Erasmus University Rotterdam
Economics & Informatics
Erasmus School of Economics

Master Thesis

The Security Risks of Mobile Payment Applications Using Near-Field Communication

Author:
Raymond Vermaas
Studentnumber 322126
info@raymondvermaas.nl

Supervisor:
Dr. Tommi Tervonen
tervonen@ese.eur.nl

Co-Reader:
Dr. Yingqian Zhang
yqzhang@ese.eur.nl

Thesis Coach:
Drs. Johanneke Siljee
TNO Technical Sciences
johanneke.siljee@tno.nl

Thesis internship at:
TNO innovation for life

March 20, 2013

Abstract

In this thesis the security landscape for near-field communication (NFC) payment applications on smartphones is investigated. The proposed model in this thesis is used to perform a robust risk assessment on a set of NFC related attack and fraud scenarios. The impact of the fraud scenarios is evaluated using information of a real-world NFC payment provider. The attack scenarios are presented to a group of experts in an expert elicitation, who gave their opinion on the likelihood of the attack scenarios based on evaluation criteria of an existing risk assessment method. The experts are also asked to answer a set of calibration questions to gain insight in the knowledge and certainty of the expert opinions. These results are aggregated into a discrete probability distribution. This distribution is used as input for a SMAA-TRI model that results in set of risk class acceptability indices for the different attack scenarios. This resulted for the NFC attack scenarios in two high risks scenarios, three medium risk scenarios and five low risk scenarios.

Raymond Vermaas - 322126 March 20, 2013

Acknowledgments

This thesis would not have been possible without the support of many people. First, I would like to express my sincere gratitude to my thesis supervisors Tommi Tervonen and Yingqian Zhang, who have dedicated their time and effort towards my research. Second, I would like to express my very great appreciation to my internship supervisor Johanneke Siljee from TNO for all the support, guidance, and valuable and constructive critique she provided me with throughout my thesis internship at TNO. I would also like to offer my special thanks to the staff of the TNO ICT Security expertise group for their hospitality and the opportunity they have given me to perform my thesis work at their company. Furthermore, I would like to thank all the experts that took the time and effort to fill out the expert elicitation. I especially want to thank the experts that provided me with additional comments on my elicitation. Last but not least, I want to thank my parents and brother, for all their love and guidance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Near-field communication (NFC) enables short-range communication between devices and tags [1, 2]. It can be used to make small payments between a mobile phone and a payment terminal or to receive extra information about a subject through a NFC tag. In the last few years, NFC has become increasingly popular. It is anticipated that NFC payments will hit mass market in 2015 [3], and that the technology will have a 448 million users and will raise the yearly total transaction value for mobile payments to 617 billion dollar by the end of 2016. Also, most of the high-end mobile phones from the major phone manufacturers already contain NFC chips. This increase in NFC-enabled hardware already gave rise to the first NFC payment applications, such as Google Wallet. This application has received much attention, since it stored Radio Frequency Identification (RFID) credit cards inside a mobile phone, enabling the user to use their phone for payments. Also, in the Netherlands there is interest for NFC payment applications. In 2008, the Dutch company Payter ran a large NFC payment pilot in Rotterdam to test this new technology. Until recently, a joint-venture of the major Dutch telecommunication providers and banks investigated a large-scale introduction of NFC payment services in the Netherlands.

However, vulnerabilities in new technologies in the payment and banking industry are an interesting target for hackers and criminal organizations. One successful attack may destroy the reputation of a payment provider or bank and possibly the loss of trust in the new technology. A payment provider should therefore pursue the goal of building a safe and trusted NFC payment method. A risk assessment is an important tool in achieving this goal, since it gives insight in the vulnerabilities, threats and risks.

1.1 Problem Statement

Since NFC is used for security sensitive applications, like payments and access control, the security of NFC is important. However, the protocol itself contains few security measures. This leaves the responsibility of security to the NFC application developer. Creating awareness on

the vulnerabilities and possible attacks of NFC in smartphones is therefore an important step in creating more secure applications. There is already research available on the vulnerabilities in NFC [4, 5], but these attacks are still quite theoretical. Companies and application developers are usually more interested in practical and tangible vulnerabilities and the likelihoods of the attacks exploiting these vulnerabilities. A good quantitative risk assessment may contribute to the awareness among companies and application developers about the vulnerabilities in NFC.

An integral in performing risk assessments is the task of quantifying the risks. The evaluation of information security risks is usually done using a standardized security evaluation framework. These framework are built upon best practices in the security domain and are mainly focused on the ease of use. However, the uncertainty involved with the quantification of risks is almost never taken into account in such frameworks, although the imprecision of risks is of critical importance when dealing with the introduction of new technologies, such as NFC. Another problem is that a risk assessments is usually performed by a limited number of experts, causing biased risk assessments. By using more experts one should get a more accurate assessment, since this will incorporate different opinions into the risk assessment, making the assessments less subjective. An expert elicitation and a multi-criteria decision analysis method might possibly solve these problems by incorporating more expert opinions in the risk assessment and by offering a theoretically solid framework for aggregating the assessment's multiple risk dimensions.

