization. Of course if somebody wants to bypass security mechanisms, he or she can do it but a well-protected target is

less attractive. Although it is impossible to eliminate operational losses, it is feasible to reduce them with the help of sound the ORM.

3.2 Importance of operational risk management

The field of risk management (RM) has its origins in the insurance industry. In the 1980s, risk management in manufacturing firms took hold with the adoption of total quality management. It was not until the 1990s that the field of risk management received greater recognition for its importance in many companies, especially those providing financial services (Chernobai & Rachev, 2007).

Modern companies are focused on ensuring of business performance and at the same time protecting investors and corporate brands. Due to this fact the executives are being prompted to re-prioritize the importance of the ORM within their organizations. Conditions of the modern world make top management and boards of directors expect a deeper understanding of how OR is being managed. Globalization that led to increase in transaction volumes and stronger reliance on IT have introduced higher degrees of complexity and uncertainty to organizations. In order to be competitive and improve performance, many organizations are trying to understand and proactively manage the risks that can influence their business.

Most organizations and, of course, financial institutions are undergoing a lot of changes nowadays. This has increased the probability of failure and mistakes from operations points of view – resulting in increasing focus on managing of OR. Some operational risk losses that have been widely discussed in media in recent years had catastrophic consequences like in case of Barings Bank and Société Générale. The regulators of finance sector are demanding greater understanding by directors of the risks they manage, and the quality of the controls that are being used to reduce or mitigate these risks. This resulted in the stronger emphasis on the importance of having a sound ORM practice in place. This makes ORM one of the most complex and fastest growing risk disciplines in financial institutions (MetricStream, 2011).

It is possible to define main drivers for this development. First of all, financial institutions acknowledge that a consistent and effective ORM framework can help them to achieve better performance. For example, by including effective ORM in certain activities a bank can help ensure that risks associated with those activities are understood and addressed. The second driver is the launch of Basel II that affected most financial services worldwide. And finally, banking failures reinforced by shareholders and regulatory pressures have made it mandatory for financial

institutions to implement systematic and continuous processes for collecting, analyzing and reporting operational risks (Chartis Research, 2010) Operational risks exist as soon as a company uses employees and can emerge long before credit or market risks come into light. (Oesterreichische Nationalbank, 2006). Therefore it is vital to have a sound ORM in place.

Of course risk management and particularly ORM is not a solution per se, but one of the most important instruments that should be genially brought into all strategic and operational decisions. According to Basel II definition, employees are one of the operational risk sources in an organization and that is why they should be aware of essential objectives and components of ORM implemented in the organization (Oesterreichische Nationalbank, 2006). In addition to awareness of employees it is important that top management has positive attitude to ORM and allocates appropriate budget funds and human resources making it possible to establish or/and maintain a good risk culture within organization.

3.3 Operational risk management and information technology

Globalization and development of IT has made information the most important asset for many companies and especially for such advanced organizations as financial institutions. Business units have become more dependent on information and real-time access to it, which makes securing of this vital asset a prerogative for many banks. With the openness of modern business comes the reality of information exposure at risk. Strong information security is necessary for the challenges modern banks face today. According to different sources like, for example, Verizon Data Breach Investigation Report (DBIR) 2010, a great share of data breaches involved staff gaining access to information. It is not surprising because data is stored in many different locations like hard drives, servers, portable devices etc. and employees need access to them in order to do their jobs. In addition, regular moving and replication of data makes security of information a real challenge.

The data breach is an operational issue that can lead to loss of reputation among customers and business partners and in extreme cases can even put organization existence at risk. Lacking precautions in information security constitute a significant operational risk. Nevertheless every kind of risk, even so unpredictable and complicated like operational, can be significantly reduced by implementing and constant improving of solid risk management.

Companies do not have much internal information about data breaches. Information that is available from external sources is often not easy to analyze due to the variety of scenarios and/or

incompleteness of cases. The situation is worsened by the fact that many companies try to hide facts of compromised data from the media or/and regulation bodies because data breaches can tarnish companies' reputation.

The most recent example of this is the Sony case that can get into the Top 10 list of all-time biggest breaches. More than 12,700 customers' credit card numbers may have been stolen. Sony Online Entertainment (SOE) believes hackers stole customer information on April 16 and April 17. Engineers and security consultants reviewing SOE systems discovered that personal information from approximately 24.6 million SOE accounts may have been stolen, as well as certain information from an outdated database from 2007. The outdated database had approximately 12,700 non-U.S. credit or debit card numbers and expiration dates. There may also have been 10,700 direct debit records stolen from customers in Austria, Germany, Netherlands, and Spain. Consumers were notified by email only on April 26, when Sony also notified the New Hampshire Attorney General's Office. The most recent news shows that Sony did not appear at the hearing to testify because the company did not have time to prepare for the hearing, in spite of the fact that the data breach occurred more than a month ago (Office of Inadequate Security, 2011). Moreover, the company did not notify the FBI until two days after they detected the breach and did not meet with the FBI until 5 days after the breach.

The Sony example shows that companies, even in case of major breaches, are trying to avoid negative publicity and even official bodies. This very data breach seems to be the work of outsiders (investigation is not over) but nevertheless tarnished the image of the company greatly. Imagine the situation if the further investigation will conclude that this very data breach was caused by a SOE employee or with the help of someone who worked at SOE. This would impact the company's image even harder because how can the customers be loyal to SOE if the company is not even able to protect their information from inside? Companies are very afraid of such scenarios and that is why some of them do not publically announce data breaches caused by insiders at all. Companies are not doing this because they may be afraid of the negative publicity or increased liability that may arise from the incident. Or, they may believe that the harm suffered would not be sufficient to warrant criminal charges (Probst, Hunker, & Gollmann, 2010). All this results in wider information gaps regarding cases related to data breaches caused by insiders and makes it very difficult for organizations to develop a more comprehensive understanding of the insider threat.

This lack of information puts a question of how to deal with the insider threat. The possible solution is to use all the available information from all available sources, but this information can unfortunately not always be homogenous. The companies face a challenge to develop an approach that draws upon information coming from different sources and traditional (statistical) methods that are broadly used for measuring, for example, credit risk are not very useful for this objective. There should be a tool that can combine different types of data in order to develop an appropriate approach to deal with this very comprehensive problem. The flexible modeling framework provided by Bayesian Networks (BN) makes it an appropriate candidate for modeling this challenging issue. In addition, BNs ability to represent complex interrelationships among entities and its mathematically sound interface can make it the best match to create a model for quantitative analysis of sensitive data breach. Further I would like to dwell upon what BN are and give a simple example of how they can be used.

4. Bayesian Networks

4.1 Overview

Information today is abundant but at the same time is often inconsistent, contradictory and of uncertain traceability and reliability. The process of information interpreting is very important, because not interpreted information has little value. That is why it is not surprising that there has been shown heightened interest in recent years for statistical approaches in order to optimally use the information (Pourret, Naim, & Marcot, 2008).

Due to the constant information flow, people face a lot of problems in the modern world. These problems need to be solved and decisions are sometimes not easy to make especially those regarded complex problems. Human cognitive abilities, memory and reason are limited and that is why it can be difficult to understand and manage the reality (Pourret, Naim, & Marcot, 2008). Besides, biological limitations of human capabilities, a variety of factors, either cultural (education, ideology), psychological (emotions, instincts), and even physical (fatigue, stress) tend to distort our judgment of the situation (Pourret, Naim, & Marcot, 2008). One way of trying to better handle reality – in spite of these limitations and biases – is to use representations of reality called models.

BN belong to the family of probabilistic graphical models. These models are used to represent knowledge of an uncertain domain. BNs consist of nodes and arcs between the nodes where the nodes represent random variables and arcs – probabilistic dependencies among the corresponding random variables. The conditional dependencies between nodes can be estimated by using combination of statistical data with qualitative data. Hence, BNs combine principles from graph theory, probability theory, computer science, and statistics.

Graphical models (GMs) with undirected edges are generally called Markov random fields or Markov networks (Ben-Gal, 2008). These networks provide a simple definition of independence between any two distinct nodes based on the concept of Markov blanket and are popular in fields such as statistical physics and computer vision (Ben-Gal, 2008).

BNs correspond to another GM structure known as a directed acyclic graph (DAG). The variables together with the directed edges form a DAG. A directed graph is acyclic if there is no directed path $A1 \rightarrow \cdots \rightarrow An$ so that $A1 = An$ (Jensen & Nielsen, 2007). When talking about the relations in a directed graph, the wording of family relations is used: if there is link from A to B, that means that

B is a child of A, and A is a parent of B. BNs are both mathematically rigorous and intuitively understandable. They enable an effective representation and computation of the joint probability distribution (JPD) over a set of random variables (Ben-Gal, 2008).

A BN represents a simple conditional independence statement. It means that each variable is independent of its parents in the graph. This property helps to reduce the number of parameters needed to characterize the JPD of the variables. This reduction provides an efficient way to compute the posterior probabilities given the evidence (Ben-Gal, 2008).

Besides the graphical structure, the model consists of parameters that describe the relationships between nodes. These parameters are presented in accordance with the Markov property where the conditional probability distribution (CPD) of each node depends only on its parents. For discrete random variables, this conditional probability is often represented by a table, listing the local probability that a child node takes on each of the feasible values – for each combination of values of its parents. The joint distribution of a collection of variables can be determined uniquely by these local conditional probability tables (CPTs) (Ben-Gal, 2008).

4.2 Bayes theorem

The fundamental rule for probability calculus is presented by the following expression:

$$P(A \mid B)P(B) = P(A \cap B)$$

This rule describes how to find the probability of seeing both A and B when we know the probability of A given B and the probability of B.

By conditioning on another event C, the fundamental rule can also be written as:

$$P(A \mid B \cap C)P(B \mid C) = P(A \cap B \mid C)$$

Since P(A ∩ B) = P(B ∩ A) (and also P(A ∩ B | C) = P(B ∩ A | C)), we get that
P(A | B)P(B) = P(A ∩ B) = P(B | A)P(A) from the fundamental rule.
This yields the well-known Bayes' rule:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Bayes' rule provides us with a method for updating our beliefs about an event A given that we get information about another event B. For this reason P(A) is usually called the prior probability of A

, whereas P(A | B) is called the posterior probability of A given B ; the probability P(B | A) is called the likelihood of A given B (BNs and Decision Graphs).

$$P(A \mid B, C) = P(B \mid A, C)P(A \mid C)/P(B \mid C)$$

4.3 Example of a Bayesian Network

Let us imagine a following situation. It is a beautiful Monday morning and you are having a delicious breakfast. You decide to find out what the weather is like and look out of the window. Suddenly you begin to worry because the grass on the lawn before the house is wet and you favorite dog was out all the night. But there is no need to worry because you do not know for sure why the lawn is wet. There two possible causes for the wet grass: either it was raining, or the neighbor forgot to turn of the sprinkler last night. A BN model can help to find out what cause is more likely.

First of all we should graphically present the problem that should be solved. It can be done as following:

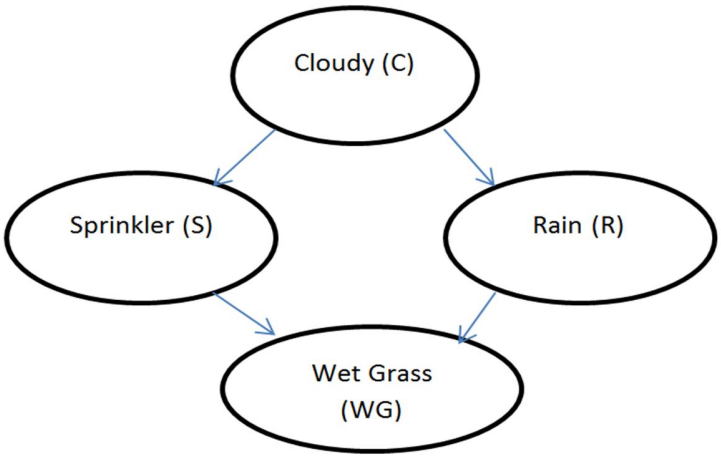

Figure 1 - Example of a Bayesian Network

We see the event (WG) has to possible causes: either the water sprinkler is on or it is raining. These two causes have a parent (C) while (WG) is a child of both (R) and (S). This conditional independence relationship encoded in the BN can be stated as follows: a node is independent of its parents, where the child/parent relationship is with respect to some fixed topological ordering of the nodes. By the chain rule of probability, the joint probability of all the nodes in the graph above is $P(C, S, R, WG) = P(C) \times P(S|C) \times P(R|C, S) \times P(WG|C, S, R)$.

By using conditional independence relationships, we can rewrite this as $P(C, S, R, WG) = P(C) \times P(S|C) \times P(R|C) \times P(WG|S, R)$ where we were allowed to simplify the third term because R is independent of S given its parent C, and the last term because WG is independent of C given its parents S and R.

The second step in solving a problem is specifying of the model parameters. For a directed model, we must specify the CPD at each node. If the variables are discrete, this can be represented as CPT, which lists the probability that the child node takes on each of its different values for each combination of values of its parents.

The CPT for each node in this example is presented as following:

Cloudy:

False	0.5
True	0.5

Sprinkler:

Cloudy	False	True
False	0.5	0.9
True	0.5	0.1

Rain:

Cloudy	False	True
False	0.8	0.2
True	0.2	0.8

Wet Grass:

Sprinkler	False		True	
Rain	False	True	False	True
False	1.0	0.1	0.1	0.01
True	0.0	0.9	0.9	0.99

When the model is graphically presented and parameters are specified we can try to find out why the grass is wet. As it was mentioned before, there are two possible reasons for this: either it is raining, or the sprinkler is on. Let us found what cause is more likely. We can use Bayes' rule to compute the posterior probability of each cause (where 0=false and 1 =true).

$$\Pr(S = 1|\mathrm{WG} = 1)\frac{\Pr(S = 1, WG = 1)}{\Pr(WG = 1)} = \frac{\sum_{c,r}\Pr(C = c, S = 1, R = r, WG = 1)}{\Pr(WG = 1)} = \frac{0.2781}{0.6471} = 0.430$$

$$\Pr(R = 1|\mathrm{WG} = 1)\frac{\Pr(R = 1, WG = 1)}{\Pr(WG = 1)} = \frac{\sum_{c,s}\Pr(C = c, S = s, R = 1, WG = 1)}{\Pr(WG = 1)} = \frac{0.4581}{0.6471} = 0.708$$

$\Pr(WG = 1) = \sum_{c,r,s} \Pr(C = c, S = s, R = r, WG = 1) = 0.6471$ is a normalizing constant, equal to the probability (likelihood) of the data. We can see that it is more likely that the grass is wet because it is raining: the likelihood ratio is 0.708/0.430 = 1.647.

