OPTIMAL INFORMATION SECURITY EXPENDITURES CONSIDERING BUDGET CONSTRAINTS

Andreas Schilling, Faculty of Management and Economics, Ruhr University Bochum, Germany, andreas.schilling@rub.de

Brigitte Werners, Faculty of Management and Economics, Ruhr University Bochum, Germany, or@rub.de

Abstract

In this paper, we present a new quantitative optimization model to support decision makers in determining how much to invest in information security and how to allocate funds. The approach considers uncertain properties of security risks and provides concrete investment recommendations. Evaluating the problem in a holistic way improves insight into the problem structure and leads to better decision making. By using methods of mathematical optimization, available budget can be utilized most effectively. An exemplary case study demonstrates how the approach is applied to increase security of a cloud-based information system. To test our model, we use very detailed as well as vague input data. In both cases, good results are produced which can be the basis for further decision making. The approach is designed to be used within the framework of an existing risk management process.

Keywords: Security Optimization, Information Systems Security Quantification, Security Economics, Risk Management, Decision Making.

1 Introduction

Due to very high and continually increasing reliance on information systems in business environments, the importance of assessing information security risks is particularly high. To assess and increase the security of such systems, most organizations apply different types of risk management. There are several approaches that are used in practice and there is no globally accepted standard. In general, to manage risk, it is necessary to identify individual threats and their potential impacts. Risks resulting from these threats have to be measured appropriately to prioritize counter measures.

To bring information security risks down to an acceptable level, risk management usually encompasses three sub-processes. First, a risk assessment is conducted to identify threats and their impacts. Based on the findings, the system is inspected to determine possible vulnerabilities related to these threats. To reduce risk, possible counter measures/security controls are proposed which reduce the exploitability of specific vulnerabilities. This sub-process is usually conducted by a team of experts who are familiar with information security in general and the system under consideration. The next step is risk mitigation which involves prioritizing between controls, selecting the most effective ones, and implementing them. Due to the fact that the elimination of all risks is naturally impossible, it is the obligation of an organization to select the most effective controls while respecting budget limitations. When controls are put in place, a continual evaluation phase follows to measure the effectiveness of deployed controls (Aagedal et al. 2002; Stoneburner et al. 2002).

The complexity of this process makes effective risk management a challenging task. It is characterized by considerable system complexity and a high degree of uncertainty. To improve existing procedures, we propose to use an integrated decision model to evaluate and increase security in a holistic way by taking several kinds of interdependencies into account. The model combines quantitative modeling and mathematical optimization to select security controls in order to minimize expected loss with respect to the complete threat landscape. The mathematical foundation of our model is designed to process fine grained data as well as more aggregated information to determine an optimal combination of security controls and damage estimations. Although the accuracy of the model is influenced by the quality of the input data, we found evidence that it also provides useful results in case of vague input information.

The remainder of this paper is structured as follows: We begin by reviewing several approaches and models that aim at quantifying information security. The following section introduces a model-based approach to security quantification. This includes analyzing the structure of the security quantification problem and identifying uncertain properties. Based hereon, a deterministic optimization model is established to provide concrete decision support. Subsequently, we thoroughly review results of the model in an exemplary case study and demonstrate how it can be used as a tool to effectively increase information security. Finally, we conclude this paper and give possible directions for future research.

2 Related work

In recent years, the interest in quantitative models to improve information security investment decisions has increased significantly. Several research streams have addressed this problem from different angles. There are basically two distinct subproblems where each subproblem is focused on one key issue. First, the right budget has to be determined, and second, the best controls within this budget have to be selected. Particularly the second problem is still not sufficiently solved. Therefore, in this paper, we propose a model to optimally solve the security control selection problem.

In general, activities concerning these problems are carried out in a risk management process of an organization. Aagedal et al. (2002) conducted a detailed analysis of the risk management process and identified key aspects of its sub-processes. As a result of their analysis, they proposed several qualitative models to carry out a risk assessment. Each model is designed to support a specific phase of the assessment and takes into account that information about the system has to be acquired by eliciting expert judgment (Ryan et al. 2012). The models enable a structured view of the risk management process but only provide

qualitative results. Hua et al. (2011) state that the objective of information security is to minimize potential losses while simultaneously considering the cost of security solutions. According to them, it is very likely that non-optimal investment decisions are made if no proper risk analysis is conducted prior to the selection of security controls.

2.1 Determining an optimal budget

Gordon et al. (2002) provide first results on the economics of information security and give recommendations on how much to invest in security. They developed a model that takes the structure of information security into account and provides information on expected losses due to security incidents.
They state that the general vulnerability of a system depends on the invested amount in security. More specifically, the exploitability of a vulnerability is defined as a function depending on the amount invested in security technology. Their approach focuses on the application of different security breach functions $S(z, v)$ to calculate a concrete amount z to invest to mitigate a specific vulnerability v. The general idea is that higher investments reduce a vulnerability and yield lower expected losses. The results give valuable insights into the economics of security investments. Although it is possible to obtain concrete amounts to invest, the model gives no recommendation to the decision maker on *how* to invest. Therefore, in our model, security investments are considered to be discrete investments. The primary problem here is not to determine the amount to invest but to select the actual controls.
Hausken (2006) examined four additional types of security breach functions with different shapes and found that different functions should be applied in different situations. Wang et al. (2008) proposed a more detailed analysis which makes use of security incident data and statistical methods like the concept of value-at-risk to support a decision. Tsiakis (2010) also analyzes the problem of determining appropriate security expenditures and emphasizes the importance of security measurement. Although qualitative approaches can contribute to this evaluation, they are not sufficient to perform a solid cost/benefit analysis. A quantitative approach, on the other hand, can provide concrete measures and gives a clear indication to the decision maker.

