

DESIGNING A RULE-BASED CYBER RISK ASSESSMENT TOOL FOR SMALL TO MEDIUM ENTERPRISES

RUBEN KOEZE

DELFT UNIVERSITY OF TECHNOLOGY

DESIGNING A RULE-BASED CYBER RISK ASSESSMENT TOOL FOR SMALL TO MEDIUM ENTERPRISES

ABSTRACT

With the extensive use of the internet nowadays, companies are becoming more and more at risk from cyber-attacks. Especially SMEs are vulnerable, as they lack in cybersecurity. This is due to time, resource and knowledge constraints. In the literature no clear solutions can be found in order to help SMEs with their cybersecurity. To overcome these issues, this research aims in developing a method with which a tool can be developed that can help SMEs in getting insights in their cybersecurity status. This tool should be capable of letting an SME do their own risk assessment without investing too much money or time. In order to do this, existing methods are analyzed. After the analysis, the TRESPASS method is used as a basis for the tool. This method is slimmed down in order to fit the needs of SMEs. By creating a knowledge based system, the SMEs can use the knowledge of cybersecurity experts without the requirement to have to knowledge in-house.

Keywords: cybersecurity, risk assessment, risk assessment tool, smes, rule-based system

INTRODUCTION

The penetration of internet in the Netherlands is at an all-time high, with only 8 percent of the population never using the internet [1]. While the Netherlands is one of the front-runners in the penetration of the internet, the rest of the European Union is not far behind. Within the whole of the European Union, the internet usage is at 82% for people between 16 and 74 years old [2].

With the rise of the internet come threats via the internet. These attacks can vary from installing malware that shows ads, stealing critical company information or bringing down ICT infrastructure. While these attacks can be aimed at the individual, they can also be aimed at companies. Everything concerning the prevention and response of and to these attacks is called cybersecurity. While this topic is also on the rise, as can be seen in Google Trends [3], it still remains a neglected subject for a lot of people, but especially for companies.

With a lack of cybersecurity, a lot of issues can arise. The causes of these issues come with different likelihoods and impacts. By using risk management, the costs of risks “firing” can be minimized. Risk management as defined by Cambridge University: “the activity of calculating and reducing risk, so that an organization does not fail or lose money” [4]. The first part of this definition is the calculation of the risk. This calculation is done by doing a risk assessment. A risk assessment is the basis for risk management and although it can be applied in many sectors, it is also essential for cybersecurity [5, 6].

The focus of this research will lie on Small and Medium-sized Enterprises, or SMEs. For this research the definition used for a SME is that the company may have a maximum of 250 employees or a maximum turnover of 50 million euros. The focus on SMEs is chosen because of the big part of the economy they represent and because the state of their cybersecurity. In the Netherlands, SMEs are a big part of the national economy. According to the *Centraal Bureau voor Statistiek* SMEs make up more than 60% of the Gross

Domestic Product and provide 70% of the employment opportunities in the Netherlands [7]. According to Eurostat, SMEs provide for 99% of the jobs within the European Union [8].

CYBERATTACKS ON SMES

In a research conducted by Capgemini and TNS Nipo, only 35% of the SMEs did pay attention to cybersecurity once in a while [9]. This is also confirmed by research by the NCSC [10]. This is no surprise when looking at the percentage of SMEs that think that it is very unlikely that they will be the target of an attack. More than half of the SMEs thought that it is very unlikely that they will become the target of a cyber-attack, mainly due to an underestimation of their asset value [11]. This might explain why SMEs do not take measures, even though, when asked, they are aware of the fact that they are not well protected and prepared against a possible cyber-attack [12]. Only 14% of the SMEs rate their ability to mitigate cyber risks above 6 on a scale from 1 to 10 [13]. One of the reasons that SMEs do not pay much attention to cybersecurity is the thought that they do not have high value assets for attackers [11, 14]. However, different researches found other reasons for this lack of security. Other reasons are: lack of investments in cybersecurity [13, 15], a lack of in-house expertise [16] and limited resources [16]. While companies often have IT staff, they are not specialized or focused on cybersecurity. But what is important to notice is that with limited resources and small measures, big results can be achieved [9].

Based on the existing literature, no tool is suited for helping SMEs in doing a cyber-risk assessment without creating a lot of overhead. Therefore a research is proposed to fill this knowledge gap. This research will be done by answering the main question:

What would a tool look like that helps SMEs do cyber-risk assessments and point out the weaknesses in their cybersecurity?

With the help of multiple sub questions this main question will be answered. The following sub questions are defined:

1. How can existing frameworks and assessment methods be tailored for SMEs?
2. Does the tool meet the requirements of SMEs?

METHODOLOGY

In order to answer the abovementioned questions, different methods are required. In this chapter, the research method for each of the sub questions will be explained.

Literature review

Sub question two will be answered by the means of a literature study. Searches for existing literature will be conducted by the use of Google Scholar, Scopus and Web of Science. A sample of the keywords, or combination, used will be: cybersecurity, framework, SME, risk assessment, cyber risk. In the found literature, the references can be used for further exploration of the subject.

Design science

The main deliverable for the proposed research is a tool. This tool shall be built based on the principles proposed by the design science theory by Hevner [17]. The theory by Hevner consists of three parts. The environment, the knowledge base and the IS research. The environment can be seen as the problem space. Within this space, everything that defines the problem that creates the urge for the to be designed artifact is present.

The knowledge base is the existing literature that defines frameworks, theories, methods and everything that is relevant for the research. Knowledge can be extracted from the knowledge base, but the research will also provide new additions to it.

The last part is the IS research. This research is done with the input from the environment and the knowledge base. In this part the actual building of the artifact is done by using all knowledge and then evaluating the artifact, also using knowledge from both the environment and the knowledge base. This is a repeating process in which the artifact is built, evaluated and then changed on the basis of what the evaluation concludes.

As a basis for this project the TRESPASS project is chosen. This includes a tool that is too complex for SMEs, but has principles and design choices that can be adopted.

MODELING AND CALCULATING THE RISK

The goal of the tool is to show a company what their risks within the cyber domain are. As said, this needs to be done with few resources. In order to accomplish this, the tool has to be easily accessible, which translates to a “simple” tool. For doing calculations and determining where the risks lie, the user of the tool has to model their organization within the tool. To accomplish this, the way of modelling has to be simple, but should contain all elements that a user needs to correctly model an organization. In the coming paragraphs, these different elements of the modelling will be discussed.