By simply setting the scenario for "wet grass" as true we get the following result (which corresponds with the results above) in AgenaRisk:

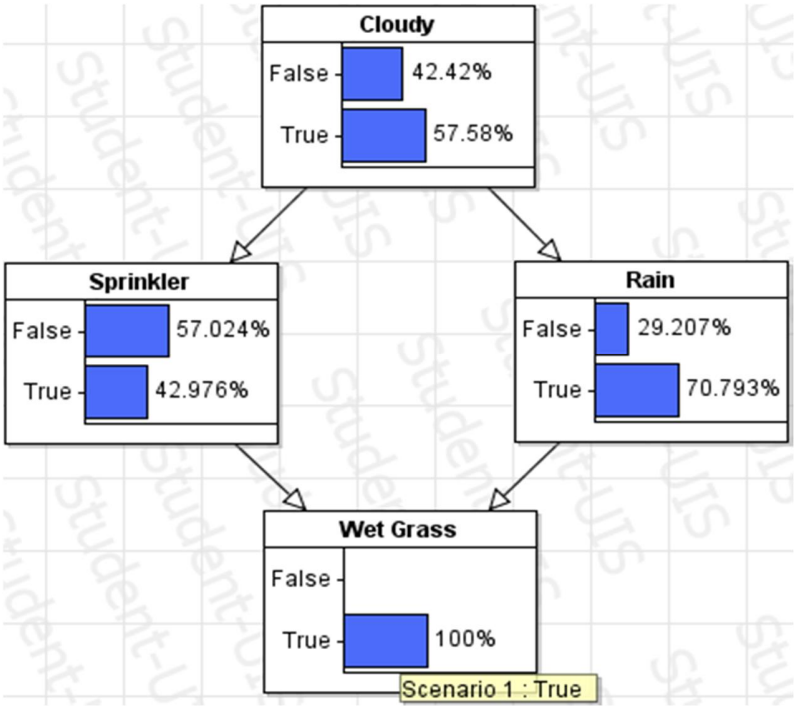

Figure 2 – Example of a Bayesian Network in AgenaRisk

Bayesian belief networks are technique for integrating qualitative data in the form of subjective beliefs and insecure knowledge into the quantitative modeling of operational risks. An advantage they offer is that they illustrate cause-effect chains that are of decisive importance for the management of operational risks. They can be applied to support scenario analyses where cause-effect relationships are important and the subjective evaluations of experts should be used due to the future-oriented nature of these analyses (Oesterreichische Nationalbank, 2006).

Bayesian inference has many great features, that is why it fits well for operational risk modeling. First of all it provides transparency for review by internal audit and/or regulators as both sources of information can be analyzed separately. Second of all its foundations rest on assumptions that fit well with operational risk, as both observations and parameters of the distributions are considered to be random (Gregoriou, 2009). And finally it provides a structural and sound statistical technique to combine two heterogeneous sources of information (subjective human opinions and objective collected data) that makes it natural candidate for quantitative analysis of data breach.

5. Nature of Data Breaches

5.1 Overview

As it was mentioned before, continuing technological innovation and competition among existing banking organizations and new entrants have allowed for a much wider array of banking products and services. However, the rapid development of new banking capabilities carries risks as well as benefits (Basel Committee on Banking Supervision, 2003).

This section is generally based on Verizon DBIRs. These reports, in the author's point of view, contain most complete information regarding data breaches that is freely available for public use. Verizon Business is the global IT solutions partner to business and government. It caters to large and medium business and government agencies and is connecting systems, machines, ideas and people around the world for altogether better outcomes. Verizon Business helps enterprises find solutions to their business issues through the use of information technology. The company helps its clients to, for example, control hardware, facilities and IT operational costs through its managed and IT consulting services.

Banks and other institutions use huge amounts of money and other resources to protect sensitive information. According to Innovation Asia Staff Global information technology spending by financial services institutions is expected to reach US$363.8 billion in 2011, an increase of 3.7% over 2010 (CFO Innovation Asia Staff, 2011). But in spite of huge investment into security, data breaches continue to occur. Breaches occur within all types of organizations but most of them happen in companies that provide financial services. Financial services, hospitality, and retail comprise the "Big Three" of industries affected (Verizon, 2010).

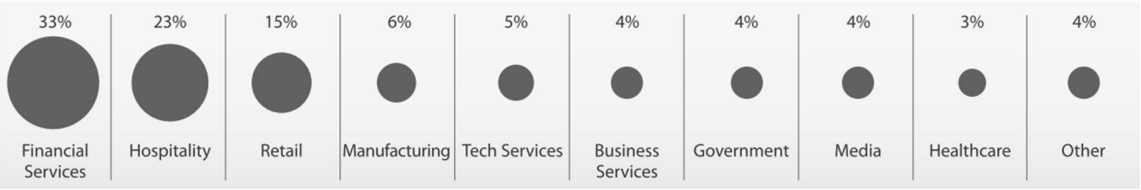

Figure 3 - Industry groups represented by percent of breaches (Verizon, 2010)

The targeting of financial institutions is quite obvious. Stealing sensitive information can be compared to the bank robbery because it represents the nearest approximation to actual cash for the criminal (Verizon, 2010). Financial organizations keep large amounts of sensitive consumer data for long periods of time that makes them more vulnerable for possible attacks.

Figure 4 shows that the number of breaches was highest in the middle size organizations that have from 1.001 to 10.100 employees. It can be explained by the fact that smaller organizations have less money to spend on security but also have fewer assets to protect. Large corporations in turn have more valuable assets but also possess more resources that can be used for information security. Information thieves seem to choose their victims according to perceived value of data and cost of attack (Verizon, 2010).

Figure 4 - Organizational size by percent of breaches (number of employees) (Verizon, 2010).

5.2 Threat agents

There are different attackers (threat agents) that cause or contribute to an incident. These can be divided into external, internal and partners (Verizon, 2010). External threats come from outside organization. These threats include hackers, organized crime groups, and government entities, as well as environmental evens such as weather and earthquakes. Internal threats come from within the organization. This includes company executives, employees etc., as well as internal infrastructure. Insiders are trusted and privileged in comparison to external entities that typically have no trust or privilege (Verizon, 2010). Partners encompass any third party sharing a business relation with the organization. This includes suppliers, vendors, hosting providers, outsourced IT support, etc. Some level of trust and privilege is usually implied between business partners (Verizon, 2010).

The statistics presented in figure 5 gives a picture of threat distribution. Figure 5 shows that external threats prevail over internal, partner and multiple threat agents. Most of the threats come from outside of the organization. Some threats have mixed origin, for example, when an outsider solicits or bribes the employee in order to embezzle or skim data and/or funds.

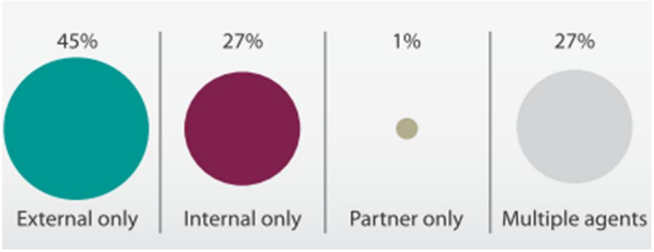

Figure 5 - Threat agents by percent of breaches (Verizon, 2010).

5.2.1 External agents

External threats originate from sources outside the organization and its network of partners. Examples include hackers, organized crime groups, and government entities, as well as environmental events such as weather and earthquakes. Typically, no trust or privilege is implied for external entities (Verizon, 2010). During the research Verizon experts discovered that external agents were involved in 70% of breaches that corresponds to 98% of records compromised in the 2009 caseload. 85% of all compromised records in 2009 attributed to organized crime. There is a level pegging between companies and criminals: while financial and other organizations invest money into security of information, criminal groups band together. Cooperation allows criminal groups to pool resources, specialize skills and distribute the work effort, among other advantages. Although a great part of cooperation between criminal organizations is largely tactical, the potential for broader alliances to undertake more complex criminal schemes in an increasingly global economy is significant (US Government Interagency Working Group, 2000).

5.2.2 Internal agents (insiders)

An insider is a person that has been legitimately empowered with the right to access, represent, or decide about one or more assets of the organization's structure (Probst, Hunker, & Gollmann, 2010). Internal threats are those originating from within the organization. This encompasses company executives, employees, independent contractors and interns, etc., as well as internal infrastructure. Insiders are trusted and privileged and some of them more than others.

The "insider threat" or "insider problem" has received considerable attention and is cited as the most serious security problem in many studies (Probst, Hunker, & Gollmann, 2010). It is also

considered the most difficult problem to deal with, because an "insider" has information and capabilities not known to other, external attackers (Probst, Hunker, & Gollmann, 2010).

Internal agents were involved in about a half of cases investigated by Verizon and United States Secret Service (USSS) in 2010. But those breaches were responsible only for 3% of compromised records. Nevertheless the possible damage caused by insiders can be much worse because Verizon did not include contributory errors into their report. Contributory error is, for example, when an employee unintentionally misconfigures an application and makes it vulnerable to attack by another agent.

5.2.3 Partner agents

Partners include any third party sharing a business relationship with the organization. This includes suppliers, vendors, hosting providers, outsourced IT support, etc. Some level of trust and privilege is usually implied between business partners.

According to 2010 DBIR partner agents were responsible for 11% of breaches (1% of records). Verizon findings show that that the majority of breaches involving partners are the result of third-party information assets and accounts being "hijacked" by another agent and then used to attack victims. This frequently involves a remote access connection into the victim's systems. If compromised, the malicious agent's actions would appear to come from a trusted source and therefore be even more difficult to detect and prevent (Verizon, 2010). Poor partner security practices usually allow or worsen these attacks (Verizon, 2010). The USSS caseload, on the other hand, shows most partner breaches stem from the deliberate and malicious actions of that partner. An example of this might be a third-party system administrator who maliciously misuses his/her access to steal data from the victim.

5.3 Threat actions

Verizon defines threat action as a description of what the threat agent did to cause or contribute on the breach. There are usually multiple actions across multiple categories during a breach scenario.

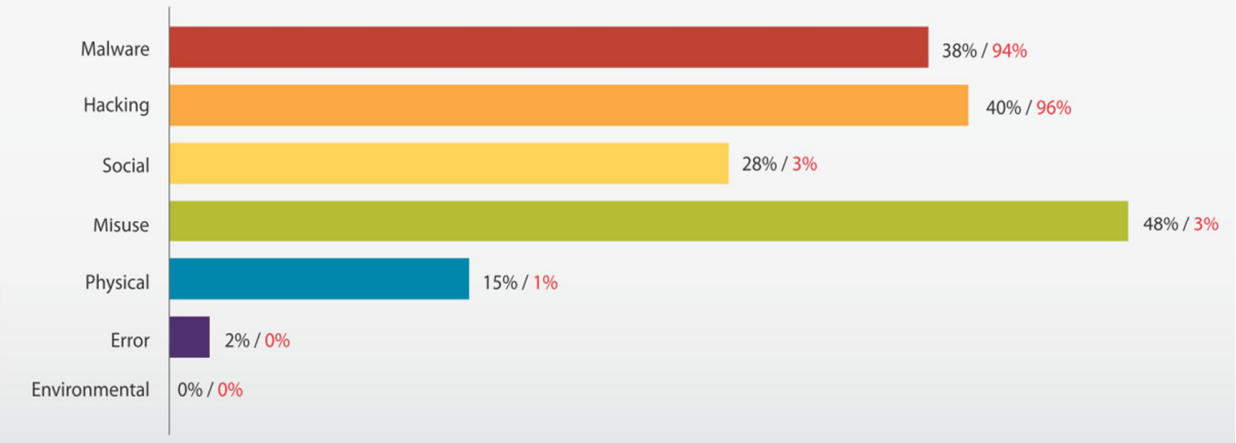

Figure 6 - Threat action categories by percent of breaches and records (Verizon, 2010).

As it comes from the figure 6, Hacking and Malware dominate with respect to number of records compromised. It can be explained by the fact that in the big breaches, the attacker hacks into the victim's network and installs malware on systems to collect data (Verizon, 2010).

5.3.1 Malware

The term malware (slang for malicious software) is also commonly used for rogue software. Malware is any type of software designed specifically to disrupt a computer or its operations (Salomon, 2006). Malware is designed to cause damage: this includes viruses (spread by attaching to other files and infecting them), worms (propagate through networks by infecting other computers), Trojan Horses (conceal their real purpose of causing damage by claiming to be a harmless program in order to be installed by the user) but also spyware and adware, especially those programs that try to reinstall themselves from an invisible copy after the original has been deleted (Salomon, 2006). While some of these programs cause damage, at worst, by consuming resources for spreading themselves, there are also cases in which data are deleted, manipulated or transferred to third parties, system crashes are triggered or programs are installed on the infected computers allowing hackers to access the system through a "backdoor" (Oesterreichische Nationalbank, 2006).