2.2 Management tools and financial measures

As basis for an optimization approach, financial measures and key economic indicators can be used to guide the decision. Many approaches and models already applied such measures, although mostly not to determine an optimal solution but to measure the performance of a given one. Bojanc et al. (2012) developed a modeling approach which uses quantitative measures to describe specific model components. On the basis of these measures, a quantitative risk metric is introduced that can be used to determine the expected damage to the system. In addition, this value is used to calculate different key economic indicators such as return on investment (ROI) and net present value (NPV) (Bojanc et al. 2008; Bojanc et al. 2012). This approach gives insights into risk management but does not provide an integrated method to reduce risk. Furthermore, to support the understanding of the economics of security investments, another quantitative model has been proposed by Aissa et al. (2010). The model is designed to calculate the mean failure cost by taking into account requirements of different stakeholders, multiple system components, and a threat vector (Aissa et al. 2010; Rabai et al. 2013). However, the model only evaluates the current state of the system and provides no further decision support. Therefore, it is not possible to derive any information on how to improve security.
Sonnenreich et al. (2006) are proposing a quantitative approach which utilizes the return on security investment (ROSI) to take economic aspects into consideration. The same idea is also shared by Böhme et al. (2008). ROSI, initially proposed by Berinato (2002), can be used to calculate a ROI for security investments. It takes the specifics of information security into account since security investments do not yield a concrete return. The return has to be determined by predicting potentially reduced losses caused by deploying security controls. To predict these losses, a quantitative model is required.

Another general class of approaches falls into the field of microeconomics and uses game theoretic models to treat security as a game between an organization and an attacker to determine an ideal investment level (Cavusoglu et al. 2004; Cavusoglu et al. 2008; Gal-Or et al. 2005; Roy et al. 2010).
In this paper, we establish a new model that is based on the principles of previous work on security quantification and additionally uses combinatorial optimization to find an optimal investment strategy. The benefit of an optimization approach is that it enables the decision maker to find the best possible solution while respecting different properties and restrictions at the same time. Our model can be integrated into an existing risk management process and may be used in conjunction with established approaches. For example, it is possible to estimate a security budget as input to our model using the method developed by Gordon et al. (2002). In addition, results may be further analyzed using different key economic indicators as demonstrated by Berinato (2002), Böhme et al. (2008), and Bojanc et al. (2012).

3 Model-based security quantification

The quantification of information security is part of an organizational risk management process. To quantify security we have to take a close look on each threat and its corresponding impact on the system. Each threat in combination with some impact constitutes an information security risk. Since impacts are usually not positive in information security, a high level definition describes risk as uncertainty of an event that may lead to a negative outcome (Hogganvik et al. 2006; Kaplan et al. 1981; McNeil et al. 2005). This understanding implies that risk can be measured, for instance in monetary terms. An uncertain negative event happens with its probability of occurrence and it is common practice to quantify a specific risk by multiplying the probability of the threat in question by its estimated financial impact (Aven 2011; Gordon et al. 2002). The result is the expected loss of the risk.
There exists a considerable market for security solutions which is why the problem is not finding security controls, but finding the right ones within a given budget. Possible controls have a wide variety of capabilities to protect against different threats and vulnerabilities and, additionally, they might complement each other (Schilling et al. 2014) leading to a situation where only a rational methodology and quantitative analysis is able to help select the best configuration of controls. To prioritize between controls it is necessary to measure their impact on security. Since this has to be done before controls are deployed, it is not possible to verify their effectiveness in a production environment. In order to solve this, we propose to use a quantitative optimization model enabling the decision maker to maximize the effect of deployed security controls before they are established. The model only represents relevant parts of the real system to allow a practical applicability.

3.1 Structural elements of the model

To design an optimization model suitable to describe information security, the essential parts defining security have to be identified. Based on a literature review of other information security models (Aagedal et al. 2002; Bojanc et al. 2008; Bojanc et al. 2012; Cavusoglu et al. 2004; Gordon et al. 2002; Schilling et al. 2013; Sun et al. 2006; The Open Web Application Security Project 2013), we identified a common set of components. Although each model uses a different approach, all share a very similar understanding of the overall problem structure.
It is common notion that the sources of security risks are threats. A threat is a specific incident that causes damage to an asset. The incident is caused by different threat agents including hackers, competitors, administrative staff, or malicious insiders. Assets typically include technical components like hardware, software, and data. To attack an asset, a threat agent can exploit one of many possible vulnerabilities. A vulnerability is a technical or organizational shortcoming of the system. This may be an insecure application programming interface (API) endpoint (Mather et al. 2009), software backdoors (Schuster et al. 2013), or even lack of, or insufficient, rules for employees (Adams et al. 1999). To prevent an attacker from taking advantage of such vulnerability, an organization can deploy security controls which affect