ELEMENTS IN THE MODEL

The goal of the tool is to model an organization in such a way that is easy to understand, easy to do, but still gives a correct representation of how the organization is structured.

As mentioned, the TRESPASS project will be used as a basis for this purpose. The TRESPASS project is a cooperation between multiple organizations and universities in Europe. The project provides an “attack navigator”. As stated on the website of the TRESPASS project: *“This navigator makes it possible to say which attack opportunities are possible, which of them are the most urgent, and which countermeasures are most effective.”* [18]. The way TRESPASS accomplishes this is the use of a visual representation of the company; called a Socio-Technical security model, or the TRESPASS-model within TRESPASS. This model consists of multiple elements that can create the structure of an organization.

An important aspect of this process is the fact that the modelling is done in cooperation with an analyst with specific cybersecurity knowledge. This is done because the inner workings of the TRESPASS attack tree navigator (and the modelling that goes with it) are so complex, that it needs specific knowledge of its workings in order to be

used. The advantage of the comprehensive structure of TRESPASS is that it is suited for big companies or projects. The disadvantage of this, is that it is not easy to use by smaller actors and the threshold for using TRESPASS is therefore high.

The goal of this tool is also to show organization their weak points and display this in a way that shows how attack might enter an organization. While this tool is well developed, it is very complex. The essence of the tool is the same as the goal of this thesis, but the complexity of the TRESPASS project makes it not suited for SMEs. While this is the case, the structure that is used can be adopted in this thesis in order to model an organization. The components used in TRESPASS are shown in Table 1 with an explanation of what each element is. In Table 2 **Error! Reference source not found.** are the real world equivalents shown with the modelling component that fits the real world component.

Component	Description
Actors	Represent human players or processes involved in the system
Assets	Can be either items or data
Locations	Represent where actors or items may be situated either physically or digitally
Edges	Describe possible relocation paths between locations
Policies	Describe access control and specify allowed actions, e.g., get some data item from a location or move between locations
Processes	Formalize certain state transition mechanisms, e.g., computer programs or virtual machines

Table 1 - Different component TRESPASS [28]

Real world	Model component
Relevant area	Locations and edges
Computer networks	Assets and edges
Human actors	Actors
Physical access control	Policies and processes
Computer access control	Policies and processes
Software processes	Processes

Table 2 - TRESPASS modeling components fitted to the real world [28]

Within TRESPASS, there is also a focus on physical access to components, this is the reason that there is a component *Location*. This component is not taken into account for this research, as it is out of the scope. The assumption is made

that SMEs operate from one location or that the location aspect is negligible. This is done in order to keep the structure simple and within the constraints that a SME has. The scope of this thesis does not focus on the physical security that a company has, but only on the security that is in the cyber domain. Furthermore, the components that are described within the model of this thesis will be discussed. Important to notice in these descriptions is the fact that names of components might be the same, but the content of the components can differ from the TRESPASS components. This is due to the simplification to fit the modelling within time constraints that fit SMEs. It also has to do with the simplification due to the technical complexity that has to be limited.

The last mentioned component, processes, are also not incorporated in this research. The inclusion of state transitions will drastically increase the complexity of the modelling. This will prevent the modelling method in reaching its goal: creating a simple graspable method for modelling an organization.

The component that remain, and of which the model will be built, are *actors*, *devices*, *assets*, *policies* and *edges*. These different components will all be shortly discussed as for what they will stand for in the modelling of an organization.

Actors

The first component is the human factor within a company. This component is the same as it is in TRESPASS, except for the fact that it cannot describe processes; it will always represent one or multiple actors. When the last is the case, actors can be grouped. In most companies there will be standard groups like HR, system admins and administrative. These groups all have different permissions which brings different risks.

It is important to notice that these are always actors within the company. These actors cannot represent the attacker.

Devices

While the *device* component does not exist in TRESPASS, but is an adoption of the asset component. In the devices category all hardware components of an organization are

described. This is done to accomplish that the tool is easy accessible for people without cybersecurity knowledge. A separate devices component category is clear to understand and gives a good overview of what devices are present in an organization. The asset category in TRESPASS might cause confusion, as the asset category was both for data and for devices.

Assets

Different from the category in TRESPASS, the assets category is the value for a company. Most of the times this is data that a company has stored on their network. As described in the Devices category, in TRESPASS the Assets category consists of both the devices and the data that is at hand. For the simplification and the easiness to understand the tool, this category is divided in two. Assets can be things like medical data, credit card data or personal customer data. In this category, the thing that is most valuable for a company (in the IT area) is defined.

Policies

While this is again a category that has the same name as in TRESPASS, it is not the same thing. In TRESPASS the policies category is a complex one, in which very specific actions can be described. Things like access control and the movement of certain data. This interpretation of policies is too complex for the scope of this thesis and is therefore simplified. It can even be seen as a complete change of the meaning policy from the TRESPASS meaning.

Within the tool, a policy is something that influences an actor or device. This can be things like, what education does an actor have or how often is a device updated. This all influences the risk a link in the model carries, but more on how these influences work will be described later on in this chapter.

Edges

In TRESPASS, the edges are the connections between different components of the model. While edges, in the sense of connections, exist in the models created within this research, they are simply called connections. How these

connections work and what kind of influence they have will be described in the next part.

STRUCTURE OF THE MODEL

The components described in the previous paragraphs can create a model that represents an organizations IT status. While the different components are simple, they still include most of the aspects of an organization that are relevant for the cybersecurity of an organization.

Still, it is important to notice that creating all different components to model an organization is not enough. These different components need connections between them in order to show the usage and data flows between the different components.

In the simplest form, the structure is: an actor has access to a device and with that device the actor can access an asset. This flow is shown in Figure 1.

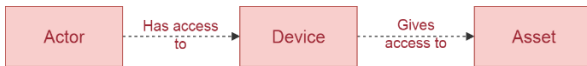


Figure 1 - Connections between different components

All of the access flows within an organization can be modelled with this structure. In a simple example, the main asset of a company is the credit card data of its customers. This data can be accessed with a certain computer. The group of system administrators has access to this type of computer, which makes that this group can access the asset (credit card data). While this is a singular flow, this model can be made more complex when for example a group support staff also has access to this computer, or when the system administrators also have a mobile phone which gives them access to the asset. No components are bound to one or two connections.