During the 1980s and 1990s, it was usually taken for granted that malicious programs were created as a form of vandalism or prank to enhance their "fame and glory". While early forms of malware damaged or crashed computer systems, most malware applications today are designed for financial gain (Van Luvender, 2011). During the past few years, considerable evidence points to the fact that the generation, distribution and use of malware is driven predominantly by economic interests

(Johannes, Eaten, & Wu, 2008). Actors in the underground malware economy will continue to pursue these activities, as long as benefits from semi-legal and illegal activities outweigh the costs of these activities, including the expected costs of sanctions (Johannes, Eaten, & Wu, 2008). Due to the relatively low cost of launching fraudulent or criminal activities in cyberspace and the high potential gains, the economic incentives to expand cybercriminal activity continue to be strong (Johannes, Eaten, & Wu, 2008). Data-stealing malware is a threat that divests victims of personal and proprietary information with the intent of monetizing stolen data through direct use or underground distribution. According to Verizon 2010 DBIR, malware is responsible for 38% of all breaches and 94% of all data lost in 2009. The hazard resulting from malware is also considerably higher in the absence of appropriate protective measures because numerous viruses and worms are detected and eliminated by up-to-date protective software without any problems, but constitute a serious threat when outdated or no anti-virus systems are used (Oesterreichische Nationalbank, 2006).

There are different ways of how malware can be installed on the victim's computer. Figure 7presents malware infection vectors in the Verizon 2010 DBIR.

Figure 7 - Malware infection vectors by percent of breaches within Malware (Verizon, 2010).

5.3.2 Hacking

Early use of the term "hacker" was applied to computer hobbyists who spent their spare time creating video games and other basic computer programs. However, this term acquired a negative connotation in the 1980s when computer experts illegally accessed several high-profile databanks (Referense for Business, 2011).The introduction of relatively inexpensive personal computers and modems made this pastime affordable. The use of regular telephone lines as access ways made it possible. Over time, the designation "hacker" came to be associated with programmers and disseminators of computer viruses (Referense for Business, 2011). However, "hacking" is not only

about creation and spreading of viruses. This definition encompasses a wide range of other computer crimes as well, many of them primarily grounded in efforts to make money. Most of information today is kept and transferred electronically that has made it a target for malicious attackers.

In the early days of hacking and breaking into computers, some security experts maintained that "hackers have done less damage to corporate computer systems than overflowing lavatories" (Salomon, 2006). Today, such a claim seems ludicrous. The damage done to computers, to networks, to individuals, and to the economy is getting worse and has become a global concern. Fighting it involves governments, law enforcement agencies, and security experts all over the world.

Hacking affords the criminal many advantages because it can be accomplished remotely and anonymously, it doesn't require direct interaction or physical proximity, and there are many tools available to automate and accelerate attacks. The latter allows even less-skilled agents to cause a lot of trouble (Verizon, 2010).

5.3.3 Types of hacking

In this section I would like to examine the types of hacking observed in Verizon's and the USSS's 2009 caseloads. According to Verizon, hacking encompasses all attempts to intentionally access or harm information assets without (or in excess of) authorization by thwarting logical security mechanisms.

Figure 8 shows that two types of hacking are responsible for the majority of records compromised. These types are "Use of stolen login credentials" and "SQL injection". The use of stolen credentials is most used type of hacking in both Verizon and USSS datasets. One of the main reasons behind this is the proliferation of password-gathering malware. Stolen credentials give a hacker an opportunity to disguise himself as a legitimate user that allows a hacker to feel more "comfortable" because authenticated activity is much less likely to be noticed by detection mechanisms (Verizon, 2010).

Structured Query Language (SQL) is a computer language used for database programming. SQL injection is a technique to maliciously exploit applications that use client-supplied data in SQL statements (Oracle, 2009). SQL injection involves entering SQL code into Web forms such as login fields or browser address fields to access and manipulate the database behind the site or system (Van

Luvender, 2011). In simple terms, it tries to fake out the login function using SQL commands instead of actual usernames and passwords. A successful attack allows criminals to access, modify or delete information from databases such as e-mail addresses, personal information and credit card numbers. SQL injection is a popular attack method in the underground economy due to its versatility. In addition to stealing database information, it gives hackers access to Web-site content, which can be easily manipulated to allow other attacks from the compromised site such as the distribution of malware (Van Luvender, 2011).

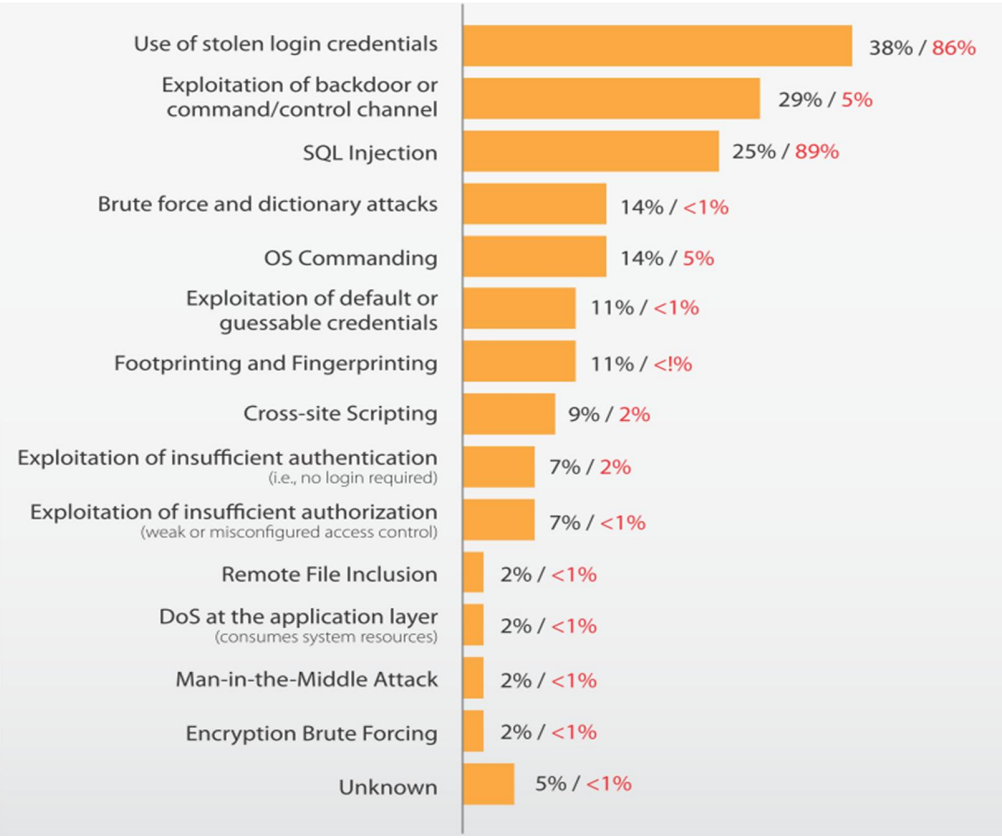

Figure 8 - Types of hacking by percent of breaches within Hacking and percent of records (Verizon, 2010).

Exploitation of backdoors is another common method of network and system intrusion. A backdoor is a remote administration utility which bypasses normal security mechanisms to secretly control a program, computer or network (F-Secure, 2011). These utilities may be legitimate, and may be used for legitimate reasons by authorized administrators, but they may also be misused by attackers (F-Secure, 2011). A backdoor is usually able to gain control of a system because it exploits

vulnerabilities, bugs or undocumented processes in the system's code. In most Verizon cases a backdoor was created by as a function of malware that was installed at an earlier stage of the attack.

5.4 Attack Pathways

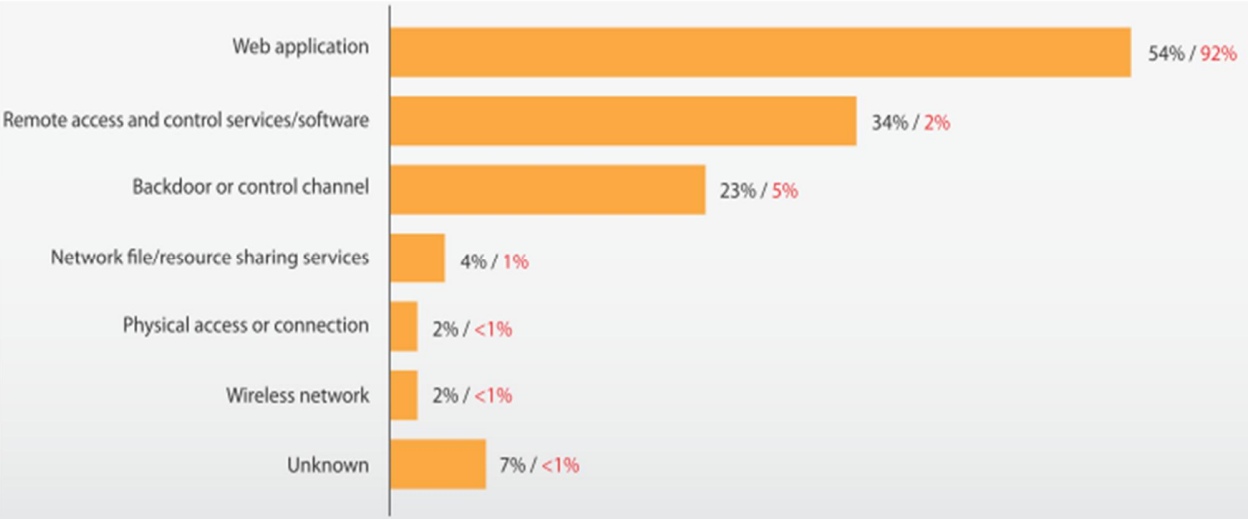

Figure 9 - Attack pathways by percent of breaches within Hacking and percent of records (Verizon, 2010).

In additions to threat categories described in the previous section, it is important to discuss the pathways exploited by attackers as they conduct their malicious activities. In context of data breach, the pathway refers to the interface through which an attacker gains access to corporate systems (Verizon, 2008).

Both Verizon and USSS cases show that “web application” is on the top of the list for number of breaches (54%) and the number of records compromised (92%) in 2009. Unlike most information assets which have limited visibility outside the organization, web applications are by design accessible to the outer world (Verizon, 2008). Application software that does not properly check the user input, could be vulnerable to compromise (Center for Strategic and International Studies, 2009). Attackers can inject specific exploits, including buffer overflows, SQL injection attacks, and cross-site scripting code to gain control over vulnerable machines (Center for Strategic and International Studies, 2009).

Attacks against vulnerabilities in web-based and other application software have been a top priority for threat agents in recent years that can be backed by the fact that the number of breaches through

web applications raised from 34% in 2004-2008 Verizon caseloads to 54% in 2009. But in spite of its "popularity" in recent years, this attack pathway was responsible only for 22% of breaches in Verizon 2011 DBIR.

In over 30 percent of the breaches investigated by the Verizon 2010 study, an attacker gained unauthorized access to the victim via one of the many types of remote access and control software. On many occasions, an account which was intended for use by vendors in order to remotely administer systems was compromised by an external entity. These vendor accounts were then used to illegitimately access enterprise information assets. This scenario is particularly problematic due to the fact that, from the victim's perspective, the attacker appears to be an authorized third party (Verizon, 2008). In many of these cases, the remote access account was configured with default settings, making the attacker's job all too easy.

5.4.1 Social

Social tactics or social engineering can be defined as an art or better yet, science of skillfully maneuvering human beings to take action in some aspect of their lives (Hadnagy, 2011). In other words it is the act of manipulating a person to take an action that may or may not be in the "target's" best interest (Hadnagy, 2011). This may include obtaining information, gaining access, or getting the target to take certain action. According to Verizon, social tactics employ deception, manipulation, intimidation, etc. to exploit the human element, or users, of information assets. These actions are often used in conjunction with other categories of threat (i.e. malware designed to look like antivirus software) and can be conducted through technical and non-technical means (Verizon, 2010). Software vendors are becoming more skilled at creating software that is hardened, or more difficult to break into. As hackers are hitting more hardened software and as software and network attack vectors, such as remote hacking, are becoming more difficult, hackers are turning to social engineering skills. Often using a blend of technical and personal skills, hackers are using social engineering in major attacks as well as in minor breaches throughout the world (Hadnagy, 2011).

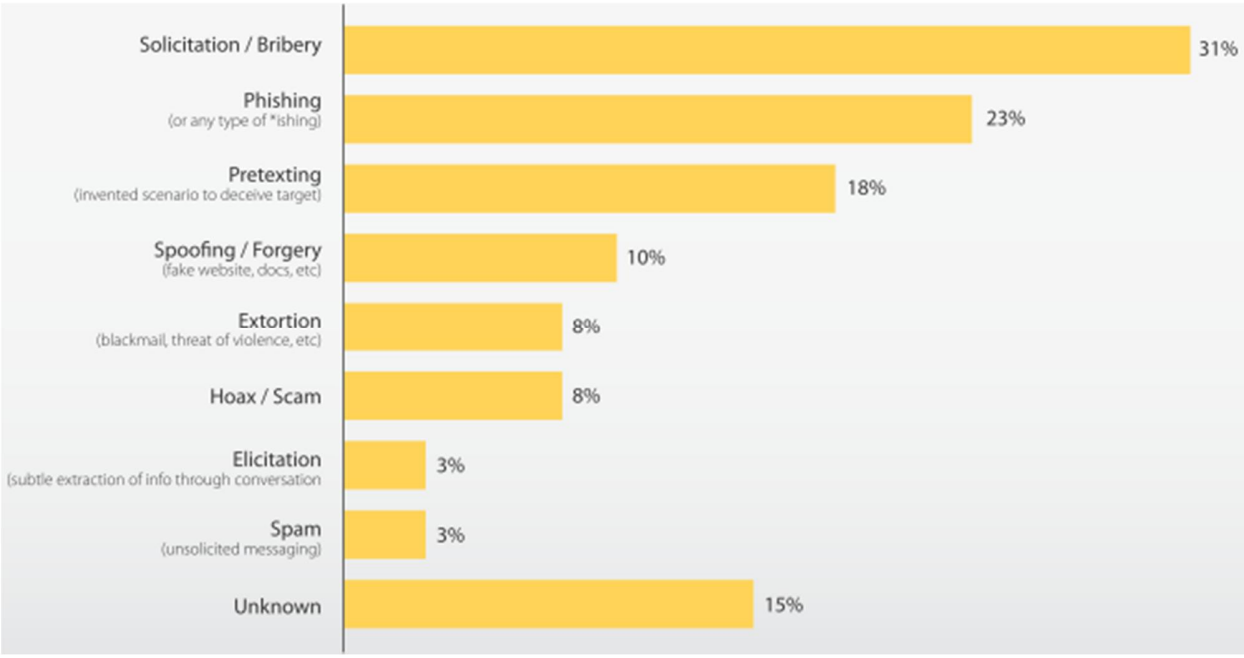

Figure 10 - Figure. Types of social tactics by percent of breaches within "Social" (Verizon, 2010).