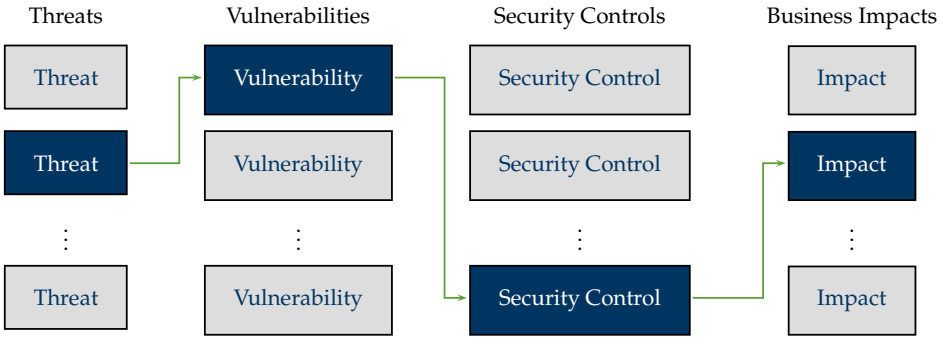

Figure 1. The severity of information security risks is depending on multiple components that are connected in a path-like structure through the system. (Adapted from The Open Web Application Security Project (2013)).

the exploitability of vulnerabilities. Concrete instances of threats, assets, vulnerabilities, and security controls are determined in a risk assessment. When this information is available, the decision maker is faced with the problem of choosing the most effective security controls. The selection problem requires a coherent method for evaluating risk and is constrained by budget limitations. Figure 1 shows how these components are connected. A security incident arises in form of a path through the system, starting with a threat and resulting in a negative business impact.

In most cases a threat can exploit a number of vulnerabilities to cause damage. For each threat a set of potential vulnerabilities can be identified. Since not all vulnerabilities are equally likely to be exploited by different threats, there exists an interrelationship subject to some degree of uncertainty. The same applies to vulnerabilities and security controls. Most security controls do not only affect a single vulnerability, but multiple ones. Therefore, the interrelation of vulnerabilities and controls leads to a situation, where different combinations of controls can have a completely different effect on security. This characteristic has to be considered to maximize the effect of controls. If the selection of controls is not performed in an integrated way considering all vulnerabilities at the same time, possible positive interdependencies of controls remain unrealized.

In practice, controls are often selected by treating vulnerabilities successively and assigning controls that are best suited for the current control in question (Stoneburner et al. 2002). As a result, the selection of controls is often suboptimal or the budget cannot be met. To maximize the effectiveness of controls, the best combination of controls within the limit of available budget has to be chosen. Thus, to consider the whole system at once during the decision process, we propose to use an integrated optimization model.

3.2 Uncertain properties

The quantification of security is a problem with several degrees of uncertainty. It is uncertain whether threats are arising or if vulnerabilities can be exploited. In addition, a security control may fail to prevent an attack and if so the resulting damage may vary. The following formulation takes these uncertain components into account by using random variables to model uncertain parameters which have to be identified as part of the risk management process.

Let $k \in K$ be security controls and let $j \in J$ be vulnerabilities. Each security control k has an effect on vulnerability j. This effect can be approximated but is still subject to some degree of uncertainty. This is modeled by a random variable $C_{j,k}$ to take these uncertain events into consideration. In a real system, this variable can only take values 0 or 1. This is due to the fact that a control can only fail or succeed and anything in between is no tangible event. If the outcome is 0, security control k fails to prevent the exploitation of vulnerability j. In case the control is successful, the variable takes the value 1. It is crucial

to define this binary construct to obtain realistic values for random realizations. A Bernoulli distributed random variable fulfills this requirement:

$$C_{j,k} \sim B\left(1; p^C_{j,k}\right). \tag{1}$$

The superscript C indicates that probability parameter $p^C_{j,k}$ is associated with random variable $C_{j,k}$. We use this notation throughout this paper to denote that a parameter is either associated with a threat (T), vulnerability (V), or security control (C). To model the exploitability of vulnerabilities and the probability of threats, we define two additional random variables $V_{i,j}$ and T_i. Both are defined in the same way as $C_{j,k}$ using a Bernoulli distributed random variable. Let $V_{i,j}$ take value 1 if vulnerability j exploits threat i (0 if not), and T_i takes value 1 if threat i arises (0 if not).

$$V_{i,j} \sim B\left(1; p^V_{i,j}\right) \tag{2}$$

$$T_i \sim B\left(1; p^T_i\right) \tag{3}$$

From these definitions we can determine the probability that a threat arises and causes damage by calculating its overall probability of success with respect to all vulnerabilities and security controls. A threat only causes damage if it arises, at least one vulnerability is exploited, and all security controls fail. To determine which controls should be implemented, the overall damage to the system depending on the deployed selection of controls has to be determined. If all deployed controls that have an effect on vulnerability j fail, vulnerability j can be exploited. Therefore, probability $e^V_{i,j}$ that vulnerability j is exploited by threat i is the joint probability that the threat is trying to exploit the vulnerability and that all controls fail:

$$e^V_{i,j} = p^V_{i,j} \cdot \prod_{k \in K} \left(1 - p^C_{j,k}\right). \tag{4}$$

Furthermore, the probability δ_i that threat i causes damage is the joint probability that it arises and that at least one vulnerability is exploited. The probability of this event is

$$\delta_i = p^T_i \cdot \left(1 - \prod_{j \in J} \left(1 - e^V_{i,j}\right)\right). \tag{5}$$

The corresponding event is binary, meaning that a threat either causes damage or not. To define this property we could again use a Bernoulli distributed random variable, however, threats may arise multiple times during a specified time period, say one year. Let n_i be the expected number of occurrences of threat i within this time period. To consider this, we introduce a Binomial distributed random variable R_i expressing the annual rate of occurrence (ARO) of threat i:

$$R_i \sim B\left(n_i; \delta_i\right). \tag{6}$$

The fact that R_i is Binomial distributed reflects that each time out of n_i times a threat arises, it causes damage with probability δ_i. This means R_i describes a sequence of Bernoulli distributed random events where each time a threat may cause damage or not. To quantify damage to the system, the outcome of R_i is multiplied by the single loss ℓ_i of the threat. The result is the expected annual loss $a\ell_i$. The total expected loss within one year is the sum of losses of all threats.