Policies come in last. When the structure of the organization is built with the actor, device and asset components, policies can influence the actor of device components. This is shown in Figure 2.

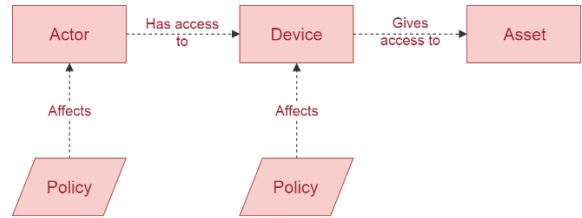


Figure 2 - Influence of policies on devices and actors

Policies have an impact on the probability of a risk firing through that component. How these calculations are done is described further on in this chapter. Examples of policies are *a device is updated every week* or *an actor is trained every year on the risks of cyberattacks*. These policies improve the security of a component, thus limiting the risk to the asset which the components are connected to.

CALCULATING THE RISK

In the structure described in the previous paragraph, paths exists. It is clear what actor can access which asset with which device. This is similar to the concept of an attack tree, as introduced by Schneier [19]. An attack tree is, as defined in the paper by Schneier: *“A way of thinking and describing security of systems and subsystems”* [19]. It represents the attacks and countermeasures on a system, displayed as a tree structure. An example of an attack tree is shown in Figure 3.

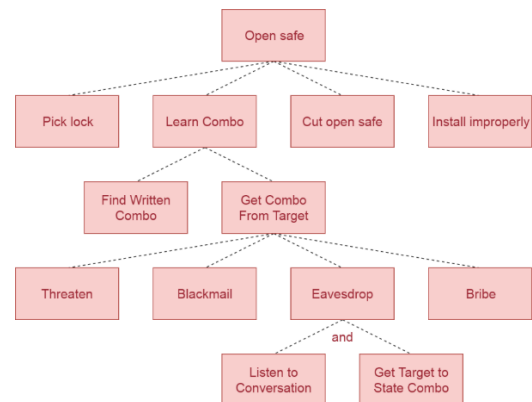


Figure 3 - Attack tree example [19]

The example shows that the main goal for the attacker is to open the safe. To do this, there are multiple options. For some options (in this case *Learn Combo*) there are again multiple options. Going down the tree it shows all possible ways for the attacker to get to their end goal of *Opening Safe*. In this case, the tree is relatively simple, but in the

bottom it shows a good example of how to accomplish to *Learn Combo by Get Combo From Target* using *Eavesdrop*. To accomplish the *Eavesdrop* method, the attacker has to *Listen To Conversation*, but that is not all. The attacker also needs to make sure that the *Target States the Combo*. What this means for the calculations will be explained in the next paragraph.

It is important to point out that the attack tree does not fit the tool that is being created for this research. This is because the attack tree method needs the perspective of an attacker. For this research, the assumption is made that the attacker has the same capabilities for each SME, more on what this means for the calculations will be discussed later on. This means that the perspective can switch from the attacker to the defender, which means that creating a structured overview of what the vulnerabilities on the defending side are is easier. With the tool being used by people that are defending and do not have specific knowledge on the possible attacks on their system, the attacker cannot be set specific to the business. This is why the structure in the previous paragraph is chosen. However, the attack tree shows an interesting way of calculating risk; one that can be adopted into the model that is chosen for this research. Where in the attack tree the methods for entering a certain node in the systems is shown, the structure remains the same in this model, although the nodes do not represent the actions an attack does, but the defending nodes. As said, because an assumption is made on the attack strength of the attacker, these odds do not differ per attacker, and the impact of different defense strategies will remain the same. Therefore, the possible calculations that can be done with an attack tree will be discussed in the next paragraph.

CALCULATIONS IN THE ATTACK TREE

In an attack tree, there are multiple ways of calculating what the highest risks are (or what the best attack paths are). In the paper by Schneier, examples like costs, attacker-skills or probability of success are given. This last one is the one that is relevant, as the goal of this research is to create a tool that can conduct a risk assessment.

In a paper by Ingoldsby [20] this calculation method with threat probabilities is further defined. He states that every node in the tree has a certain probability of succeeding. This chance is determined by looking at multiple factors to succeed in that attack. In the example given the cost for that attack, the technical ability necessary and the noticeability that comes with the attack are taken into account. Those three factors are all a number between 0 and 1. When these three numbers are multiplied, the ease of the attack is determined. Or in other words; the probability that this attack is conducted by an attacker. This gives a probability for every node that this method is used for an attack.

ADAPTING THE ATTACK TREE RISK CALCULATIONS

As said, while an attack tree is something different than the model used in this thesis, the structure remains similar. Where an attack tree is viewed from the perspective of the attacker, the structure used in this research is on the viewpoint of the defender. This means that for an attack tree, the probabilities are displayed as a probability that an attacker succeeds in doing that one component. In the structure that is used in this thesis, the components in the “tree” are “defending” elements while the “attack” elements are not mentioned explicit (different from for example an attack-defense tree). As stated, an actor is connected with a device, which is again connected with an asset. This creates path from the actor, through a device, to an asset. This means that if either the actor or the device is breached, the asset is accessible. This can be better explained by using an example. If an actor can access an asset via a device, this means that both the actor and the device have access to that asset. This means that even if the device is perfectly secure, if the actor gets breached, the attack can use the actor to get to the asset. The other way around this works the same way. Even if the actor does everything secure from a cybersecurity perspective, when the device is not secure and can be accessed by an attacker without the involvement of an actor, the attack can still reach the asset. So, that means the structure from the model can be translated to the diagram as shown in Figure 4.