Solicitation and bribery account for 34% of social tactics recognized by the VERIS framework in the 2009 caseload. These were scenarios in which someone outside the organization conspired with an insider to engage in illegal behavior (Verizon, 2010). According to USSS, these are usually organized criminal groups conducting similar acts against numerous organizations. They recruit, or even place, insiders in a position to steal sensitive data, usually in return for some cut of the score (Verizon, 2010). According to recently published Verizon 2011 DBIR, solicitation and bribery remains the most common type of social tactic in 2010, but by a much wider margin than before. This type of social tactics is responsible for 74% of breaches within "Social" in Verizon 2010 dataset. This frequently entails collusion between an external agent and an insider, though other combinations occur as well (Verizon, 2011).

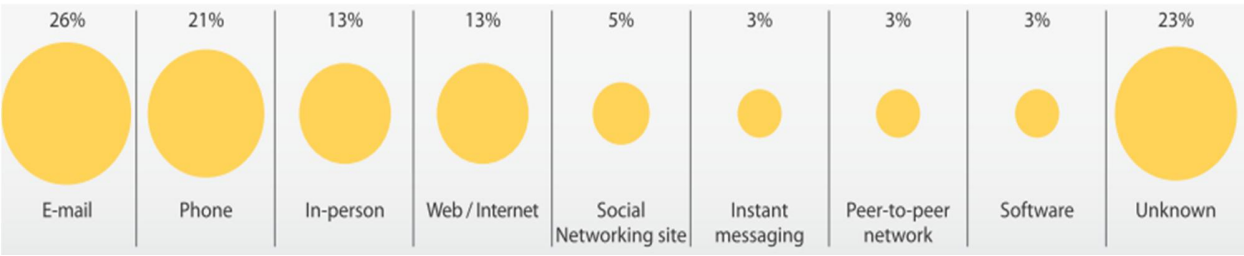

Figure 11- Paths of social tactics by percent of breaches within Social (Verizon, 2010).

As it comes from figure 11 e-mail and phone were most often used paths for social attacks in 2010 DRIB. E-mail as a social engineering tool often contains a topical subject that is supposed to trigger an emotion that leads to unwitting participation from the target (Malcolm, 2006). This type of social engineering involves malicious code, such as that used to create a virus. This code is usually hidden within a file attached to an email and the intention is that an unsuspecting user will click/open the file.

E-mail is also dominating in Verizon 2009 DRIB but if we will look at results of the 2011 investigation we can find out they are quite surprising. In 2011 DBIR in-person contact tops the number of breaches (78%) within social. According to this result it is possible to conclude that attackers staked on the personal touch with their victims and this is quite understandable. Even in modern high-tech business world many deals won't get done without an in-person "meet-and-greet" (Verizon, 2011).

5.4.2 Misuse

Misuse implies use that is contrary to expected operational behavior (Probst, Hunker, & Gollmann, 2010). But this definition is very broad and in some degree is oversimplified. In practice, the concept of misuse is meaningful only with respect to a policy that defines what usage is acceptable and what is not. Unfortunately, a basic gap exists between use that is intended to be acceptable and use that is actually possible (Probst, Hunker, & Gollmann, 2010). Sometimes it is quite difficult to distinguish between what is possible (because of design flaws and implementation bugs) and what is actually authorized, as well as limitations that result from inadequate granularity and expressiveness of access controls (Probst, Hunker, & Gollmann, 2010). Both Insiders acting maliciously and insiders acting unintentionally can affect negatively the organizations they are part of. Hence there are always going to be gray areas in how security policies define both insider misuse and proper behavior. Furthermore the apparent success of what might be considered accidental but tolerated misuse could easily inspire subsequent malicious misuse (Probst, Hunker, & Gollmann, 2010).

Verizon RISK team states that both misuse and hacking can utilize similar vectors and achieve similar results but in case of "Misuse" the agent inappropriately uses granted access whereas with hacking, access is obtained illegitimately. The category of misuse is exclusive to parties that enjoy a degree of trust from the organization like insiders and partners (Verizon, 2010).

Embezzlement (act of dishonestly appropriating or secreting assets by one or more individuals to whom such assets have been entrusted), skimming (theft of credit card information used in an otherwise legitimate transaction) and related fraud were seen more often than other forms of misuse in 2009 dataset and were exclusive to cases worked by USSS. These actions are typically perpetrated by employees entrusted with the oversight or handling of financial transactions, accounts, record keeping, etc.

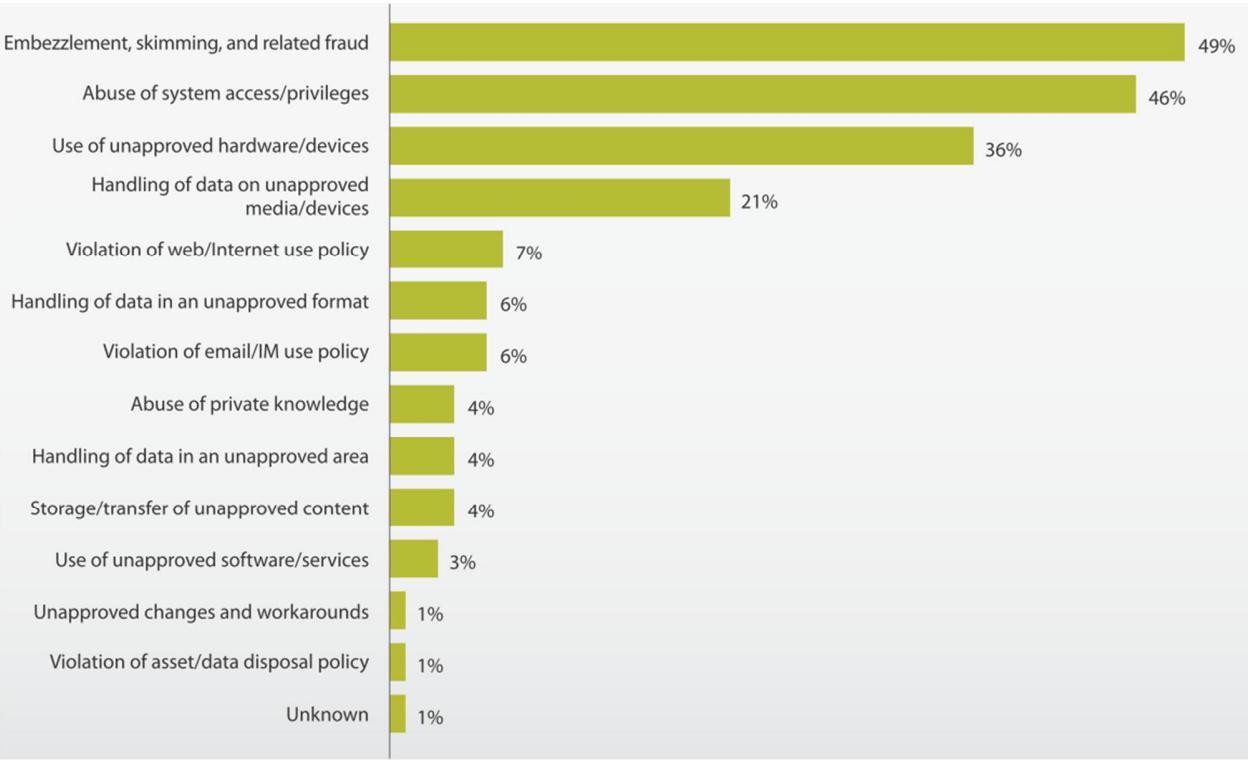

Figure 12 - Types of misuse by percent of breaches within Misuse (Verizon, 2010).

Misuse was the most common of all threat actions (48%) in Verizon 2009 dataset while being responsible only for 3% of records breached. This can be explained by the fact that an employee engaging in this type of fraud has different goals than a hacker because he or she has a clear interest in keeping their job, remaining undetected and avoiding prosecution (Verizon, 2010). Stealing small amounts of data or monetary assets over a longer period of time is more suited to this than a "grab as much as you can and run" approach usually used by outsiders. Insiders also have the luxury of targeting exactly what they want in the amount they want and when they want it.

Abuse of system access and privileges follows a close second behind embezzlement. As the name implies, it involves the malicious use of information assets to which one is granted access. Though

not apparent from figure 11 above, the abuse of system access tends to compromise much more data than embezzlement and other types of misuse (Verizon, 2010). It often involves privileged insiders like, system and network administrators (especially the larger breaches) but also other types of employees.

Handling of data on unapproved media and devices was the next common type of misuse in both Verizon and USSS caseloads. Sometimes the devices themselves are contraband but more often the data in question is not approved for storage on an otherwise sanctioned device. Verizon findings show that success of a breach does not depend on the perpetrator ability to use a certain portable device. Unfortunately, insiders have plenty of choices when it comes to media and devices fit for stealing data out of their employer (Verizon, 2010).

5.4.3 Error

Error refers to anything done (or left undone) incorrectly or inadvertently (Verizon, 2011). Given this broad definition, some form of error could be considered a factor in nearly all breaches. Poor decisions, omissions, misconfigurations, process breakdowns, and the like inevitably occur somewhere in the chain of events leading to the incident (Verizon, 2010). For this reason, it is important to distinguish between errors as a primary cause of the incident vs. contributing factor. If error is a primary cause it independently and directly progresses the event chain leading to an incident. On the other hand, if error is a contributing factor, it creates a condition that allows the primary chain of events to progress (Verizon, 2010). For example, a misconfiguration that makes an application vulnerable to attack is a contributing factor whereas one that immediately crashes the server is the primary cause.

Variety of errors is never-ending. An interesting case happened at Lincoln National Life Insurance when usernames and passwords for agents and authorized brokers were printed in a brochure (Jones, 2009). The brochure was also posted on an agent's public website. The login information enabled access to a website containing medical records and other personal information from individuals seeking life insurance. Applicant name, Social Security number, address, policy number, driver's license number and credit information was also on the website. About 27000 records were compromised because of the accident (Office of Inadequate Security, 2010). Another example describes a 30-year-old man who made more than 1,000 fake ID cards that he used to rip off people,

stores and banks. The criminal said that it was easy to find new victims: All he needed to do was to search the dumpsters of local banks. The sensitive information he found in the trash was used to mock up fake identification cards and blank checks.

5.4.4 Physical

This category includes human-driven threats that employ physical actions and/or require physical proximity. In almost half of the Verizon and USSS cases involving physical actions, theft was the type. Typically, the assets that were stolen were documents, but also frequently included desktop or laptop computers. According to Verizon, theft typically occurred in non-public areas within the victim's facility like offices and storage rooms, although there were a few exceptions to this rule.

5.4.5 Environmental

This category not only includes natural events like earthquakes and floods but also hazards associated with the immediate environment (or infrastructure) in which assets are located. The latter encompasses power failures, electrical interference, pipe leaks, and atmospheric conditions. Although environmental hazards most often affect the attribute of availability, they do occasionally factor into scenarios resulting in the loss of confidentiality as well. Verizon, for instance, investigated incidents in the past in which a power outage led to a device rebooting without any of the previously-configured security settings in place. An intruder took advantage of this window of opportunity, infiltrated the network, and compromised sensitive data (Verizon, 2010).

6. Insider threat to organization security

Previous section gives an idea of different threat actions and threat agents. In this chapter I would like to drill down into to the issue of the insider threat. According to Verizon report, insiders were not responsible for a lot of compromised records but at the same time they were involved in 48% of cases investigated by Verizon and USSS in 2010. This fact is supported by Celent, a research and consulting firm focused on the application of information technology in the global financial services industry. Celent claims that approximately 60 percent of bank fraud cases where a data breach or theft of funds has occurred are the work of an insider (Jeghler, 2008). As it was mentioned in previous chapter, insider malicious fraud accounts for a relatively small percentage of all data breaches within the financial services industry. From 2005 to 2008 insider fraud accounted for about 9% of all data breaches investigated by Celent (Jeghler, 2008). This begs the question of how many incidents are actually communicated to affected customers. Celent estimates that up to 50% of these incidents go unreported. Although a fair percentage of incidents are communicated, those that are not communicated in a timely manner are also problematic. This delay poses a substantial risk, as the public backlash can be strong. Incidents that go undetected pose the greatest risk to financial institutions (Jeghler, 2008). Statistics vary regarding the prevalence of cases perpetrated by insiders compared to those perpetrated by individuals external to the targeted organizations. Nevertheless, insiders pose a substantial threat by virtue of their knowledge of and access to their employers' systems and/or databases, and their ability to bypass existing physical and electronic security measures through legitimate means (U.S. Secret Service & CERT Coordination Center, 2004).

Insider bank fraud is perpetrated by someone who works inside, or has access to restricted areas or information inside of the financial institution. This kind of fraud can be difficult for banks to defend against, since so many people are put in a position of responsibility with the bank's assets. There is an opinion that it is difficult to say for sure how big of a problem is insider bank fraud because banks don't have established industry standards for measuring these losses (Financial Fraud Law, 2009). As it mentioned before in this work, the financial damage can be difficult to quantify because in addition to direct remediation costs from fraud incidents there are also indirect costs from for example damaged reputation.

Organized criminals and hackers may dominate the news headlines, but employees are by far the most significant source of fraud exposure for banks today (Inscoe & Krishna, 2009). Insider fraud is

a silent crime that is only publicized when the perpetrators are caught, and those represent only a small fraction of the total number of employees committing fraud today (Inscoe & Krishna, 2009). In many cases, the people committing the crimes are the banks' top performers because the same creativity, attention to detail and intelligence that made them successful at banking also makes them successful at fraud (Inscoe & Krishna, 2009). Unfortunately, employees and contractors who access financial institution systems during the course of work know the system better than anyone else and they are better positioned to exploit the systems' vulnerabilities.