3.3 Optimizing the selection of security controls

To find an optimal selection of security controls we use an optimization model which is set up in replacing random variables by their respective expected values. The resulting optimization model makes it possible to

minimize the total expected loss depending on different control configurations according to the introduced structure. The minimization is carried out using a successive linear approximation technique (FICO 2015) to find the best possible selection of security controls. To apply this technique we establish a mathematical formulation of the problem providing a quantifiable objective (total expected loss) and restricting possible control configurations by introducing a budget constraint.

Indices and sets	
I	Index set of threats (indexed by i)
J	Index set of vulnerabilities (indexed by j)
K	Index set of security controls (indexed by k)
Parameters	
B	Budget for security controls
c_k	Costs for security control k
ℓ_i	Single loss caused by threat i
n_i	Number of occurrences of threat i within one year
p_i^T	Expected probability that threat i arises $\left(p_i^T = \mathrm{E}\left[T_i\right], \forall i \in I\right)$
$p_{i,j}^V$	Expected probability that threat i is exploiting vulnerability j $\left(p_{i,j}^V = \mathrm{E}\left[V_{i,j}\right], \forall i \in I, j \in J\right)$
$p_{j,k}^C$	Expected probability that security control k prevents a threat from exploiting vulnerability j $\left(p_{j,k}^C = \mathrm{E}\left[C_{j,k}\right], \forall j \in J, k \in K\right)$
Decision variables	
$a\ell_i$	Annual loss expectancy of threat i, $a\ell_i \geq 0$
$e_{i,j}^V$	Expected probability that vulnerability j is exploited by threat i, $e_{i,j}^V \in [0,1]$
δ_i	Expected probability that threat i successfully occurs, $\delta_i \in [0,1]$
sc_k	Selection of security control k to be established, $sc_k \in \{0,1\}$

Model formulation:

$$\min \quad \sum_{i \in I} a\ell_i \tag{7}$$

subject to:

$$a\ell_i = \ell_i \cdot \delta_i \cdot n_i \qquad \forall i \in I \tag{8}$$

$$\delta_i = p_i^T \cdot \left(1 - \prod_{j \in J} \left(1 - e_{i,j}^V\right)\right) \qquad \forall i \in I \tag{9}$$

$$e_{i,j}^V = p_{i,j}^V \cdot \prod_{k \in K} \left(1 - p_{j,k}^C \cdot sc_k\right) \qquad \forall i \in I,\ j \in J \tag{10}$$

$$\sum_{k \in K} c_k \cdot sc_k \leq B \tag{11}$$

$$sc_k \in \{0,1\} \qquad \forall k \in K \tag{12}$$

The objective function (7) minimizes the total expected loss. This is the sum of all expected annual losses $a\ell_i$ caused by threats $i \in I$. The expected annual loss is calculated for all threats in constraint (8) by multiplying single loss ℓ_i by its estimated rate of occurrence n_i and probability δ_i. δ_i is the probability that threat i successfully occurs (9) and is based on the probability $e_{i,j}^V$ that at least one vulnerability is exploited. The value of $e_{i,j}^V$ depends on the selection of security controls and is calculated in constraint (10). To determine whether a security control is selected, the binary decision variable sc_k is introduced. If

$sc_k = 1$, control k is selected and if $sc_k = 0$ it is not. Due to the fact that the selection of sc_k influences the expected loss, the best possible selection is determined when solving the model. In constraint (11) costs are limited to a specified budget B. Constraint (12) defines sc_k to be binary.
In addition to finding the optimal selection of controls for a given budget, the resulting optimization model can be used to determine how the system reacts when fixing or varying certain input parameters. This means, for instance, calculating several optimal solutions for varying budgets or testing a specific configuration of controls by setting the corresponding sc_k values to 1. The decision maker receives fine grained information on the current state of different system parts. This includes, among other things, how much loss can be expected and which parts of the system are most vulnerable (Schilling et al. 2013). How accurate this information is depends on the input information and may reach from a good prediction to a rough estimate.

4 Case study and computational results

The advantages of the proposed optimization model are shown in a realistic case study of an exemplary information system. The model is implemented using the standard optimization software Xpress Optimization Suite which allows modeling and solving of complex and large nonlinear problems. The problem is solved applying successive linear programming provided by the Xpress-SLP solver `mmxslp` (FICO 2015).

4.1 Application scenario

In the following, we analyze an exemplary cloud-based system to demonstrate how the model can be applied in a realistic setting. The data for the case study are based on the authors' practical experience in developing and maintaining real-world information systems. The system in question mainly provides an information based service to end users. All data are stored in a dedicated database. The service is provided over the Internet and can be accessed by end users with a web browser. The system is created on top of an existing cloud platform and hosted by an external cloud service provider (CSP). A third-party API is used to obtain information needed for the provisioning of the core service. In addition, several other CSPs are integrated to add additional functionality (e.g., sending transactional emails, application performance measurement, payment processing, etc.).
Threats can be very technical like cross-site scripting (Heiderich et al. 2012) or arise from an organizational level like social engineering attacks (Luo et al. 2011). Our approach enables decision makers to consider different kinds of threats simultaneously in one unified model. The same applies to vulnerabilities which may be identified deeply in the source code of an application or at a very high level of an organization. To reduce the exploitability of vulnerabilities, the decision maker can suggest any number of security controls. The more alternative controls are proposed, the better available budget is utilized.
For this case study of a cloud-based system we identified as most important 20 threats, 18 vulnerabilities, and 25 security controls. Financial losses caused by successful threats range between 220 € and 8,900 €. Threats are estimated to occur between 5 and 90 times within one year. Costs for security controls start at 2,900 € and go up to 50,200 €. Solving an instance of this problem takes between 0.12 and 1.02 seconds using the standard solver Xpress-SLP. Solution time is expected to increase for bigger problems and for very large applications it may be useful to improve performance by using a custom heuristic approach (e.g., tabu search, genetic algorithms).