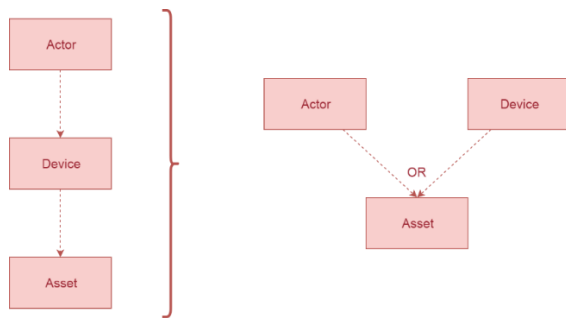


Figure 4 - Translate model to diagram for probabilities

The goal of this diagram is to calculate the probability on the asset, after which the risk on the asset can be calculated. Just like in an attack tree, every node will have a probability that it will be used for a breach. Assuming that a breach on actor and a breach on device are both mutually exclusive events, the following formula can be used in order to determine the probability that the asset will be breached:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Which translates in this particular case to:

$$P(\text{Asset}) = P(\text{Actor}) + P(\text{Device}) - P(\text{Actor} \cap \text{Device})$$

For example, if the probability that the actor in this case is breached is 0.6 and the probability that the device is breached is 0.4, the probability that the asset is breached is:

$$P(\text{Asset}) = 0.6 + 0.4 - 0.6 * 0.4 = 0.76$$

This calculation gives a probability of 0.76 (or 76%) that the asset will be breached. However, as these are calculations for a risk assessment, the risk on the asset needs to be calculated. The general formula for calculating risk is $Risk = Probability * Impact$. This formula is also used in these calculations, however, the impact is not yet determined. The impact will be determined by the user of the tool, indicating the impact of a breach on a certain asset from 1 to 5, where the numbers 1 to 5 correspond to number between 0 and 1.

Impact inputted by user	Impact in calculation
1	0.20
2	0.40
3	0.60

4	0.80
5	1.00

Table 3 - User input conversion






Risk	Name	Color
0.00 – 0.20	Very low	
0.21 – 0.40	Low	
0.41 – 0.60	Medium	
0.61 – 0.80	High	
0.81 – 1.00	Very high	

Table 4 - Risk names and corresponding colors

In Table 3 the conversion from the user input to the number that is being used for calculation is shown. The reason that the impact is chosen from 1 to 5 is because of the ease to understand for the user. Because all calculations are done in number from 0 to 1, the conversion is also done to conform to this standard. This conversion is a linear conversion. To follow up on the example that was just given, if the asset used for calculations was given an impact score of 4, the calculation for the final risk would be:

$$Risk \text{ for Asset} = 0.76 * 0.8 = 0.61$$

This means that the risk for the asset will fall in the category *High* as can be seen in Table 4, which is a severe risk and should ring a bell at the user end. Again, this conversion is done to make the interface easier for the user. This conversion is again linear.

The calculations that are done determine the probability and risk for the breach of an asset. However, these calculations are done with the probabilities of a breach of an actor or a device. How these probabilities are determined will be explained in the next paragraph.

DETERMINING THE PROBABILITIES OF COMPONENTS

As said, the different components have different probabilities of being breached. These probabilities are the base for the calculations on determining the risk on the asset (as discussed in the previous paragraph). However, the calculations that work on the different components are not yet discussed. In the *Structure of the model* part of this chapter, it is explained that different policies have an impact on the components. This impact will have an effect on the probability that a component will be breached. The probability of a component being breached will influence

the probability of an asset being breached, following the structure shown in Figure 4 earlier on in this chapter. One problem with the fact that the probability of all these components need to be determined is that there is knowledge needed in order to do this. A quote from a research by McGraw illustrates this perfectly: *“The key to an effective risk assessment is expert knowledge of security”* [21]. The problem with this is that the goal of this research focusses on creating a risk assessment method that is accessible to people that do not have this particular knowledge. This means that these two aspects have to be decoupled, the knowledge of what are the probabilities that a certain node will be breached needs another origin than the user. Somehow the knowledge of certain probabilities need to be put into the system, this way the user just has to select different options that are pre-programmed in the system.

This kind of system is called a knowledge-based system (or KBS). In such a system, the knowledge that is required for making decision is put into a knowledge-base. This makes that the system can make decision without the user having to input certain knowledge, furthermore it is flexible as it can be easily extended and refined [22]. A knowledge-based system works with certain rules, most of the times these are IF-THEN rules. An example of a rule, as given by Smith [22]:

```
IF
    there exists a normal fault with class
    unknown, and
    there exists a red pattern
        with length < 50 ft.,
        with bottom above the top of the
        fault,
        with azimuth perpendicular to the
        fault strike
THEN
    the fault is a late fault with direction
    to downthrown block equal to the azimuth
    of the red pattern
```

As can be seen, the knowledge base must contain certain knowledge to conclude the fact that is written in the THEN statement. As can be seen, there are multiple conditions that conclude into the THEN statement. This is not completely in line with the structure of the model used in

this research. The rule as abovementioned can be translated in the following form for this research:

```
IF
    policy 1 works on component, and
    policy 2 works on component, and
    policy 3 works on component
THEN
    the probability of a breach on the
    component is impacted with impact(policy
    1, policy 2, policy 3)
```

While this works, it is not a rule-based system as proposed Smith. The difference lies in the fact that the policies in the IF statement have an impact on what happens in the THEN statement. In the THEN statement, calculations are required in order to determine what the impact on the probability of a breach on the component is. While this might not be the use as Smith intended it, it still has the advantages which are needed for this model: it can use knowledge of experts and perform them on a component without the need for the user to have knowledge on the risks that work on a component or policy.

In the end, it means that it is not a rule-based system in the traditional sense of the word: it uses only one rule with different inputs. But by choosing these different inputs per policy, the rules serve the purpose of a well-structured knowledge-base than can be easily read and easily extended.

How these rules and calculations that come with the rules come together is explained in the following paragraphs of this chapter.

CALCULATIONS ON THE POLICIES

An important aspect of the knowledge-based system is how the rules are structured. Because no such a system has ever been used, it is hard to find relevant literature on this subject. Most literature found on the subject is based on fuzzy rules [23] or are hardcoded rules that do not use probabilities (just like the given example above) [24]. While this means that there are no calculations that can be used from a knowledge-based system, the system remains suited for this purpose as the expert knowledge can be incorporated in the model. For the calculations, the Gordon-Loeb model will be used in combination with the rules [25]. This model is been widely accepted as

determining what the effects on successive investments in cybersecurity are. In this case, this will be used to determine the diminished effect [25, 26] of more policies on a node, as these can be aligned: more policies is more investment.