As with other forms of fraud, the face of internal fraud is changing due to the flourishing and increasingly sophisticated underground economy. Historically, employee fraud involved account skimming and other small-scale attacks that put money in the employee's pocket. Today, with access to the online fraud forums, employees can advertise and sell customers' personal and financial information and make money without stealing directly from accounts. Because it seems less direct, employees have an easier time rationalizing this type of fraud, especially those who are acting out of desperation (Van Luvender, 2011).

Verizon researches help to understand what categories of insiders cause most of the breaches. According to Verizon regular employees were responsible for 51 percent of the breaches in 2009 caseload and 85 percent in 2010 caseload. Regular employees are bank tellers, retail cashiers and other similar personnel taking advantage of their everyday job duties to steal data. Next category is finance and accounting staff. These insiders have usually higher privileges than regular employees. Their oversight and management of accounts, records and finances affords them great propensity for harm (Verizon, 2010). Third place goes to system/network administrators who were responsible only for 12% of breaches but at the same time were responsible for most of the compromised records caused by insiders in 2009. This is quite obvious because higher privileges usually give bigger opportunity for abuse. The Verizon 2011 DBIR shows that for the second year in a row, it is regular employees who are behind the majority of data compromises.

Table 1- Types of internal agents by percent of breaches within Internal

Position Year	2009	2010
Regular employee/end-user	51%	85% (+34%)
Finance/accounting staff	12%	22% (+10%)
System/network administrator	12%	3% (-9%)
Executive/upper management	7%	11% (+4%)
Helpdesk staff	4%	4%
Software developer	3%	2% (-1%)
Others	1%	1%
Unknown	9%	1% (-8%)

For classification purposes, Verizon groups insiders into three major categories. They either acted deliberately and maliciously, inappropriately but not maliciously, or unintentionally without malice. Much like in 2009 dataset, investigators determined that nearly all internal breaches (93%) in 2010 dataset were the result of deliberate malicious activity. This may seem odd, but it can be explained by the fact that Verizon DBIR's are based on data loss cases investigated by either a third party forensics group or a law enforcement agency. In addition, if the insider's only involvement was related to a conditional event, Verizon did not consider insiders as a primary threat agent and thus not depicted in the statistics above.

Study of data breaches occurred during 2005-2009 conducted by Suzanne Widup provides and interesting statistics for data breaches that are not only malicious in nature. The study is based on incident reports from the Open Security Foundation, the Privacy Rights Clearinghouse, Sound Assurance and the Identity Theft Resource Center (Widup, 2010). These reports were combined and normalized for the time period of January 2005 through December 2009. The final data set contained 2,807 incidents from these sources.

Figure 13 shows that when the incident involves insiders, it is more than twice as likely to be accidental in nature as someone behaving maliciously. Insider's mistakes caused far less damage in the first three years of the study than they did in 2008 and 2009, when they caused millions of records to be disclosed.

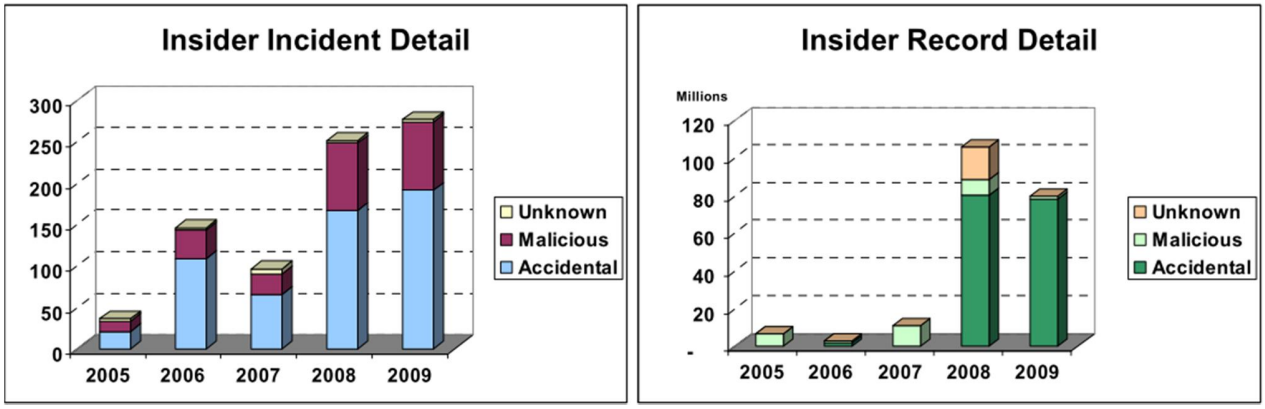

Figure 13 - Insider incident detail and Insider record detail (Widup, 2010).

Table 2 presents distribution of compromised records between groups of insiders and as we can see the prevailing majority of compromised records are accidental in nature.

Table 2 - Insider records detail (Widup, 2010).

	2005 - 2009
Insider	
Accidental	160,359,601
Malicious	28,358,365
Unknown	17,232,779
Total	**205,950,745**

The analysis conducted by Suzanne Widup gives a clearer picture of the insider threat to the organization because it includes accidental data breaches. Nevertheless it is difficult to say for sure how big a problem of insider threat is. As it was mentioned previously, reputational impact worries organizations the most. In addition, the admission of a significant vulnerability could flag other attackers, so very few companies are willing to be public about intellectual capital losses. McAfee (leading security solutions vendor) claims that around half of organizations reported that a data breach involving sensitive information or intellectual property is their number one concern. One in seven organizations has not reported data breaches and/or losses to outside government agencies or authorities, or stockholders (McAfee/SAIC, 2011). Only three in ten organizations report all data breaches/losses suffered, while one in ten organizations will only report breaches/losses that they are legally obliged to, and no more (McAfee/SAIC, 2011). Six in ten organizations currently “pick and choose” the breaches/losses they report, depending on how they feel about them (McAfee/SAIC, 2011).

Bayesian Network Modeling for Analysis of Data Breach in a Bank

Internet is making information more accessible to public and easier to distribute. Different sites like WikiLeaks, for example, pose new threats to businesses, as insiders will be increasingly tempted to release their company's secrets for financial or technological gain, to increase the level of transparency of organizations, or to expose what they believe is wrongdoing (McAfee/SAIC, 2011).

7. Critical security areas and controls

As it was mentioned previously, an insider is a person that has been legitimately empowered with the right to access, represent, or decide about one or more assets of the organization's structure. The previous section proves that insider threat is a very comprehensive issue that affects all organizations. The "insider problem" should receive considerable attention because it is one of the most serious security issues in sensitive data perspective. It is also may be considered the most difficult problem to deal with due to the fact that insiders are trusted and therefore present a great threat to information security. They typically have greater knowledge than outsiders about system vulnerabilities. Therefore, the chances of a successful attack can be greater for an insider attack than for an outsider attack. For instance, the knowledge that a malicious insider has about the sensitivity of information gives him/her a better chance to breach information confidentiality. Even more challenging is the malicious expert insider. This type of user can perform harmful actions while behaving almost indistinguishably from normal users, making detection very difficult (AlGhamdi, Wright, & Barbará, 2005). This section has an objective to give an overview of most critical areas and controls that are common for the majority of organizations and provide a basis for the developing of a BN.

After each DBIR the Verizon RISK gives a list of conclusions and recommendations to the companies. Verizon experts claim that creating a list of recommendations gets progressively more difficult every year. But at the same time the recommendations are never new and unexpected. This can be explained by the fact that the security woes are not caused by the lack of something new. They almost surely have more to do with not using, under using or missing something old (Verizon, 2011). According to Verizon 2011 DRIB the 63% of breaches could have been prevented by implementing simple and cheap measures, 33% by intermediate preventing measures and only 4% of data breaches needed difficult and expensive preventive measures. As is the case elsewhere, various approaches to coping with insider threats may also be applicable to coping with outsider threats and vice versa - various approaches to coping with external attackers may also be applicable to coping with insiders (Probst, Hunker, & Gollmann, 2010).

7.1 Preconditions for data breach

A policy violation is a deviation that basically breaks the rules of the organization. Violations can come in all shapes and forms and some are obviously more severe than others. Verizon RISK team

recommends watching for “minor” policy violations. Verizon researches show that there is a correlation between “minor” policy violations and more serious abuse. That is why Verizon recommends that organizations be wary of and adequately respond to policy violations. Based on case data, the presence of illegal content, pornography, etc. on user systems (or other inappropriate behavior) is a reasonable indicator of a future breach. Actively searching for such indicators rather than just handling them as they pop up may prove even more effective (Verizon, 2010).

Internal whistleblowers are employees who bring wrongdoing at their own organizations to the attention of superiors. To benefit from internal whistleblowing, an organization should create a culture that encourages employees to ask questions early - to point out issues and show courage in confronting unethical or illegal practices. It is very important to encourage internal whistleblowing - that is, to an authority within the organization - to preclude external whistleblowing and the resulting damage to an organization.

An interesting case happened at TJX Companies, the mammoth US retailer whose substandard security led to the world's biggest credit card heist (Goodin, 2008). TJX Companies fired an employee after he left posts in an online forum that made disturbing claims about security practices at the store where he worked. Security was so lax at the TJ Maxx outlet located in Lawrence, Kansas, that employees were able to log onto company servers using blank passwords. Because of this policy a massive network breach had leaked the details of more than 94 million customer credit cards. The employee was fired after managers said he disclosed confidential company information online. Other security issues included a store server that was running in administrator mode, making it far more vulnerable to attackers. The employee said he brought the security issues to the attention of a district loss prevention manager in 2006, and repeatedly discussed them with store managers. Except for a stretch when IT managers temporarily tightened password policies, the problems went unfixed.

7.2 Security controls

As it was mentioned before, Verizon DBIRs show that for the second year in a row, it is regular employees who are behind the majority of data compromises. These employees aren’t normally escalating their privileges in order to steal data because they don’t need to. They simply take advantage of whatever standard user privileges were granted to them by their organizations. Verizon

case findings suggest that regular employees typically seek “cashable” forms of information like payment card data, bank account numbers, and personal information.

Probably the most dangerous of all are privileged and highly sophisticated users and that is why they should be restricted and monitored more than others. Such users like executives, system administrators, and developers are fully aware that the system is baited and will employ sophisticated tools to try to analyze, disable, and avoid security mechanisms entirely. They usually steal larger quantities and more valuable forms of information. These super users should not be given more privileges than they need and separation of duties should take place. User activity and especially privileged one should be logged and generate messages to management. Unplanned privileged use should generate alarms and be investigated.

Users typically need a means of moving data once they have misappropriated it. Some use corporate or personal e-mail to send it to external parties or accounts. Some smuggle it out on various types of personal devices or media. Others use approved devices, but for unapproved purposes or in an unsanctioned manner. Verizon RISK team claims that the success of a breach does not hinge on the perpetrator being able to use a certain portable device (i.e., plugging up USB slots doesn’t eliminate the problem). Unfortunately, users have plenty of choices when it comes to media and devices fit for secreting data and removing it from their employer. For this reason, it is generally easier to control data at the source with access control mechanisms than it is to block a virtually limitless array of potential destinations.

It is also very important to have security software in place. According to the fact that software is never flawless, it is vital that latest patches are regularly installed and antivirus definitions are up-to-date. The insider can, for example, install malware to capture login credentials (in order to avoid access controls) but updated systems make the probability of success of this attempt lower that in case when definitions are outdated. But even if the latest software is installed and it is up-to-date, it does not eliminate the problem of malware. Even online solutions like, for example, www.virustotal.com where one can upload a file and scan it with over forty different antivirus engines do not guarantee the harmlessness of a file because currently there is no solution that offers 100 percent effectiveness in detecting viruses and malware. The way to deal with this problem can be monitoring, logging of user activity and timely analysis of user’s activity for signs of abnormal behavior.

7.3 Access Control

Defending against an insider who attempts to misuse his access privileges is one of the most significant problems facing information's security. An authorized insider can violate a system security policy for several reasons and in a multitude of ways but not all violations are true threats. An insider's privileges may range from those of a novice user to a system administrator. Hence, more attention must be paid to insider users allowed access to system resources in order to reduce risks imposed by them. Sensitive data and applications need to be protected so that only authorized individuals can access them.

One of the most interesting cases that resulted in about 1,2 million records-at-risk happened at Lincoln National Corporation (Financial Industry Regulatory Authority, 2011). Lincoln Financial Securities, Inc. (LFS) and Lincoln Financial Advisors Corporation (LFA) failed to adequately protect non-public customer information. From 2002 through 2009, between the two firms, more than 1 million customer account records were accessed through the use of shared user names and passwords. Since neither firm had policies or procedures to monitor the distribution of the shared user names and passwords, they were not able to track how many or which employees gained access to the site during this period of time. As a result of the weaknesses in access controls to the firms' system, confidential customer records including names, addresses, social security numbers, account numbers, account balances, birth dates, email addresses and transaction details were at risk. The Web-based system both firms used, combined non-public customer account information from various sources and allowed employees to view the customer account information within a single site. Home office personnel from both firms could access the system either by clicking on a link on the firm's website or could gain access through any Internet browser by going directly to the system's website and logging in with one of the shared user names and passwords. This incident resulted in imposing fines of 600 thousands of dollars not to mention the blemished reputation of the company.

There are different ways to obtain other user credentials that can be later used to get "authorized" access to sensitive information. That is why implementation of measures to thwart stolen credentials should take place in organizations and keeping credential-capturing malware off systems is priority number one. Organizations should consider two-factor authentication where appropriate and restricting administrative connections (i.e., only from specific internal sources). A "last logon" banner and training users to report/change passwords upon suspicion of theft also have promise

(Verizon, 2010). Policy-based controls enrich the traditional identity and roles-based access control model, enabling more complex rules based on other criteria such as time-specific restrictions or location of the requestor (Entrust, 2004).

7.4 Data breach

According to U.S. Secret Service and CERT Coordination Center case-study analysis, the insider threat activity examined in the banking and finance sector appears to involve an interaction among organizational culture, business practices, policies, and technology, as well as the insiders' motivations and external influences (U.S. Secret Service & CERT Coordination Center, 2004).