4.2 Optimal Selection of Security Controls

From a business perspective, the main problem with security investments is the absence of a real return on investment. If security expenditures are increased, the number of incidents is (usually) reduced. As a consequence, it becomes more difficult to measure the effectiveness of deployed security controls. The

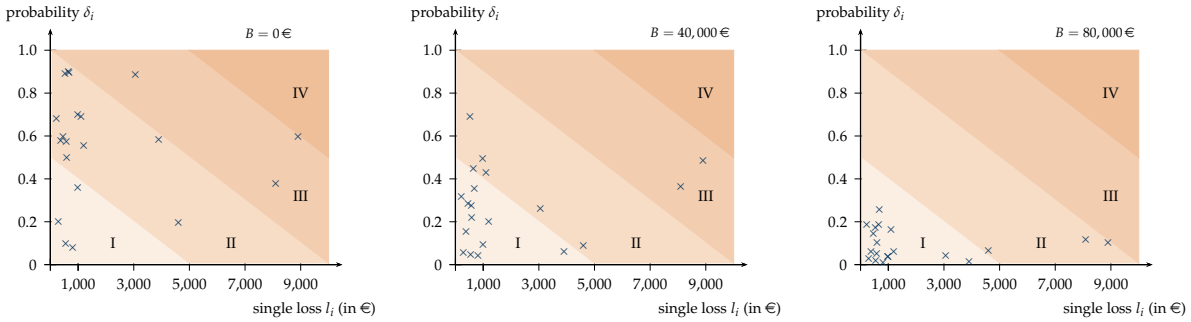

Figure 2. Severity of threats according to different budgets visualized in three risk matrices with four risk levels.

only real solution to this problem is to try to predict what may happen and measure the consequences of controls before they are deployed. Our model enables the decision maker to do so.
Based on the introduced case study, we first use the model to calculate how much loss $\overline{L}$ can be expected if no security controls are deployed (worst-case scenario). The severity of each threat is shown in the first risk matrix in Figure 2. The four risk levels indicate within the range of existing threats, how severe a threat is. In this case the risk matrix gives no absolute indication on the criticality of a threat to the system but how it compares in relation to other threats. The risk levels are evenly distributed in the range of single losses. The maximum value is set slightly above the highest single loss to ensure best readability. Under the assumption that a threat always causes the same loss if it occurs, its expected loss only decreases if its probability δ_i is reduced. Therefore, a threat can only move vertically downwards in the matrix.
The second risk matrix in Figure 2 reveals how the distribution of threats changes if we deploy security controls according to the optimal solution given in Table 1 This solution is obtained by solving the model to optimality with a budget of $B = 40,000$ €. As can be seen in the third risk matrix, the probability of threats is significantly reduced when choosing a budget of 80,000 €. There are no more threats located in level III or IV and the majority is now located in level I.

Expected loss $\overline{L}$ without investment in security:		1,197,118 €
Optimal solution:	$sc_k^{opt} = (\underline{0},0,0,0,1,0,0,0,0,\underline{1},\underline{0},1,0,0,0,1,0,0,0,0,0,0,0,0,0)$	
Expected loss L^{opt} according to optimal solution:		634,608 €
Costs C:		38,100 €
Gross loss reduction:	$(\overline{L} - L^{opt})$	562,510 €
Net benefit:	$(\overline{L} - L^{opt} - C)$	524,410 €

Table 1. Optimal selection of security controls and corresponding expected loss for a fixed budget of $B = 40,000$ €.

Table 2 represents the optimal solution with an increased budget of $B = 80,000$ €. It is evident that, in comparison to the solution given in Table 1, the expected loss is further reduced. This is achieved by deploying one additional security control ($sc_1 = 1$) and by switching from control 10 to 11 ($sc_{10} = 1, sc_{11} = 0 \rightarrow sc_{10} = 0, sc_{11} = 1$). The previous solution depicted in Table 1 resulted in a gross loss reduction of 562,510 € (from 1,197,118 € down to 634,608 €) which is now increased to a total loss reduction of 1,014,093 € (from 1,197,118 € down to 183,025 €). Although the expected loss is further reduced by doubling the budget, the reduction only corresponds to an additional 38 percentage points. In comparison, the expected loss according to the first solution was reduced by 47 percentage points. This observation leads to the assumption that losses decrease with increasing investments, but at a decreasing rate.

To explore which solution has the highest net benefit, the costs of security controls have to be considered as well. The net benefit is the difference between the gross loss reduction and resulting costs. The deployment of the first solution leads to a net benefit of $524,410 €$ which is increased to $939,392 €$ according to the second solution. The benefit rises because additional costs of controls are overcompensated by the realized loss reduction. However, at some point increasing control costs cannot be compensated by an additional loss reduction and security investments become pointless from an economic point of view.