The Gordon-Loeb model uses the formula shown below. Where *Effect* is the effect of the policy on the improvement of the probability of the breach on the component. The *impact factor* is the impact of the policy as determined by expert, on a scale from 0 – 1. Lastly, *l* is the number of the policy, where the first policy will have a bigger impact than the next one.

$$Effect = 1 - \frac{impact\ factor}{(l + 1)}$$

When plotting this formula, it shows the impact of multiple policies implemented on the same component with the same impact (0.8 for this example). The plot of this formula with *impact factor* = 0.8 can be seen in Figure 5.

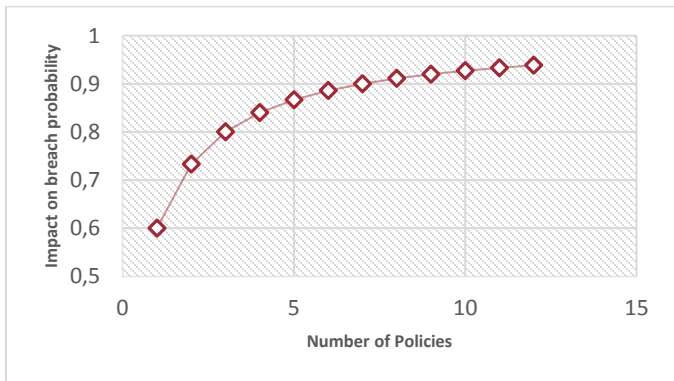


Figure 5 – Diminishing marginal returns effect of number of policies

This *Effect* will then be used to impact the probability that the component will be breached. This will look like the following formula:

$$P(device\ with\ Policy_1\ to\ Policy_n) = P(device) * Effect_1 * Effect_2 * ... * Effect_n$$

As said, and shown in Figure 5 this includes the effect of diminishing marginal returns. This means that every extra policy has relatively less impact than the previous one. Using this example of the effect of policies with the same effect on a component with base risk 0.7 is shown in Figure

6. This graph clearly shows that the first policy has a big effect (decreasing the breach probability from 0.7 to 0.28) while the following policies have less of an effect. The second policy decreased the breach probability from 0.28 to 0.19.

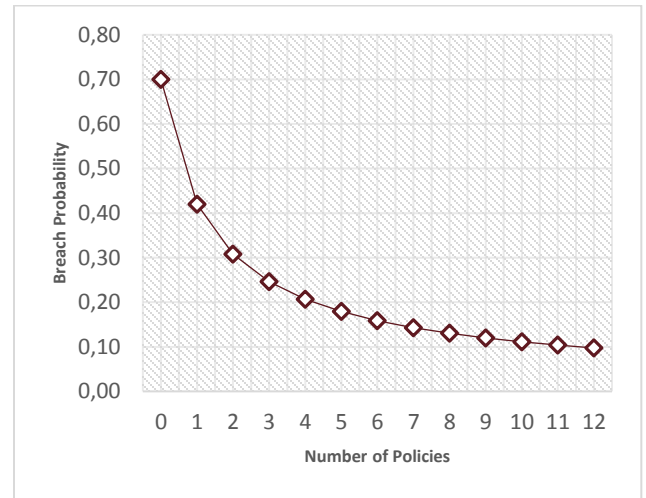


Figure 6 - Impact on breach probability

STRUCTURING OF THE RULES

To apply these calculations on the different actors and devices, a clear structure has to be defined in order to create a knowledge base. This knowledge base needs to be easy in maintenance and editing or adding rules. In order to do this, each rule has (like most KBS) an IF-THEN structure. As said, every rule will influence the actor or device with a value of 0 to 1.

IF *name* IS *value* THEN PROBABILITY ON *works on* IS DECREASED BY FACTOR *factor*

As an example, the patch frequency of a device is given. In this case the variables in the rule will look like:

name = patch frequency
works on = device
value = weekly
factor = 0.8

Which results in the following rule:

IF *patch frequency* IS *weekly* THEN PROBABILITY ON *device* IS DECREASED BY FACTOR **0.8**

Because in this case, the value and risk have multiple options (value can be weekly, monthly, yearly, etc.) this creates the following rules:

IF *patch frequency* IS *weekly* THEN
 PROBABILITY ON *device* IS DECREASED BY FACTOR *0.8*
 IF *patch frequency* IS *monthly* THEN
 PROBABILITY ON *device* IS DECREASED BY
 FACTOR *0.6*
 IF *patch frequency* IS *yearly* THEN
 PROBABILITY ON *device* IS DECREASED BY
 FACTOR *0.4*

These rules can be displayed in the form of a table in the following form:

#	Name	Works on	Value	Factor
1	Patch frequency	Device	Weekly	0.8
			Monthly	0.6
			Yearly	0.4

Table 5 - Policy rules in table form

DETERMINING THE RULES

For a proof of concept, the knowledge base should have an initial set of rules that can be used in order to create a model. It is important to stress that the determining of the rules is not a core part of this research, it is merely intended in order to proof the model's concept. Therefore the ruleset that is created is not exhaustive in any way. The way the model was set up is in such a way that the rules in the knowledge-base can be edited or removed with ease. It is also easy to extend the knowledge base with extra actors, devices, assets or policies.

To create an initial set of rules that can be used for the proof of concept, an expert session was held. In this expert session, two KPMG consultants determined the base value of different components and the possible policies that can work on those components. These two consultants were selected on basis of their expertise. The first one is familiar with the quantification of breach probabilities for customers of KPMG while the second one has an expertise on the penetration testing of systems, and therefore can estimate what common attack vectors are. A summary of the expert session can be found in Appendix G. The results of the session are shown in Appendix H.

VALIDATION

The validation of the model could have been done with a case study. However, because of time constraints and the complexity of this way of validation, this is out of the scope of this research. With the help of an interview and an expert

session, the model is validated. In the expert session, it was confirmed that the structure of the proposed model is good. While it could use expansion, due to the simplicity and scope this is a good start. This is also confirmed in the second interview. In the second interview it also was stated that the asset component might be of less value than the other components. Also the way of using experts in order to determine the probabilities of different aspects is one that is already been used in production, as confirmed by the expert session. This separation of knowledge is also confirmed to be a wise modelling decision by the second interview. In order to calculate the probability that an asset is breached, the assumption is made that when the device or actor that is connected to that asset is breached, the asset is breached as well. While there is some doubt about this structure, it is agreed upon that it is a perspective issue.

RESULTS

The implementation of this model is done in the form of a web-based tool. The tool incorporates the different components of the model as visual blocks that can be connected. This ensures the ease of use and gives a good visual representation of the company that has to be modelled.