Motivation in general is an important question when dealing with insider threats and their consequences. This can cover the whole range from "innocent action", "fun", "technical challenge", "criminal intentions", to "espionage", or a combination of each of these factors. Surprisingly, even though one would expect the contrary, the effect of actions can be equally devastating for each of these motivations (Probst, Hunker, & Gollmann, 2010). This, of course, makes detecting a threat even more important—but also more complicated. Insiders might launch the attack for profit or sabotaging the target organization. Some of them might mount an attack for personal reasons such as taking revenge against the enterprise or even satisfy their plans to invoke some policy change inside an organization (Probst, Hunker, & Gollmann, 2010)

Insider threats represent an especially insidious threat to organizations. As trusted employees, they are given access to information that could compromise the organization if it falls into the wrong hands. Despite much research into the psychology and motivation of insiders, the fact remains that it is extremely difficult to predict insider attacks (Probst, Hunker, & Gollmann, 2010). This presents organizations with a dilemma. On the one hand, most cases of insider espionage could have been prevented by timely and effective action to address the anger, pain, anxiety, or psychological impairment of perpetrators well in advance of the crime (Probst, Hunker, & Gollmann, 2010). On the other hand, mistaken prediction of potential insider crime may have severe negative consequences for the individuals under scrutiny as well as the organization (Probst, Hunker, & Gollmann, 2010). A threat user is motivated in order to make some sort of adverse effect against the system. This motivation will influence the user's overall intention toward the system. A normal user, however, is not expected to be maliciously motivated and, therefore, his/her threat intention against

the system is expected to be very low (AlGhamdi, Wright, & Barbará, 2005) Apart from motivation an individual may be receiving pressure from friends, family or organized crime syndicates for reasons such as financial gain self-interest and/or revenge (Hadnagy, 2011).

8. Developing of a Bayesian Network

8.1 Model description

The development of a BN for quantitative analysis of insider threat is a challenging and difficult task. An insider is a potential threat to organization because he or she knows the company better than anyone who is outside of it. The situation is worsened by the fact that the insider cannot be absolutely restricted. Insiders have responsibilities and in order to do their jobs they need to have access to information and other company's assets. Companies have to trust their employees because without trust it is not possible to run business in the modern world. This trust is an operational risk that should be reduced as much as possible in order to prevent potential losses and save company's reputation. As it was mentioned before in this work, information is the most important asset for most of the modern companies and especially for financial institutions. The importance of information security and necessity of trust to insiders, should make the insider problem one of the prerogatives for every financial institution. The previous section gives overview of the most common problem areas for the companies that had experienced data breaches. The objective of this section is to summarize the previous findings in order to develop a BN that gives a clear presentation of malicious in nature data breach caused by insiders.

There are not many researches available describing how to deal with the insider problem, especially insider problem in the light of information security. Some companies like Symantec, Verizon and others publish different reports that can give a deeper understanding of information security problem. These reports are usually based upon data breaches investigated by one or another company. Verizon reports, in author's opinion, give the clearest picture of what threats modern companies encounter. Nevertheless, Verizon DBIRs reports published before 2010 were based purely on cases investigated by Verizon RISK team that can put a question of data reliability. However as from 2010 the Verizon DBIRs also include caseloads provided by USSS. This makes the data presented in those reports more reliable due to the fact that USSS caseloads cover data breaches that happened in many different organizations. The including of USSS's caseloads in Verizon 2010 DRIB raised the number of data breaches involving insiders up to 48 percent. Verizon reports focus not only on insider problem but the security of information at large. Nevertheless, various approaches to coping with insider threats may also be applicable to coping with outsider threats and vice versa (Probst, Hunker, & Gollmann, 2010).

Bayesian Network Modeling for Analysis of Data Breach in a Bank

As it was mentioned before, modeling of insider threat is a challenging task, especially in light of the fact that little relevant data is available. The purpose of the model presented in this work is not to encounter all possible reasons for data breaches conducted by insiders. This is a very challenging task demanding, for example, deep understanding of technical side of information security. This understanding goes well beyond either the scope of this master thesis or author's education and competence in this subject. The Bayesian model developed here is intended to show the most common and the same time most critical factors that were discovered during analysis of different reports and theoretical sources. The presented model can be viewed as a basis for further work. For example, the node "measures to thwart stolen credentials are implemented" is an input node in the model. However this node can also have such parent nodes as: anti-credential capturing malware, two-factor authentication, implementation of time-of-use rules, restricting of administrative connections and many other nodes that can also have their parents. To understand how all these factors influence the "measures to thwart stolen credentials are implemented" node, one should have a clear understanding of how do the IT systems work. It also seems to be quite problematic to find relevant historical data that can be used to set in probabilities into the model. The solution in this situation can be using of expert opinions. However these opinions can be quite subjective and not relevant to all organizations.

In order to make this model a substantial instrument for prevention of data breaches caused by insider, the group of experts from different fields should be involved. For example, the reasons for "insider's motivation to steal data" are quite difficult to identify. Identifying of these reasons per se can be a topic for an independent research, but nevertheless insider's motivation should be taken into account. However, despite much research into the psychology and motivation of insiders, the fact remains that it is extremely difficult to predict insider attacks (Probst, Hunker, & Gollmann, 2010). The node "advanced control of privileged users" is also presented as an input in the model. Privileged insiders are very difficult to control. They are trusted more than others and that is why they pose a more substantial risk to organizations. They should be "excessively" controlled and not given more privileges than needed in order to do their jobs. Different mechanisms can be used to control these insiders. These mechanisms also depend upon the organization and insider's role in the organization. For example, both CFO and system administrator are privileged insiders but very different preventive measures should be implemented to control them. Maybe, the most important in controlling of privileged users is to create a system in which the privileges are extensively partitioned. In other words, create the system where no user holds all the privileges, and where the

granted privileges are insufficient to gain possession of all other privileges (Probst, Hunker, & Gollmann, 2010). Modern security approaches are very advanced but nevertheless none of them is flawless. If there was a solution in the market that provides 100 percent security from data breaches, everybody should have implemented it and have forgotten about security issues. It is rather an ideal case when there are no "complete" insiders because in the absence of meaningfully secure systems and fine-grained access controls that are properly defined, properly implemented, and properly administered, that ideal is still practically not possible (Probst, Hunker, & Gollmann, 2010). Every organization should be aware of the fact that super users do pose very high risks and implement advanced control over them in accordance with organization specifics and insider's role in the organization.

As it was mentioned before the insider can act maliciously or accidentally. Accidental actions are actions without motivation to maliciously use data. It can be non-malicious policy violations or errors that lead to data breaches. Designing a model that captures both malicious and accidental data breaches is a very challenging task. Additional influencing factors like, for example, level of employee's training and competence should be also taken into consideration. Organization culture is something one should not forget about while designing such a model. Including of these factors makes the designing of the model nearly an impossible-doing task in the scope of a master thesis. Error, for example, is a very broad definition that can be a part of the majority of data breaches if not as a primary cause but at least as contributory factor. During the analysis of reports and theoretical sources the author came to the conclusion to design a BN that shows the probability of data breach that is malicious in nature. The model that presented in figure 14 was designed with the help of AgenaRisk software.

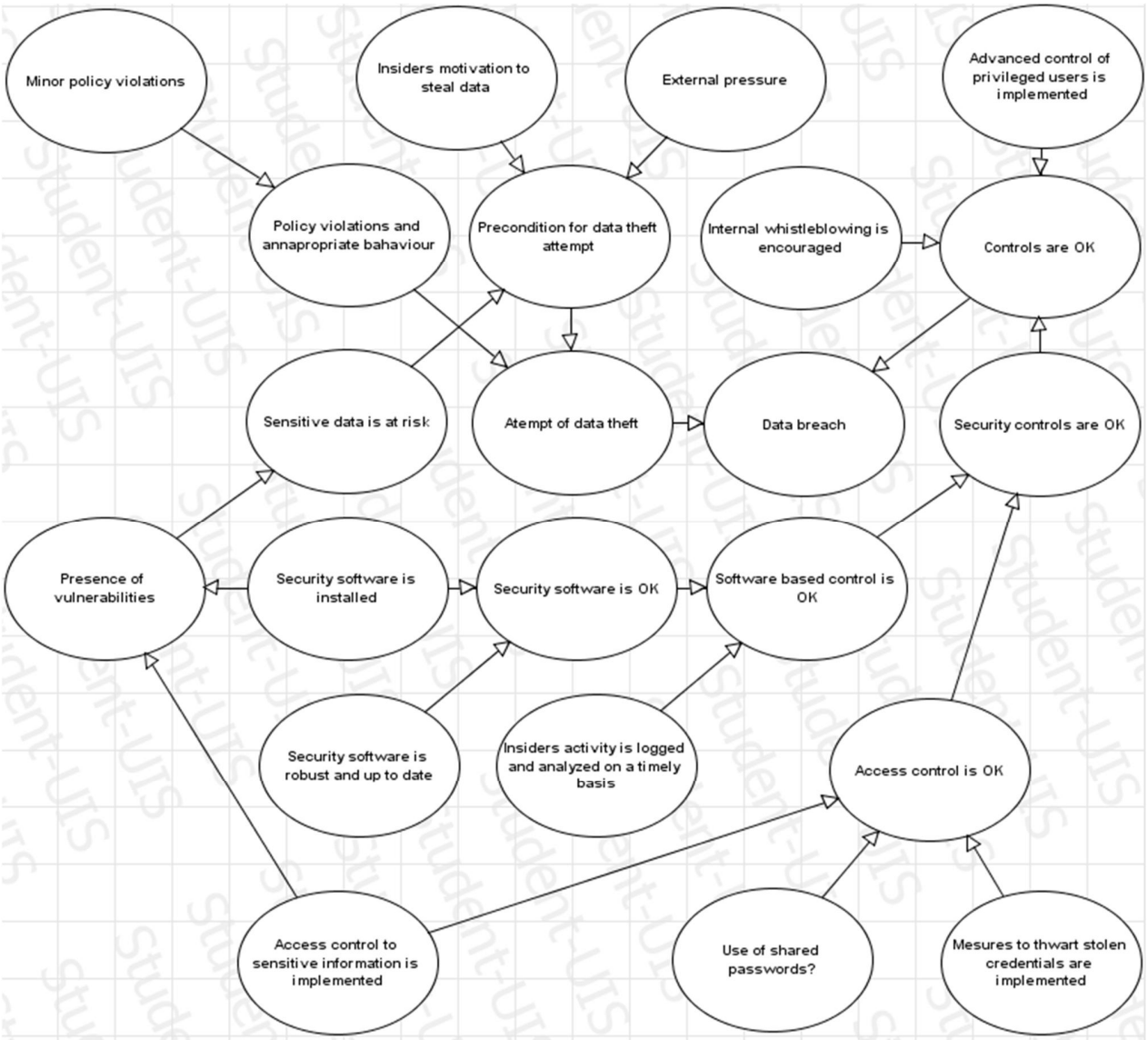

Figure 14 – Bayesian Network for a data breach in malicious insider perspective

Table 3 – Description of the nodes used in the Bayesian Network

Node Name	Node Description	States
Precondition for data theft attempt	Describes whether there is a precondition for data theft attempt in an organization.	True/False
Presence of vulnerabilities	Refers to an organization robustness to possible malicious activity.	True/False
Attempt of data theft	Refers to the whether attempt of data theft will occur in a company or not.	True/False
Sensitive data is at risk	Refers to whether sensitive information of a company is at risk.	True/False
Internal whistleblowing is encouraged	Refers to the organization attitude to whistleblowers; whether internal whistleblowing is encouraged or not.	True/False
Policy violations and inappropriate behavior	Describes whether policy violations and/or other inappropriate behavior occur in the organization.	True/False
Minor Policy Violations	According to Verizon there is high correlation between minor policy violations and major policy violations. Node refers to the degree of minor policy violations.	Many/Few
Controls are OK	Refers to the state of all the control presented in the model.	True/False
Access Control is OK	Refers to whether the access control is implemented and maintained in order to prevent unauthorized access.	True/False
Access control to sensitive information is implemented	Refers to whether the sensitive data is protected so that only authorized individuals can access it.	True/False
Security software is OK	Refers to whether the security software is installed and effective	True/False
Software based control is OK	Refers to whether the software based control is effective	True/False
Use of shared passwords?	Presence of shared passwords in the organization	True/False

	disputes the effectiveness of access control.	
Measures to thwart stolen credentials are implemented	Lack of measures to thwart stolen credentials disputes the effectiveness of access controls.	True/False
Security Controls are OK	Node describes whether the security controls are implemented and effective.	True/False
Advanced control of privileged users is implemented	Regular security controls are not very effective in relation to privileged users. Node refers to whether additional measures are implemented to control these users	True/False
Insiders activity is logged and analyzed on a timely basis	Logging of user activity without analyzing on a timely basis can be considered as useless in preventing of data breaches. Node refers to whether the logging is enabled and properly used	True/False
Security software is installed	Describes whether the security software is installed in the organization.	True/False
Security is robust and up-to-date	Describes whether the security software is regularly updated and robust to attempts of misconfiguration	True/False
Data breach	Gives the probability of data breach in an organization.	True/False
Insiders motivation to steal data	Describes whether the motivation to obtain sensitive information is high or low	High/low
External pressure	Refers to situation when an individual may be receiving pressure from friends, family or organized crime syndicates for reasons such as financial gain self-interest and/or revenge.	True/False

According to Verizon 2011 DRIB the 63% of investigated breaches could have been prevented by implementing simple and cheap measures, 33% by intermediate preventing measures and only 4% of data breaches needed difficult and expensive preventive measures. This can be explained by the fact that the security woes are not caused by the lack of something new. They almost surely have more to do with not using, under using or missing something old (Verizon, 2011). The model

presented in this work is designed in accordance with recommendations found in reports published by Verizon, Symantec and other investigators of data breaches. The model captures most common factors that have influence on malicious in nature data breaches and intends to be applicable to most financial institutions. The designed BN is not intended to present an ultimate solution for elimination of insider problem. The model summarizes findings from different data breach investigations and intends to be used as an instruction to evaluate the most common and, as different researches show, most critical areas in insider threat to data security perspective.