Expected loss $\overline{L}$ without investment in security:		$1,197,118 €$
Optimal solution:	$sc_k^{opt} = (\underline{1},0,0,0,1,0,0,0,0,\underline{0},\underline{1},1,0,0,0,1,0,0,0,0,0,0,0,0,0)$	
Expected loss L^{opt} according to optimal solution:		$183,025 €$
Costs C:		$74,700 €$
Gross loss reduction:	$(\overline{L} - L^{opt})$	$1,014,093 €$
Net benefit:	$(\overline{L} - L^{opt} - C)$	$939,392 €$

Table 2. Optimal selection of security controls and corresponding expected loss for a fixed budget of $B = 80,000 €$.

The fact that expected losses are reduced with increasing investments in security is also clearly visible in Figure 3. The curves are generated by calculating multiple optimal solutions for a stepwise increasing budget B. The calculation of 40 solutions took 15.36 seconds. Each data point is obtained from the solution of the optimization model according to the budget displayed on the x-axis. The figure shows that the expected loss curve drops steadily, however, the slope continuously decreases. Whereas relatively small investments yield a significant reduction of loss, at some point additional investments cause only marginal improvements. If investments are very high, the expected loss is approaching zero, but a total reduction to zero loss cannot be achieved. Since the objective of the model is to minimize expected losses, costs are also increasing if the budget increases. To make optimal use of the available budget, total costs C are always slightly below the budget limit B. This fact results in an almost linear shape of the cost curve. If no further controls are available to be deployed, costs remain at a constant level. The same applies to the loss and benefit curves which depend on the configuration of deployed controls. Somewhere between zero and this constant level the benefit function takes its maximum. The maximum is reached right before additional costs for controls cannot be compensated with reduced losses. From an economic point of view, it is advisable to invest in security where the benefit is at its peak. Although, this seems to be the best option, it should be noted that this only applies to a risk neutral decision maker. It is important to realize that the extent to which losses are reduced and budget is made available depends on the decision maker and his risk preference.

4.3 Model Applicability Considering Inaccurate Input Parameters

So far we used the model as if every input parameter can be estimated accurately based on historical data or expert knowledge. In practice, we can observe a lack of extensive data regarding security incidents and it is reasonable to assume that in most cases parameters have to be estimated by a team of experts. Even though there are methods and studies available which demonstrate that expert judgment can be used to obtain model parameters in information security (Ryan et al. 2012) and other areas (e.g., project selection (Hassanzadeh et al. 2012) and valuation (Beer et al. 2013)), it remains a difficult task to obtain exact parameter values for real system properties. In addition, the values of such parameters are critical to the solution of the model. While our approach benefits from exact input data, it also produces valuable results and insights using imprecise input parameters. This is due to the fact that the structure defined by input data remains mostly intact even if several parameters have been estimated vaguely.

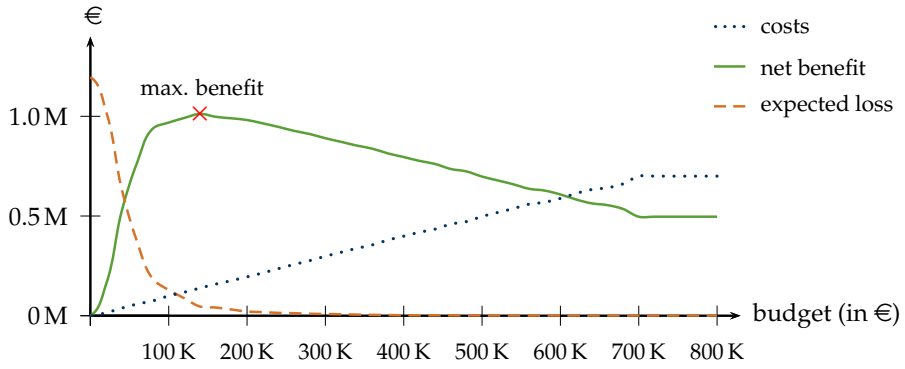

Figure 3. Expected loss is reduced with increasing budget, but at a decreasing rate.

B	Input	Optimal selection of controls	Matches
40,000 €	exact	$(0,0,0,0,\underline{1},0,0,0,0,1,\underline{0},1,0,0,0,1,0,0,0,0,0,0,0,0,0)$	23/25
	vague	$(0,0,0,0,\underline{0},0,0,0,0,1,\underline{1},1,0,0,0,1,0,0,0,0,0,0,0,0,0)$	
80,000 €	exact	$(1,0,0,0,1,0,0,0,0,0,1,1,0,0,0,1,0,0,0,0,0,0,0,0,0)$	25/25
	vague	$(1,0,0,0,1,0,0,0,0,0,1,1,0,0,0,1,0,0,0,0,0,0,0,0,0)$	

Table 3. Comparison of two optimal solutions for different budgets B with exact and vague input parameters.

To examine how the solution of the model is affected by parameter changes, we will use a more vague estimation scale to define model parameters. Instead of using exact numerical values, parameters are estimated according to five verbal probability levels: remote, rare, unlikely, possible, and probable. Each verbal level corresponds to a concrete probability value that guarantees a clear distinction between each level. To compare the results obtained with this valuation, the same data as in the previous example is used except that all probability parameters are condensed to the values of the five levels to be seen in Figure 4. For example, the value 0.24 is replaced by 0.3 which corresponds to the verbal level "rare".