This visual representation is also used for displaying the results of the calculations. The risks and probabilities are mapped and shown for the different components.

USE OF THE TOOL

The source code of the tool is open source and can be downloaded from the GitHub repository. Before use, this however means that the code needs to be run on a web server with PHP installed. The documentation on how to install the tool on a web server can be found on the GitHub repository. This documentation, together with the code, can be found on: <https://github.com/roebenk/thesis>.

After installation, the crucial part of the effectiveness of the tool is the knowledge base. As described, the knowledge base is filled with rules to run a proof-of-concept, but these rules are not meant for a production environment. Because the tool is open source, the tool can be implemented and

adapted by everyone that sees fit. This means that the knowledge base can be filled by individuals and used for their own purpose, keeping the structure of the tool. It of course also means that the knowledge base can be extended in the open source repository. Details on how to add value to the knowledge base can also be found in the repository.

SCREENSHOTS

The results of the development cycles give a proof of concept. In order to give an idea on how the translation from requirements to actual design is done, screenshots from the actual proof of concept are shown below.

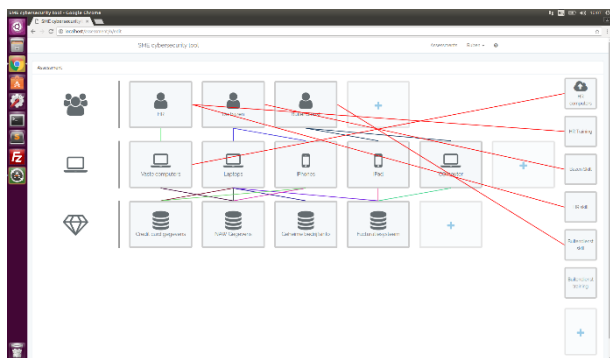


Figure 7 - Complete modelling overview

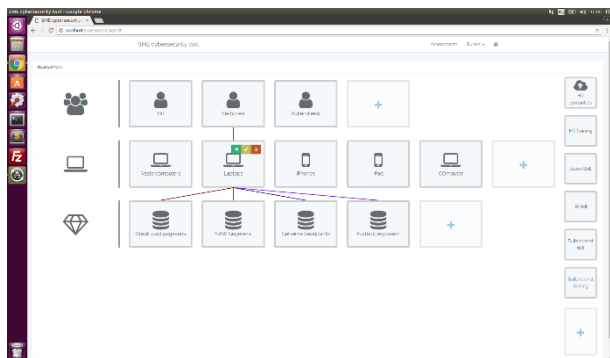


Figure 8 - Partial modelling overview

Figure 7 shows the overview in which the user can model the structure of the company. On the left side, from top to bottom, the actors, devices and assets are shown. On the right side, the policies are shown. The colored lines are the connections between these different blocks. Because the image can get cluttered due to the amount of connections, an extra option is built, in which the view can be uncluttered. Figure 8 shows that when hovering one of the

blocks, only the connections for that specific block become visible.

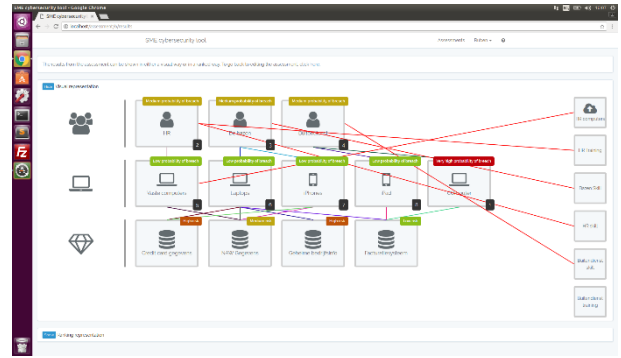


Figure 9 - Complete risk assessment overview

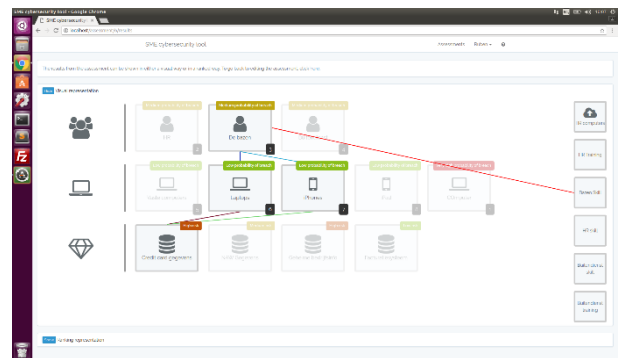


Figure 10 - Partial risk assessment overview

The same principle works for the risk assessment results page. In Figure 9 the different blocks are shown with the corresponding risks. The connections between the different blocks can clutter the view, thus a similar feature as in **Error! Reference source not found.** is implemented. When hovering over an asset, the complete influence on that asset can be viewed. This is shown in Figure 10.

As an addition to the visual representation in the results screen, a list with the different risks can also be shown. This can be seen in Figure 11. On the left, all the actors and devices with their corresponding breach probability are shown. On the right, the assets with their corresponding risks are shown.

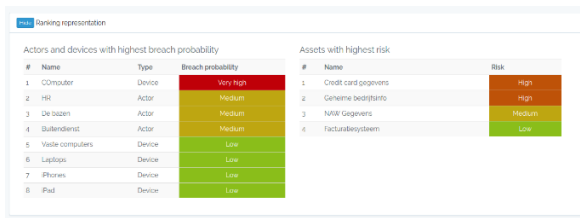


Figure 11 - Ranked visualization of risks

CONCLUSION

With the help of the different research questions, the main research question will be answered. Each paragraph will handle one or more research questions and discusses how these questions are answered.

WHAT DOES THE EXISTING LITERATURE SAY ABOUT CYBERSECURITY FOR SMES?

There is quite some literature on cybersecurity in SMEs. However, this literature mostly confirms the problem that cybersecurity within SMEs is not at the desired level. A lot of literature also confirms that most cybersecurity frameworks and assessments are lacking in applicability for SMEs. They are either too complex, or too require an external expert, something that is too expensive.

HOW CAN EXISTING FRAMEWORKS AND ASSESSMENT METHODS BE TAILORED FOR SMES?