The model is developed in order to find out the probability that a data breach will happen or not. It consists of two major parts, “attempt of data theft” and “controls” that are implemented in the company to reduce the probability of data breaches. As the model is developed only for malicious in nature data breaches, the probability of data theft attempt is zero if insider has no personal motivation or there is no external pressure on him or her. Taking this fact into consideration the model was designed in the way that these two nodes are most influential input nodes. As it was mentioned before an independent research can be conducted to estimate the probability of user motivation. This can be done, for example, with the help of the HR department or a company psychologist if there is any.

“Precondition of data theft attempt” depends upon insider’s motivation, external pressure and whether the sensitive data is at risk or not. If the sensitive data is at risk, the motivation of a malicious insider is higher and accordingly the precondition for data theft. However there is no correlation between external pressure and sensitive data at risk. It can be explained by the fact that if the criminals threaten to hurt an insider’s family member it does not matter for the insider whether the data is easily available or not. He or she will do whatever it takes to protect his or her family.

Attempt of data theft depends on “precondition for data theft” and “policy violations and other inappropriate behavior”. Logic in here is that insider must break some established rules or behave inappropriately in order to maliciously obtain data. Quite often insider malicious activity could have been prevented by timely and effective action to address the anger, pain, anxiety, or psychological impairment of perpetrators, who exhibited signs of vulnerability or risk well in advance of the crime (Probst, Hunker, & Gollmann, 2010). Nevertheless some people are too cautious and imperturbable and that is why may not show any signs of malicious activity.

Unfortunately, users have plenty of choices when it comes to media and devices fit for secreting data and removing it from their employer. For this reason, it is generally easier to control data at the source with access control mechanisms than it is to block a virtually limitless array of potential destinations. Lack of access control in an organization, especially in a big one, is a serious breach of security. Absence of access control to sensitive information gives an opportunity to get access to any information available on the company's servers. It makes this kind of control most important in data security perspective but implementing of this control is not a solution per se. An insider can install credential-gathering malware in order to bypass this kind of control. Stolen credentials give an insider an opportunity to disguise himself as another employee and get access to sensitive information. This allows a malicious insider feel more "comfortable" because authenticated activity is much less likely to be noticed by detection mechanisms. Nevertheless this disguised activity can be detected by the node "insider's activity is logged and analyzed on a timely basis". This node is intended to detect changes in user activity template. For example, user activity is considered normal when a user accesses, for example, up to ten sensitive documents during working day and not more than five documents within an hour. When the user activity log shows that the number of accessed documents is above 10 during the working day or more than 5 within one hour it can be a sign that user account has been compromised. Activity logs can also provide information (for example IP address) about the source from where the document was accessed and help to find out who is responsible for excessive use of sensitive information. Use of shared passwords is also a security breach. The situation when, for example, two users share the same password to access sensitive data can be considered as minor violation of security policies but imagine the situation when hundreds of insiders use the same password. Lincoln National Corporation and TJ Maxx (discussed previously) are striking examples of such security negligence. The Credential capturing malware can be detected by security software if the latest is up-to-date. To bypass the security software mechanisms the insider can try to change security software settings by, for example, hacking. That is why security software should also be robust to such activities.

"Access control" and "software based controls" are parents of the node "security controls are OK". The latest can prevent most of the malicious activity caused by "regular" insiders. However this node is not very effective to capture malicious activity if insider is a privileged one. That why the node "advanced control of privileged user is implemented" was introduced in the given model. "Internal whistleblowing is encouraged" node provides additional security against data breaches.

This node is a “supplementary” and comes into effect when one or both other controls fail. When two other controls are set into the “true state” this node has no effect on prevention of data breaches.

8.2 Validation of the model

The model validity can be checked with the help of sensitivity analysis. AgenaRisk has a built-in function for such analysis. Sensitivity analysis is helpful for quick identifying and visualizing variables that have the most impact. This analysis plays an important role in model validation and is especially useful when historical data is lacking. It helps to verify whether the input nodes within the network have significant influence on the probabilities of hypothesis nodes.These nodes should correspond to the expert’s own beliefs regarding the significant influences within the environment. If they do not, then review of the model should be performed to check if probability adjustments are required. The sensitivity analysis was run several times in order to eliminate nodes that have no sufficient impact on the target node. The model was also redesigned several times in order to get the picture that corresponds with the reality because only those models that reflect the reality of the world should be accepted.

In order to run a sensitive analysis one should select a target node and one or more sensitivity nodes. The author choses to see the impact of nine input nodes on the target node – data breach. Input nodes “insider motivation to steal data” and “external pressure” were eliminated from the analysis. Both nodes are very influential because setting both of them into the “false state” changes the probability of “data attempt” to zero. These nodes have most impact on data breach because without “motivation” or “pressure” there will be no data breach at all. As the model is designed to capture only malicious activity the influential power of these nodes is quite understandable.

The objective of the analysis is to get a visual representation of the impact of different sensitivity nodes on the selected target node. The sensitivity analysis results provided in figure 15show that the model design corresponds with the description provided in the previous section. The length of the bars corresponding to each sensitivity node in the tornado graph is a measure of the impact of that node on the target node. The results show that the “implementation of access control to sensitive information” and “installing of security software” have more impact on data breach than other nodes. “Advanced control of privileged users” followed by “internal whistleblowing is encouraged” also reflects the reality of data breach problem in the malicious insider perspective. Even if security

software is installed and access control is implemented there are different ways to bypass them. Regular analysis of logged insiders' activity can prevent the malicious activity in its early phase. Implementation of measures to thwart stolen credential and elimination of use of shared passwords are nearly as important as previous node. This can be explained by the fact that implementing of access control is not a solution in its self. It should be supported by supplementary measures.

"Security software is robust and up-to-date" and "minor policy violations" are on the bottom of the tornado graph. These nodes are less influential then others but nevertheless are important part of the model. Verizon researches show that there is a correlation between "minor" policy violations and more serious abuse. That is why it is recommended that organizations be wary of and adequately respond to policy violations. Based on case data, the presence of illegal content, pornography, etc. on user systems (or other inappropriate behavior) is a reasonable indicator of a future breach (Verizon, 2010). Actively searching for such indicators rather than just handling those as they pop up may prove to be even more effective.

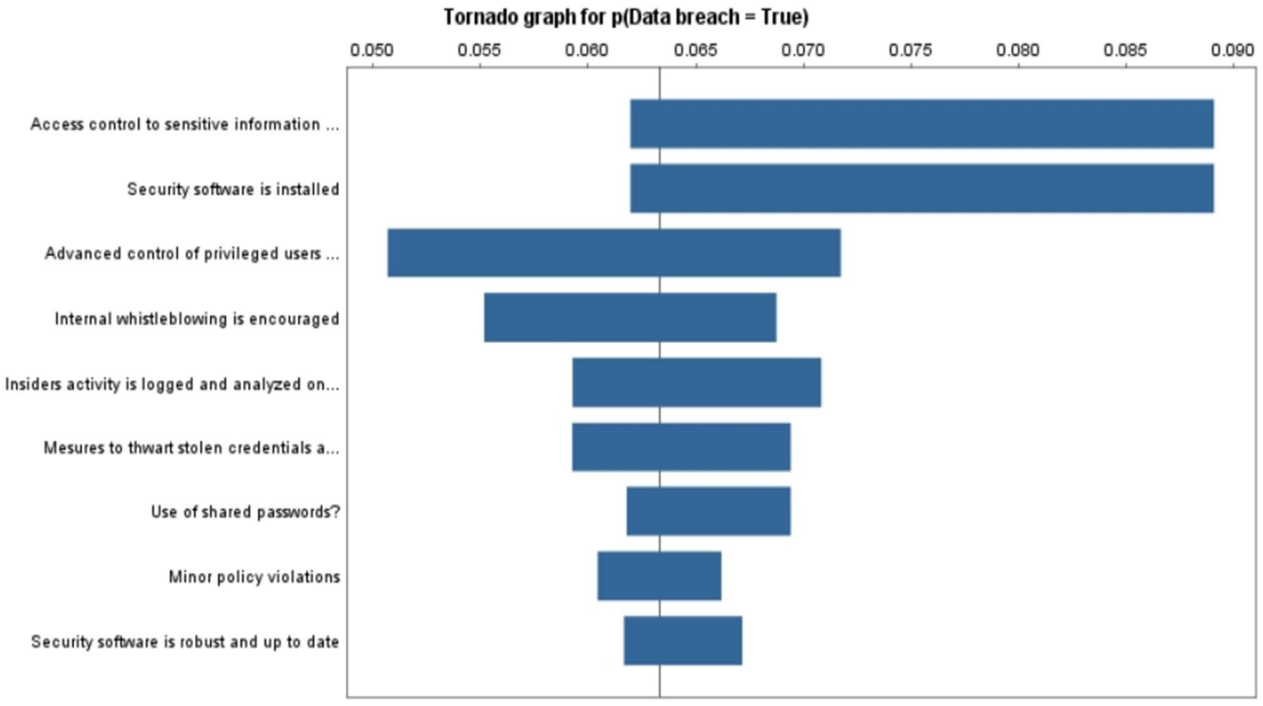

Figure 15 – Sensitivity analysis of the developed model

The results from sensitivity analysis can be used not only for pure validation purposes. They can help management to prioritize monitoring activities. Sensitivity analysis can also provide an assessment of which management options will have the greatest impact on the target nodes.

8.3 Running of scenarios

The model in its initial state is shown in figure 16. The model is based on data available from different data breach reports and it is quite understandable why the model in its initial state represents nearly the worst case scenario. If we set the insider motivation as “true” and “external pressure” as “false” we get the probability that data breach is “true” of56,6%. It makes sense because security controls presented in the initial model are quite weak. Setting the motivation as “false” and “external pressure” as true gives the probability of data breach of 72,9%. It also makes sense because when an insider is threatened and forced to steal data, the attempt of data theft does

not depend on whether the sensitive data is at risk or not.

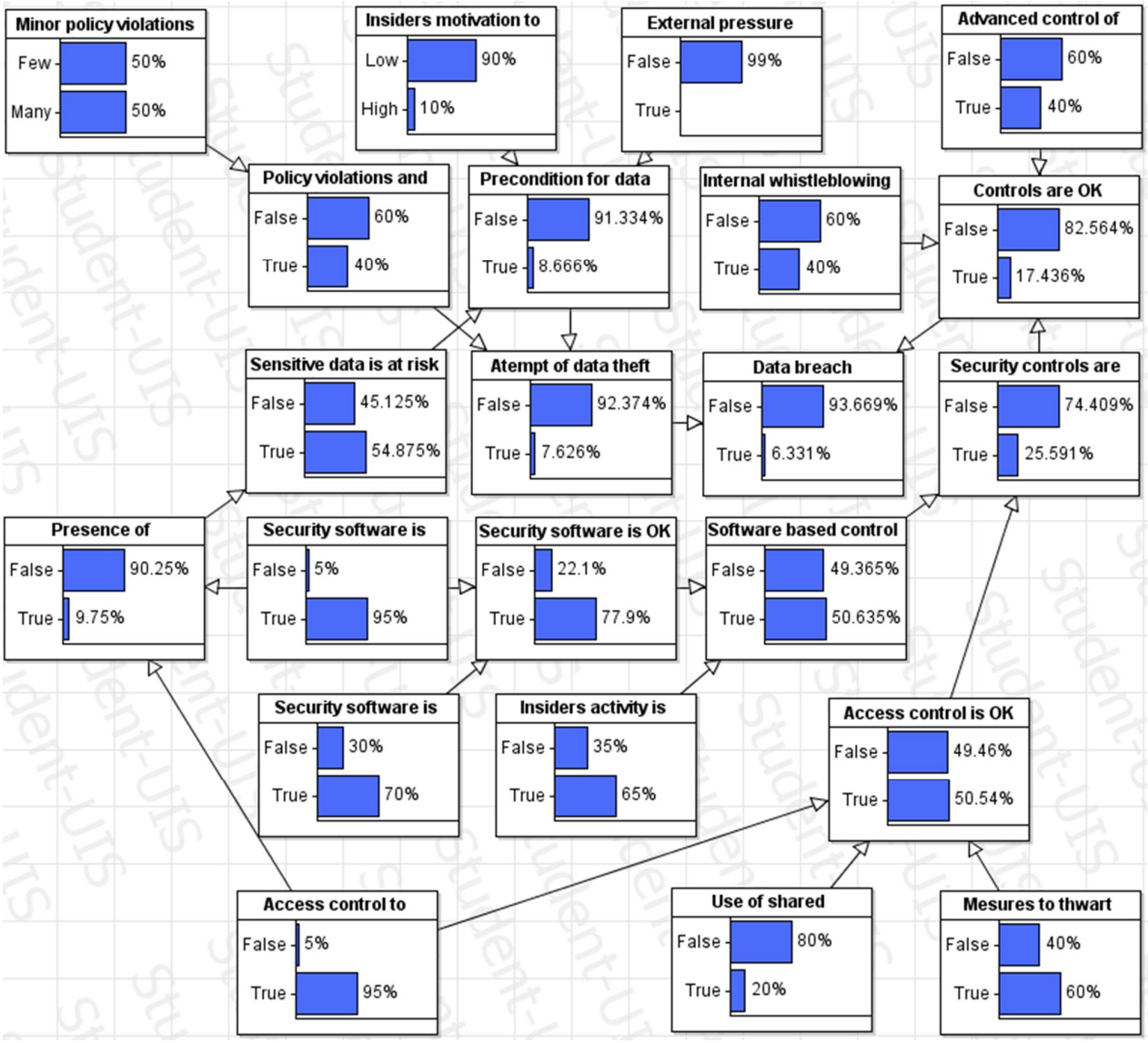

Figure 16 – Developed Bayesian Network with risk graphs on risk maps

The model ability to prevent data breaches can be illustrated through activating and combining of different sets of controls. Insider motivation is set as “true” and “external pressure” as “false”. All other input nodes are in their initial state. The results of the simulation can be found in table 4 below.