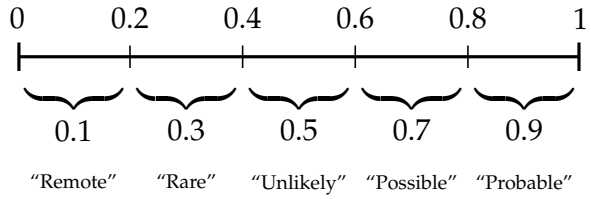

Figure 4. Probability parameter consolidation using five probability levels that correspond to a verbal estimation scale.

In Table 3 the optimal selection of controls of the two previously examined solutions are compared. The solutions were obtained by solving the model once with exact parameter values and once with vague estimations. The table reveals that the optimal selection of controls is mostly unimpaired if exact parameters are replaced with vague ones. In fact, in case of $B = 80,000$ €, both solutions match exactly. A test of 34 different solutions showed that in 76 % of the cases more than 20 out of 25 controls were unaffected. We found no case where more than 8 controls deviated.
Clearly, the advantage of easier parameter estimation comes with the drawback of more imprecise results. However, if it is difficult to provide precise input parameters an estimation according to the proposed

verbal scale (or a similar one) is reducing workload and complexity during a risk assessment. In any case, a quantitative estimation of the underlying structure of risks will yield better and more consistent results than a purely qualitative and non-integrated approach.

5 Conclusion

In this paper, we introduce a new approach to support decision makers in selecting security controls in a cost-effective manner. The approach includes a quantitative optimization model that takes into account how threats are arising in information systems and adds the element of uncertainty. By defining a structure that considers uncertain elements, it is possible to determine how much loss can be expected within a specified time period. In addition, by leveraging well-established methods of mathematical optimization, the best possible investment strategy can be selected. This new approach is a valuable method to comprehend complex investment decisions in information security and hence achieve significant improvements over the initial situation.

Based on a case study, we demonstrate and discuss the advantages of using a quantitative optimization approach to support decision making in information security. The presented model enables automatic generation of different risk assessment tools (e.g., risk matrices) and can be used to calculate specific key performance indicators (e.g., ROSI). These tools can either be used for detailed What-if analyses or as the basis of management decisions. Results of our study indicate that already relatively small investments yield a significant risk reduction. This characteristic is consistent with the principle of diminishing marginal utility of security investments and emphasizes the importance of profound business decisions in the field of information security. Furthermore, we showed that our model also produces good results even if input parameters are estimated imprecise or according to a vague estimation scale.

We are aware that the problem at hand may include additional factors we have not yet considered. This may involve compatibility of controls, in particular if controls are mutually exclusive or diminish each other's effectiveness. It is also a widely shared notion that security controls can induce additional (hidden) costs by reducing user productivity. A model that takes all those factors into account might better mimic reality, but at the same time requires more information and increases complexity. Besides including additional factors, it may be advantageous to apply different concepts to treat uncertain properties of the problem. If additional information is available, it is possible to use more detailed distributions to specify model parameters and thus improve the representation of uncertainty. Scenario approaches, robust optimization techniques, or stochastic programming may be applied. Moreover, the extension to a multi-period model is conceivable to consider dynamic aspects and medium- to long-term decision making. As a next step, we are going to investigate and compare respective models and their results.

Acknowledgment

This work was partially supported by the Horst Görtz Foundation.

References

Aagedal, Jan Øyvind, Folker den Braber, Theo Dimitrakos, Bjørn Axel Gran, Dimitris Raptis, and Ketil Stølen (2002). "Model-based risk assessment to improve enterprise security." *Proceedings of the Fifth International Enterprise Distributed Object Computing Conference*, 51–62.

Adams, Anne and Martina Angela Sasse (1999). "Users Are Not the Enemy." *Communications of the ACM* 42 (12), 40–46.

Aissa, Anis Ben, Robert K. Abercrombie, Frederick T. Sheldon, and Ali Mili (2010). "Quantifying Security Threats and Their Potential Impacts: A Case Study." *Innovations in Systems and Software Engineering* 6 (4), 269–281.

Aven, Terje (2011). *Quantitative Risk Assessment: The Scientific Platform*. Cambridge: Cambridge University Press. ISBN: 978-1139496438.

Beer, Martina, Gilbert Fridgen, Hanna-Vera Mueller, and Thomas Wolf (2013). "Benefits Quantification in IT Projects." In: *11. Internationale Tagung Wirtschaftsinformatik, Leipzig, Germany, February 27 – March 1, 2013*, pp. 707–720.

Berinato, Scott (2002). *Finally, a real return on security spending*.

Böhme, Rainer and Thomas Nowey (2008). "Economic Security Metrics." In: *Dependability Metrics*. Vol. 4909. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 176–187. ISBN: 978-3540689461.

Bojanc, Rok and Borka Jerman-Blažič (2008). "An Economic Modelling Approach to Information Security Risk Management." *International Journal of Information Management* 28 (5), 413–422.

Bojanc, Rok, Borka Jerman-Blažič, and Metka Tekavcic (2012). "Managing the investment in information security technology by use of a quantitative modeling." *Information Processing and Management* 48 (6), 1031–1052.

Cavusoglu, Huseyin, Birendra Mishra, and Srinivasan Raghunathan (2004). "A Model for Evaluating IT Security Investments." *Communications of the ACM* 47 (7), 87–92.

Cavusoglu, Huseyin, Srinivasan Raghunathan, and Wei Yue (2008). "Decision-Theoretic and Game-Theoretic Approaches to IT Security Investment." *Journal of Management Information Systems* 25 (2), 281–304.