1.2 Research question

The research question for this thesis is:

What are the risks in payment applications for NFC-enabled smartphones?

This research question results in the following sub questions:

- *What are vulnerabilities in using NFC-enabled smartphones for payment applications?*
- *What is the likelihood these vulnerabilities will be exploited?*
- *What is the impact when these vulnerabilities are exploited?*
- *What are the expert views on the risks connected to these vulnerabilities?*

1.3 Scope

Near-field communication has many possible applications, from smart posters to payment applications. This also makes the security landscape rather large. However, most research in the

NFC security landscape is done on payment applications. This is not that strange, since most NFC applications in development right now are payment applications (e.g. Google Wallet). In addition, the perceived risk in payment applications is much higher than in, for example, smart posters. With this knowledge, it is only logical to use NFC payments as the domain for my research.

In the research for vulnerabilities in NFC-enabled smartphones, the focus will mainly be on designing practical ways of exploiting the relay attack and card emulation. Finding practical attacks in these two areas might create awareness of the security issues in using NFC among application developers, application architects and publishers. Furthermore, this research will try give an overview of the NFC payments security landscape by presenting a set of attack and fraud scenarios. This might contribute to more secure implementations of NFC in payment applications. The risk assessment performed in this thesis will therefore focus on the security risks incurred by the publisher of NFC payment applications (e.g., banks and other payment service providers). The vulnerabilities and the scenarios found in the first part of the thesis serve as a basis for the risk assessment.

1.4 Motivation

Several companies are investigating the possibility to integrate near-field communication in their business. With the introduction of (NFC-enabled) applications on the market, it is common to perform a risk assessment on the application to be released. The TNO ICT Security group expects a demand for risk assessment on a mass-market NFC-enabled smartphone application within the next couple of years. In order to give a fair assessment, they need to gain insight in the threats and the risks facing NFC applications in smartphones. The aim of this research is to provide this insight.

1.5 Structure

This thesis starts with the Chapter 1. This chapter presents the problem statement and the research question. It also provides the reader with some background information on NFC, mobile payments and risk assessment. In Chapter 2, the methods for the multi-criteria decision analysis and risks assessment are introduced. The elicitation technique for handling multiple experts and the preference aggregation technique are also discussed is this chapter. Next, Chapter 3 proposes a model for risk assessment with an expert elicitation that makes use of the methods presented in Chapter 2. In Chapter 4, the model from the previous chapter is applied to a case study on the security of using NFC-enabled smartphones for payments. Therefore, this chapter will introduce the reader with the different attacks and fraud scenarios for this case and other implementation decisions. The results of the risk assessment model for the NFC case study are presented and interpreted in Chapter 5. This thesis ends with the conclusion in Chapter 6. It answers the

research questions and give insight in the risks for NFC payment applications on smartphones. The last chapter will also contain a discussion on the security of NFC payments and the security evaluation model.

1.6 Background

1.6.1 Near-Field Communication

NFC is a short-range wireless communication protocol [1, 2, 6] between devices, like smartphones. The signal carry range of NFC is between 0 and 10 cm, although ranges up to 66 meter are reported when using special directional antennas [7]. NFC is part of the Radio-Frequency Identification (RFID) technology, and is described in ISO 18092 and ISO 21481 [1, 2]. Just as other RFID technology, it uses the 13.56 MHz radio frequency band to communicate. This allows it to communicate with other RFID technology, like ISO 14443 [8] cards. It also has some differences from other RFID technology.

The main difference between NFC and other RFID technologies are the two modes of NFC: the 'active' mode and the 'passive' mode [1, 4]. In the passive mode, an NFC device behaves as an RFID token, which allows it to be read by an NFC reader. It can also emulate other RFID technologies in this mode, like ISO 14443 (Mifare cards), ISO 15693 and FeliCa cards, so it can communicate with legacy hardware that does not support the ISO 18092 NFC standard. Furthermore, this mode allows the NFC device to act like reader, so it can power and read other RFID tokens. In the active mode, it supports peer-to-peer transaction between NFC devices. This works by bringing two NFC devices [2] near each other (0-10 cm), where one of the NFC devices sets up a Radio frequency (RF) field to transmit data. The other device turns off its own RF field and receives the data by the first device. When the second device wants to send data, the roles of devices switch. The NFC messages are sent in the NDEF format [9]. The NDEF messages can be compared to MMS on the mobile network, since both technologies provide the means to transmit rich data over a wireless protocol. A NDEF message can contain data, like text, website links and vCards.

The bandwidth of NFC is, with a bandwidth between 106 and 424 kbps, rather low compared with other wireless communication channels, such as WiFi, GSM, Bluetooth, and UMTS. On the other hand, NFC has a connection set-up time of only a couple of milliseconds, which is quite fast in comparison with the previously mentioned technologies. This makes NFC useful for a quick transmission of small amounts of data. Near-field communication is mainly specified in four ISO standards:

- ISO 18092, Near Field Communication Interface and Protocol (NFCIP-1) [1] is the basis of the near-field communication technology. On the low-level, it sets the NFC requirements for the transfer speeds, codings, modulation schemes, and frame format of the RF interface. It also describes some NFC low-level processes, like initialization schemes, conditions required

for data collision control during initialization of a new NFC connection and the active and passive modes of NFC.