Table 4 – Simulation of different scenarios with the help of the developed model

Simulation	(A) Access control is OK	(B) Software based control is OK	(C) Whistleblowing is encouraged	(D) Advanced control of privileged insiders	PDB Probability of data breach
Initial state (0)					56,6%
1	X				49,1%
2	X	X			35,2%
3	X	X	X		28.1%
4	X	X	X	X	0.65%
5				X	45,3%
6	X			X	30%
7	X	X		X	0,65%
8	X		X	X	24%
9		X	X	X	24%
10			X		49,3%
11			X	X	36,2%

It can be observed from the table 4 above that none of the controls taken independently gives the total protection from the data breach. Only combination of controls gives sufficient data protection. We can see that simulation (4) and (7) give the same probability of 0,65%. It can be explained by the fact that control (C) has reasonably no effect when three other controls are enabled. It was mentioned previously in model description section that this control has a “supplementary” function and comes into effect when other controls fail. However, the effect of control (C) should not be underestimated. Taken independently (simulation 10) it can reduce the PDB by 6,3% (PDB(0) – PDB(10)= 6,3%).PDB(5) - PDB(11) = 45,3% - 36,2 % = 9,1 %. This reduction of PDB by 9,1 % is achieved by enabling of control (C). PDB(2) – PDB(3) = 35,2% - 28,1% = 7,1%. Here the control (C) has less effect. This can be explained by the fact that malicious activity of privileged insiders is more difficult to capture by this kind of control due to the fact that they are more “isolated” from the other employees due to their status. They usually have their own offices, that makes it more difficult to “monitor” their activity. Their status also prevents them from whistleblowing activity. It is more challenging for a regular employee to report on, for example, the line/department manager than on a

colleague who has the same position in the company. The author choses to presents the simulation only of the control nodes. These nodes are most influential in the presented model (if not taking "motivation" and "external pressure" into consideration) and that is why simulation of these nodes gives a clearer picture of how the model works.

9. Conclusions and future work

Data breach is a very comprehensive problem but the most important aspects of this problem were covered in this work. This master thesis gives overview of different threat agents and threat actions that companies should be aware of. The analysis of data breach investigations and theoretical sources confirms that insiders pose a great threat to organizations. Their accidental or malicious activity can lead to negative consequences for a company. Data breaches can result in huge financial losses, blemished reputation and loss of customers or/and market positions. The data breach caused by insiders is an operational issue that should be taken into consideration while managing the operational risk. Different researchers claim that BNs are potentially powerful tools for managing of operational risk. One of the objectives of this study has been to prove this opinion.

The modeling of data breach in general is a very comprehensive and time consuming task that is not practically feasible within the scope of a master thesis. In order to develop a model that reflects the reality at most it was decided to focus on data breaches caused by malicious insiders. The presented model was validated and different scenarios were run to show the model in work. The conducted study shows that BNs can be applied to modeling of data breaches caused by malicious insiders. The BN presented in this work includes the most common and the same time most critical factors that were discovered during analysis of different reports and theoretical sources. The designed BN is not intended to present an ultimate solution for elimination of insider problem. The model summarizes findings from different data breach investigations and intends to be used as an instruction to evaluate the most common and, as different researches show, most critical areas in malicious insider threat to data security perspective.

The developed model can be viewed as a basis for further work. The model can be used as a support for decision making in operational risk management perspective. The initial model generally reflects the state of a company that is supposed to encounter a data breach if there are insiders who are motivated to steal sensitive data or forced to do so. The management can evaluate whether the probabilities inbuilt into different input nodes reflect the reality of the company and eventually adjust them. Input nodes like, for example, “motivation” and different “control” nodes can be supplemented with parent nodes that are specific for a certain organization. According to the fact that there is not a lot of historical data available on the analyzed subject, the management can establish a working group of experts in order to finalize the model in accordance with the company’s specifics.

The finalized model is supposed to help management to better understand what problem areas are most relevant to their company and decide what improvements are worth doing. The model provides different options that can reduce the probability of data breach. The management can, for example, compare the total expected costs of each option against the total expected benefits, to see whether the benefits outweigh the costs. According to Verizon 2011 DRIB the 63% of investigated data breaches could have been prevented by implementing simple and cheap measures, 33% by intermediate preventing measures and only 4% of data breaches needed advanced and expensive preventive measures. Improving of critical areas can help the company in both lowering the probability of loss and decreasing the institution's capital requirements.

Bibliography

AlGhamdi, G., Wright, E., & Barbará, D. (2005). *Modeling Insider User Behavior Using Multi-Entity Bayesian Network.* Retrieved April 17, 2011, from George Mason University: u2.gmu.edu:8080/bitstream/1920/539/1/C4I-05-09.pdf

Allen, L. (2003, December). *The Basel Capital Accords and International Mortgage Markets: A Survey of the Literature .* Retrieved March 10, 2011, from New York University Stern: http://pages.stern.nyu.edu/~lallen/mortgage.paper.pdf

Altal Security. (2005). *International Security Standards*. Retrieved May 15, 2011, from Altal Information Security: http://www.altalsec.com/International_Security_Standards.php

Anil , P. (2005, December). *Basel II Focuses More on Operational Risk*. Retrieved May 17, 2011, from Network Magazine : http://www.networkmagazineindia.com/200512/inperson01.shtml

Basel Committee on Banking Supervision. (2003, July). *Risk Management Principles for Electronic Banking.* Retrieved March 28, 2011, from Bank for International Settlements: http://www.bis.org/publ/bcbs98.pdf

Basel Committee on Banking Supervision. (2010, April 16). *Strengthening the Resilience of the Banking Sector.* Retrieved May 17, 2011, from Bank for International Settlements: http://www.bis.org/publ/bcbs164.pdf

Ben-Gal, I. (2008). *Bayesian Networks. Encyclopedia of Statistics in Quality and Reliability.* John Willey & Sons.

Center for Strategic and International Studies. (2009, August 10). *Publications.* Retrieved April 2011, from Center for Strategic and International Studies: http://csis.org/files/publication/Twenty_Critical_Controls_for_Effective_Cyber_Defense_CAG.pdf

CFO Innovation Asia Staff. (2011, February 2). *Financial Services Institutions to Increase IT Spending*. Retrieved March 28, 2011, from Strategic Intelligence for CFOs, Finance Directors, Controllers and Treasurers in Asia: http://www.cfoinnovation.com/content/financial-services-institutions-increase-it-spending

Chartis Research. (2010, July). *Operational Risk & GRC Software Solutions 2010*. Retrieved April 18, 2011, from ChartisResearch: http://www.chartis-research.com/research/reports/operational-risk-grc-software-solutions-2010

Chernobai, & Rachev. (2007). *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Entrust. (2004, September). *Protecting Your Most Important Asset: Information .* Retrieved April 27, 2011, from Entrust: download.entrust.com/resources/download.cfm/21157/

Financial Fraud Law. (2009, November 16). *Expert Shirley Inscoe Cites "Billions Of Dollars" Disappearing Every Year Due To Internal Bank Fraud*. Retrieved March 16, 2011, from Financial Fraud Law.

Financial Industry Regulatory Authority. (2011, February 17). *News Release*. Retrieved May 29, 2011, from FINRA: http://www.finra.org/Newsroom/NewsReleases/2011/P122940

F-Secure. (2011). *Virus Encyclopedia.* Retrieved May 15, 2011, from F-Secure: http://www.f-secure.com/en_EMEA-Labs/virus-encyclopedia/encyclopedia/backdoor.html

Goodin, D. (2008, May 23). *TJX employee fired for exposing shoddy security practices.* Retrieved May 29, 2011, from The Register: http://www.theregister.co.uk/2008/05/23/tjx_fires_whistleblower/

Gregoriou, G. (2009). *Operational Risk toward Basel III: Best Practices and Issues in Modeling, Management and Regulation.* New Jersey: John Wiley & Sons.

Hadnagy, C. (2011). *Social Engenieering: The Art of Human Hacking.* Indianapolis: Wiley.

Inscoe, S., & Krishna, B. (2009). *Insidious: How Trusted Employees Steal Millions and Why It's So Hard for Banks to Stop Them.* Memento Press.

Israeli Software. (2011, November 12). *The psychology of data security*. Retrieved May 15, 2011, from Israeli Software: http://www.software.co.il/wordpress/2010/11/the-psychology-of-data-security

Jeghler, J. (2008, November 4). *Internal Fraud: Big Brother Needs New Glasses*. Retrieved March 15, 2011, from Celent: http://reports.celent.com/node/27080

Jensen, F., & Nielsen, T. (2007). *Bayesian Networks and Decision Graphs (Second Edition).* New York: Springer.

Johannes, B., Eaten, M., & Wu, Y. (2008). *ITU Study on the Financial Aspects of Network Security: Malware and Spam.* Geneva: International Communication Unity.

Jones, R. (2009, June 13). *CA: Man Fueled ID Theft Scheme With Dumpster Diving*. Retrieved May 17, 2011, from SecLists.Org: http://seclists.org/dataloss/2009/q3/5

Mackey , R. (2008, February 18). *Basel II's impact on information security*. Retrieved May 17, 2011, from Financial Service Information Security : http://searchfinancialsecurity.techtarget.com/tip/Basel-IIs-impact-on-information-security

Malcolm, A. (2006, June). *Social Engineering: A Means to Violate a Computer System.* Retrieved May 14, 2011, from SANS Institute Reading Room: http://www.sans.org/reading_room/whitepapers/engineering/social-engineering-means-violate-computer-system_529

McAfee/SAIC. (2011, March). *Underground Economies: Intellectual Capital and Sensitive Corporate Data Now the Latest Cybercrime Currency.* Retrieved May 28, 2011, from McAfee: http://www.mcafee.com/us/resources/reports/rp-underground-economies.pdf

MetricStream. (2011). *A Roadmap to the Advanced Measurement Approach (AMA) and Better Business Performance.* Retrieved April 18, 2011, from MetricStream Inc: http://www.metricstream.com/pdf/ORM_Solution_brief.pdf

Moosa, I. (2008). *Quantification of Operational Risk under Basel II: The Good, Bad and Ugly.* Hampshire: Palgrave Macmillan.

Oesterreichische Nationalbank. (2006, August). *Guidelines on Operational Risk Management.* Retrieved April 2011, from Oesterreichische Nationalbank: http://www.oenb.at/en/img/operational_risk_screen_tcm16-49652.pdf

Office of Inadequate Security. (2010, July 21). *Lincoln National Life Insurance notifies over 26,000 of breach after user/pass distributed in brochure and on the web*. Retrieved May 17, 2011, from DataBreaches.Net: http://www.databreaches.net/?p=12582

Office of Inadequate Security. (2011, May 4). *Congress not happy with Sony, Sony not happy with Anonymous, and gamers just unhappy, period.* Retrieved May 17, 2011, from DataBreaches.net: http://www.databreaches.net/?p=18114

Oracle. (2009). *Tutorial on Defending Against SQL Injection Attacks.* Retrieved May 15, 2011, from Oracle: http://download.oracle.com/oll/tutorials/SQLInjection/index.htm

Pourret, O., Naim, P., & Marcot, B. (2008). *Bayesian Networks: A Practical Guide for Applications.* West Sussex: John Willey & Sons.

Probst, C., Hunker, J., & Gollmann, D. (2010). *Insider Threats in Cyber Security.* New York: Springer.

RatingsDirect on the Global Credit Portal . (2010, November 17). *Tougher Capital Requirements Under Basel III Could Raise The Costs Of Securitization.* Retrieved May 17, 2011, from Standard & Poor's: http://www2.standardandpoors.com/spf/pdf/media/TougherCapitalRequirementsUnderBaselIIICouldRaiseTheCostsOfSecuritization.pdf

Referense for Business. (2011). *Computer Crimes*. Retrieved May 15, 2011, from Referense for Business: Encyclopedia of Small Business: http://www.referenceforbusiness.com/small/Co-Di/Computer-Crimes.html

Salomon, D. (2006). *Foundations of Computer Security.* London: Springer.

SpamLaws. (2009). *What is Data Security?* Retrieved May 15, 2011, from SpamLaws: http://www.spamlaws.com/data-security.html

Symantec. (2009, October 23). *Anatomy of a Data Breach. Why Breaches Happen and What to Do About It.* Retrieved May 2, 2011, from Symantec: http://eval.symantec.com/mktginfo/enterprise/white_papers/b-anatomy_of_a_data_breach_WP_20049424-1.en-us.pdf

Thirlwell, J. (2010, November). *Basel III and operational risk: the missing piece?* Retrieved April 14, 2011, from Authority on Operational Risk and Business Risk: http://www.johnthirlwell.co.uk/FS_Focus.pdf

U.S. Secret Service & CERT Coordination Center. (2004, August). *Insider Threat Study: Illicit Cyber Activity in the Banking and Finance Sector.* Retrieved April 15, 2011, from CERT: www.cert.org/archive/pdf/bankfin040820.pdf

US Government Interagency Working Group. (2000, December 1). *International Crime Threat Assessment.* Retrieved March 25, 2011, from Federation of American Scientists: http://www.fas.org/irp/threat/pub45270index.html

Van Luvender, R. (2011, April). *Fraud Trends in 2010: Top Threats From a Growing Underground Economy.* Retrieved Mars 2011, from Bank Systems & Technology: http://www.banktech.com/whitepaper/Risk-Management-Security/Fraud/fraud-trends-in-2010-top-threats-from-a-growing-wp1271273655734

Verizon. (2008). *2008 Data Breach Investigations Report.* Retrieved Mars 2011, from Verizon: http://www.verizonbusiness.com/resources/security/databreachreport.pdf

Verizon. (2010). *2010 Data Breach Investigations Report.* Retrieved February 2011, from Verizon: http://www.verizonbusiness.com/go/2010databreachreport/

Verizon. (2011, May). *2011 Data Breach Investigation Report.* Retrieved May 7, 2011, from Verizon: http://www.verizonbusiness.com/go/2011dbir

Widup, S. (2010). *The Leaking Vault - Five Years of Data Breaches.* Digital Forensics Association .