FICO (2015). *FICO Xpress Optimization Suite*. Version 7.6. URL: `http://www.fico.com/en/products/fico-xpress-optimization-suite`.

Gal-Or, Esther and Anindya Ghose (2005). "The Economic Incentives for Sharing Security Information." *Information Systems Research* 16 (2), 186–208.

Gordon, Lawrence A. and Martin P. Loeb (2002). "The Economics of Information Security Investment." *ACM Transactions on Information and System Security* 5 (4), 438–457.

Hassanzadeh, Farhad, Mikael Collan, and Mohammad Modarres (2012). "A practical R&D selection model using fuzzy pay-off method." *International Journal of Advanced Manufacturing Technology* 58 (1-4), 227–236.

Hausken, Kjell (2006). "Returns to information security investment: The effect of alternative information security breach functions on optimal investment and sensitivity to vulnerability." *Information Systems Frontiers* 8 (5), 338–349.

Heiderich, Mario, Marcus Niemietz, Felix Schuster, Thorsten Holz, and Jörg Schwenk (2012). "Scriptless Attacks: Stealing the Pie Without Touching the Sill." In: *Proceedings of the ACM Conference on Computer and Communications Security*. CCS '12. ACM, pp. 760–771. ISBN: 978-1450316514.

Hogganvik, Ida and Ketil Stølen (2006). "A Graphical Approach to Risk Identification, Motivated by Empirical Investigations." In: *Model Driven Engineering Languages and Systems*. Ed. by Oscar Nierstrasz. Vol. 4199. Lecture Notes in Computer Science. Berlin and Heidelberg: Springer Berlin Heidelberg, pp. 574–588. ISBN: 978-3540457725.

Hua, Jian and Sanjay Bapna (2011). "Optimal IS Security Investment: Cyber Terrorism vs. Common Hacking." In: *Proceedings of the International Conference on Information Systems*. Ed. by Dennis F. Galletta and Ting-Peng Liang. Association for Information Systems. ISBN: 978-0615559070.

Kaplan, Stanley and B. John Garrick (1981). "On The Quantitative Definition of Risk." *Risk Analysis* 1 (1), 11–27.

Luo, Xin, Richard Brody, Alessandro F. Seazzu, and Stephen D. Burd (2011). "Social Engineering: The Neglected Human Factor for Information Security Management." *Information Resources Management Journal* 24 (3), 1–8.

Mather, Tim, Subra Kumaraswamy, and Shahed Latif (2009). *Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance*. O'Reilly Media, Inc. ISBN: 978-0596802769.

McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts (2005). *Quantitative risk management: Concepts, techniques and tools*. Princeton series in finance. Princeton and N.J: Princeton University Press. ISBN: 978-0691122557.

Rabai, Latifa Ben Arfa, Mouna Jouini, Anis Ben Aissa, and Ali Mili (2013). "A cybersecurity model in cloud computing environments." *Journal of King Saud University - Computer and Information Sciences* 25 (1), 63–75.

Roy, Sankardas, Charles Ellis, Sajjan Shiva, Dipankar Dasgupta, Vivek Shandilya, and Qishi Wu (2010). "A Survey of Game Theory As Applied to Network Security." In: *Proceedings of the 43rd Hawaii International Conference on System Sciences*. HICSS 10. IEEE Computer Society, pp. 1–10. ISBN: 978-0769538693.

Ryan, Julie J. C. H., Thomas A. Mazzuchi, Daniel J. Ryan, Juliana Lopez de la Cruz, and Roger M. Cooke (2012). "Quantifying information security risks using expert judgment elicitation." *Computers & Operations Research* 39 (4), 774–784.

Schilling, Andreas and Brigitte Werners (2013). "A Quantitative Threat Modeling Approach to Maximize the Return on Security Investment in Cloud Computing." In: *Proceedings of the International Conference on Cloud Security Management*. Ed. by B. Endicott-Popovsky. Reading: Academic Conferences and Publishing International. ISBN: 978-1909507678.

— (2014). "Optimizing information security investments with limited budget." In: *Operations Research Proceedings 2014*. Springer, pp. 1–6.

Schuster, Felix and Thorsten Holz (2013). "Towards reducing the attack surface of software backdoors." In: *Proceedings of the ACM Conference on Computer and Communications Security*. CCS '13. ACM, pp. 851–862. ISBN: 978-1450324779.

Sonnenreich, Wes, Jason Albanese, and Bruce Stout (2006). "Return On Security Investment (ROSI) – A Practical Quantitative Model." *Journal of Research and Practice in Information Technology* 38 (1).

Stoneburner, Gary, Alice Goguen, and Alexis Feringa (2002). *Risk management guide for information technology systems*. Vol. 800-30. NIST special publication Computer security. Washington and DC: U.S. Gov. Print. Off.

Sun, Lili, Rajendra P. Srivastava, and Theodore J. Mock (2006). "An Information Systems Security Risk Assessment Model Under the Dempster-Shafer Theory of Belief Functions." *Journal of Management Information Systems* 22 (4), 109–142.

The Open Web Application Security Project (2013). *OWASP Top Ten: The Ten Most Critical Web Application Security Risks*. Ed. by The Open Web Application Security Project.

Tsiakis, Theodosios (2010). "Information Security Expenditures: a Techno-Economic Analysis." *International Journal of Computer Science and Network Security* 10 (4), 7–11.

Wang, Jingguo, Abhijit Chaudhury, and H. Raghav Rao (2008). "A Value-at-Risk Approach to Information Security Investment." *Information Systems Research* 19 (1), 106–120.