The conclusion from the previous paragraphs is that both the existing risk assessments and frameworks do not fit SMEs. To fit the needs of SMEs, the method to model the structure of a company is adapted from the TRESPASS project. Because TRESPASS has a lot of options that are not relevant for SMEs, parts of the structure are used to fit the scope of SMEs better. The structure that remains consists of *policies*, *actors*, *devices* and *assets*. With these components, an organization can be structured. With this structure, the risks on the different assets needs to be determined. This is done by determining a base probability for a device that it will be breached, after which policies will decrease that breach probability. The calculations for this improvement are based on the Gordon-Loeb model, which states that every extra investment in cybersecurity will (relative to the previous) have less effect. This means that

every extra policy will have less of an improvement than the previous one. In order to calculate the probability that works on the asset, the probabilities of all the actors and devices that work on the asset will be combined in order to determine the final probability. This probability will be multiplied with the value of the asset, which gives the final risk.

An important conclusion of this model is that the base probabilities and policy probability improvements need to be determined by experts. It cannot be expected that the user has the knowledge to estimate these numbers. Therefore, only the structure and the value of the assets needs to be determined by the user. This ensures that the knowledge of the risk probabilities is separated from the knowledge of the company.

DOES THE TOOL MEET THE REQUIREMENTS OF SMES?

For SMEs to be able to work with the tool, it should be simple modelling of the organizations structure. The model created has been validated by the interview and expert session that were conducted. While there were areas for improvement, the model as it is constructed now will work and will have added value to SMEs.

WHAT WOULD A TOOL LOOK LIKE THAT HELPS SMES DO CYBER-RISK ASSESSMENTS AND POINT OUT THE WEAKNESSES IN THEIR CYBERSECURITY?

Finally, the main research question of this thesis needs an answer. It can be concluded that no tool exists that fills the void of cybersecurity assessments coming SMEs. To create a tool that is suited for SMEs, the different requirements as determined should be met. But more importantly, the knowledge for estimating the probabilities and risks of the system should be included in the tool. With other words, the user should not have to worry about this, but should only construct a model of their company. With the created model, this constraint is met, and it is possible for SMEs to do this risk assessment themselves. With the visual representation of both the model and the risks, a clear understandable risk assessment can be executed.

DISCUSSION

The added value of this research lies in the structuring of a model (and tool) that can be used by SMEs. While there is a lot of literature on the fact that SMEs are lacking in the cybersecurity aspect [27-33], there is nothing that solves this problem. These researches indicate that the problems lies within the scope, resources and knowledge that are required for the current methods. However, this problem still remains not solved in the scientific literature. This research focusses on determining what is necessary in order to create a tool that fits SMEs and thus overcomes the limitations that are already found in existing literature. It is important to notice that this is in no way an exhaustive set of requirements, as not all sectors or different kinds of SMEs are included in the research. However, it does provide a good start on which cybersecurity assessments for SMEs can be build.

The second part that adds value is the model that calculates the risks for the different components. While there were existing methods available, they were not suited for SMEs due to complexity reasons. For this reason, the adaption of these models creates a simplification that can be used specifically for SMEs. This is the reason that TRESPASS is used as a basis [18]. The essence of TRESPASS is the same, but SMEs require a simplified version for their use. This is combination with the adapted Gordon-Loeb model [25] makes it a suited risk calculation method specifically for SMEs.

LIMITATIONS

Even though the research brings forward meaningful results, compromises have been made in order to stay within the scope and time-limit.

Concerning the model that determines where the risks in the system lie, this has not been validated by means of a case study. To do a case study, it would require the full cooperation of an SME, a lot of data concerning possible breaches or attacks, a complete mapping of their IT structure and a long period of time to confirm results. These were all aspects that would not fit in the scope of this research, therefore a case study is not done and the

validation is done with the help of experts. In this validation session, different aspects came forward that additions to the model could be done, but that this should be done carefully in order to ensure the accessibility for SMEs.

Furthermore, the rules that fill the knowledge base are not exhaustive. This is a clear limitation of the working of the tool right now, but is not limitation for the theory and workings behind the model and tool. While the tool delivered for this thesis merely serves as a proof of concept, with the expansion of the knowledge base, the tool could be more widely used. Again, this is a clear choice in this research, as the creation of a big knowledge base would have been out of the scope and would not contribute to the validation of the proof of concept.

RECOMMENDATIONS AND FUTURE WORK

Based on the conclusion, discussion and limitations there are recommendations for future research.

With the end product of this research, a validation by the means of a case study should give insights in the actual effectiveness of the model and tool. What kind of effect does it have on SMEs and are those positive effects? This will also give insights in whether the model should be expanded or not. When using the model with SMEs in a real situation, the need for extra components will become clear.

As mentioned in the limitations, the perspective of the attacker should be incorporated (implicitly) in the model. In the expert session it was addressed that one solution to keep it generic (and thus accessible) is the implementation of what the industry of the SME is. This is more specific than the implementation now, but still keeps it generic enough to be used by SMEs. The structure of the model and tool are created in such a way that the implementation of an attacker profile is easy to do, however, it should be confirmed in future research how to take this into the calculations.

Because of time limitations, the determining of the rules is a brief process in this research. In future research this knowledge base could be extended in such a way that the

risk assessment process will be more complete. With the addition of extra rules, this tool could go in production and could be tested by real SMEs.

REFERENCES

- [1] CBS. (2016). *Acht procent van de Nederlanders nooit op internet*. Retrieved from: <https://www.cbs.nl/nl-nl/nieuws/2016/22/acht-procent-van-de-nederlanders-nooit-op-internet>.
- [2] Eurostat, "Digital economy and society statistics - households and individuals," 2016, Retrieved from: http://ec.europa.eu/eurostat/statistics-explained/index.php/Digital_economy_and_society_statistics_-_households_and_individuals#Internet_usage, Accessed on: 08-06-2017.
- [3] Google. (2017). *Google Trends - Cyber Security*. Retrieved from: <https://trends.google.nl/trends/explore?q=cyber%20security>. Accessed on 08-06-2017.
- [4] Cambridge University, "Cambridge Business English Dictionary," in *Cambridge Business English Dictionary*, C. U. Press, Ed., ed. Cambridge: Cambridge University Press, 2011, p. 958.
- [5] T. R. Peltier, *Information Security Risk Analysis*. Boca Raton: Auerbach Publications, 2005.
- [6] J. A. Jones, "An Introduction to Factor Analysis of Information Risk (FAIR)," in "Risk Management Insight," 2005.
- [7] CBS, "De staat van het MKB 2015," 2015, Retrieved from: <https://www.cbs.nl/nl-nl/publicatie/2015/48/de-staat-van-het-mkb-2015>, Accessed on: 07-03-2017.
- [8] Eurostat, "Dependent and independent SMEs and large enterprises," 2015, Retrieved from: http://ec.europa.eu/eurostat/statistics-explained/index.php/Statistics_on_small_and_medium-sized_enterprises#Main_statistical_findings, Accessed on: 08-06-2017.
- [9] Interpol, T. Nipo, and Capgemini, "Cybersecurity in het MKB," 2015, Retrieved from: https://www.interpol.nl/~media/files/ebook_cybersecurity_in_het_mkb.pdf, Accessed on: 07-03-2017.
- [10] NCSC, "Cybersecuritybeeld Nederland 2016: Beroepscriminelen steeds groter gevaar voor digitale veiligheid in Nederland | NCSC," 2016, Retrieved from: <https://www.ncsc.nl/binaries/content/documents/ncsc-nl/actueel/cybersecuritybeeld-nederland/cybersecuritybeeld-nederland-2016/1/CSBN2016.pdf>, Accessed on: 07-03-2017.
- [11] KPMG, "Small Business Reputation & The Cyber Risk," 2015, Retrieved from: <https://home.kpmg.com/content/dam/kpmg/pdf/2016/02/small-business-reputation-new.pdf>, Accessed on: 07-03-2017.
- [12] I. Ilvonen, "Information security management in Finnish SMEs," in *Proceedings of the 5th European Conference on Information Warfare and Security, Helsinki, Finland, 1-2 June 2006*: Academic Conferences Limited, 2006, pp. 161-168.
- [13] Ponemon Institute LLC, "2016 State of Cybersecurity in Small & Medium-Sized Businesses (SMB)," 2016, Retrieved from: https://keepersecurity.com/assets/pdf/The_2016_State_of_SMB_Cybersecurity_Research_by_Keeper_and_Ponemon.pdf, Accessed on: 07-03-2017.
- [14] T. Kurpjuhn, "The SME security challenge," *Computer Fraud & Security*, vol. 2015, no. 3, pp. 5-7, 3// 2015.
- [15] B. Blakely. (2002). *Lock IT Down: Consultants can offer remedies to lax SME security*. Retrieved from: <http://www.techrepublic.com/article/lock-it-down-consultants-can-offer-remedies-to-lax-sme-security/>. Accessed on 03-04-2017.
- [16] D. Kelleher. (2009). *SME security: SME mindset must change*. Retrieved from: <https://www.scmagazine.com/sme-security-sme-mindset-must-change/article/555835/>. Accessed on 03-04-2017.
- [17] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Quarterly: Management Information Systems*, Article vol. 28, no. 1, pp. 75-105, 2004.
- [18] The TRESPASS Project. (2017). *The TRESPASS Project*. Retrieved from: <https://www.trespas-project.eu/>. Accessed on 09-10-2017.
- [19] B. Schneier, "Attack trees," *Dr. Dobbs's journal*, vol. 24, no. 12, pp. 21-29, 1999.
- [20] T. R. Ingoldsby, "Attack tree-based threat risk analysis," *Amenaza Technologies Limited*, pp. 3-9, 2010.
- [21] G. McGraw and J. Viegas, "Building secure software," in *RTO/NATO Real-Time Intrusion Detection Symp*, 2002.
- [22] Smith, Reid G., "Knowledge-Based Systems - Concepts, Techniques, Examples", ed, 1985.
- [23] K. Goztepe, "Designing fuzzy rule based expert system for cyber security," *International Journal of Information Security Science*, vol. 1, no. 1, pp. 13-19, 2012.
- [24] K. Ilgun, R. A. Kemmerer, and P. A. Porras, "State transition analysis: A rule-based intrusion detection approach," *IEEE transactions on software engineering*, vol. 21, no. 3, pp. 181-199, 1995.
- [25] L. A. Gordon and M. P. Loeb, "The economics of information security investment," *ACM Trans. Inf. Syst. Secur.*, vol. 5, no. 4, pp. 438-457, 2002.
- [26] L. A. Gordon, M. P. Loeb, and L. Zhou, "Investing in Cybersecurity: Insights from the Gordon-Loeb

- Model," *Journal of Information Security*, vol. 7, no. 02, p. 49, 2016.
- [27] V. Dimopoulos, S. Furnell, M. E. Jennex, and I. Kritharas, "Approaches to IT Security in Small and Medium Enterprises," presented at the 2nd Australian Information Security Management Conference, Perth, 2004.
- [28] A. Fielder, E. Panaousis, P. Malacaria, C. Hankin, and F. Smeraldi, "Decision support approaches for cyber security investment," *Decision Support Systems*, vol. 86, pp. 13-23, 2016.
- [29] J.-Y. Park, R. J. Robles, C.-H. Hong, S.-S. Yeo, and T.-h. Kim, "IT Security Strategies for SME's," *International journal of software engineering and its applications*, vol. 2, no. 3, pp. 91-98, 2008.
- [30] S. Pritchard, "Navigating the black hole of small business security," *Infosecurity*, vol. 7, no. 5, pp. 18-21, 2010/09/01/ 2010.
- [31] I. Lopes and P. Oliveira, "Implementation of Information Systems Security Policies: A Survey in Small and Medium Sized Enterprises," in *WorldCIST (1)*, 2015, pp. 459-468.
- [32] A. Santos-Olmo, L. Sánchez, I. Caballero, S. Camacho, and E. Fernandez-Medina, "The Importance of the Security Culture in SMEs as Regards the Correct Management of the Security of Their Assets," *Future Internet*, vol. 8, no. 3, p. 30, 2016.
- [33] L. E. Sánchez, A. Santos-Olmo, E. Fernández-Medina, and M. Piattini, "Security Culture in Small and Medium-Size Enterprise," in *ENTERprise Information Systems: International Conference, CENTERIS 2010, Viana do Castelo, Portugal, October 20-22, 2010, Proceedings, Part II*, J. E. Quintela Varajão, M. M. Cruz-Cunha, G. D. Putnik, and A. Trigo, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 315-324.