- ISO 21481, Near Field Communication Interface and Protocol (NFCIP-2) [2] specifies the detection and selection of RFID cards, so the NFC device knows with which type of tag it is dealing with.
- ISO 28361, Near Field Communication Wired Interface (NFC-WI) [10] describes the signal processing between the antenna (front-end) and the NFC controller (transceiver).
- ISO 16353, Front-end configuration command for NFC-WI (NFC-FEC) [11] describes the NFC-WI commands for information exchange between the NFC wired interface and the NFC front-end.

The RFID ISO standards (like ISO 14443 [8]) also contribute to the NFC technology. The European Computer Manufacturers Association (ECMA) is also working on new NFC standards, like NFC-SEC [12] that provides a secure channel for NFCIP-1. However, the secure channel described in this standard only works in the peer-to-peer modes of NFC. Therefore, its implementation possibilities are somewhat limited.

Applications

Mobile phone manufacturer Nokia introduced the first mass-market NFC-enabled smartphone, the Nokia 6131, back in 2007 [13]. Nowadays, almost all major mobile phone manufacturers, like Research In Motion, HTC and Samsung, enable NFC in their top-range models. The growing number of available NFC-enabled smartphones makes it interesting for companies to apply NFC in their business.

One application of NFC are smartposters. These are regular advertisement posters that also contain a NFC tag. Owners of a NFC devices can scan the tags to receive extra information, like the website of the advertiser or location-based offers, for the advertised product. In the United Kingdom, movie posters in cinemas are fitted with NFC tags [14]. These tags provide extra information about the movie and additional digital content.

Another application is the storage of e-tickets for events or for public transport. In this case the user of NFC buys a ticket that is stored on the NFC device. When entering the event or public transport, the user simply presents the NFC device to an NFC reader. This application can reduce waiting time for ticket offices and ticket machines. The Dutch soccer club Roda JC ran a pilot [15] with NFC tickets, where the seasonal tickets of 50 supporters were replaced with NFC e-tickets on their smartphones.

One of the most interesting applications is the ability to perform mobile payments over NFC. The implementation is similar to that of the e-tickets, where the NFC device acts as a credit card, or a virtual wallet is stored within the phone. The user swipes the device in front of a payment terminal to complete the transaction. The largest mobile phone NFC payment pilot in Europe

was held in the Netherlands by the Dutch company Payter in 2008 [16]. This company provided NFC-enabled phones with a virtual wallet to 2200 participants in Rotterdam and installed a number of payment terminals at retailers and parking garages. They also placed smart posters in the city center of Rotterdam to test the marketing potential of NFC tags. The aim of this pilot was to test the technology, gain insight in the user experience of NFC and boost the user adoption of NFC.

The user adoption of NFC payments with smartphones in Japan [17] is already quite far. Before the introduction of NFC, Japan was a cash-based society, since other payment methods, such as credit cards, never caught on in Japan. NFC allowed the Japanese people to pay faster. It was first introduced in vending machines, kiosks and train tickets in Tokyo. Nowadays almost every store in Japan supports payments by NFC. The current NFC trend in Japanese companies is to have a NFC-based customer loyalty program, where a single mobile device is used to store the participation details of multiple loyalty programs from different companies. The NFC development in Japan is fueled by Sony with their FeLiCa cards.

The latest development in NFC payments is Google Wallet [18] for the mobile operating system Android. This system is available for all NFC-enabled Android devices and made in collaboration with the credit card provider MasterCard. The application allows users to store all RFID-enabled credit cards in the application and use their phone instead of the cards for financial transactions. Google also offers storage for their own prepaid payment cards in the wallet and a single sign-on interface for online shopping.

Another NFC development by Google is Android Beam, which was introduced in version 4.0 of its mobile operating system Android. This application uses NFC's peer-to-peer technology to quickly share digital business cards (vCards) and bookmarks between NFC-enabled smartphones.

Security

The technical specification of NFC defines very few security measures. The implementation of security measures is left to the developer on the application level of NFC. The NFC Forum [19] defined three functionalities required for the safe execution of NFC applications on mobile phones, as shown in Figure 1.1. The NFC controller offers the low-level interface for transmitting and receiving NFC data, consisting of three stacks: the card emulation, the read/write and the peer-to-peer stacks. The second functionality is the Application Execution Environment (AEE). The AEE is the default execution environment for applications on the phone, like the dialer and the contact application. NFC applications, like a peer-to-peer app or a simple tag app, are also allowed to run in the AEE.

The last functionality is the Trusted Execution Environment (TEE). The TEE provides a secure and trusted environment for data storage, application storage and application execution. By placing an application firewall around certain applications, the applications are protected from any malicious activity the phone. Payment applications, such as Google Wallet, are supposed to run in the TEE. This TEE is provided by a secure element (SE). A secure element can be

implemented in hardware, like a micro SD card, SIM card or in an embedded smart card, or in the software (Soft-SE) as a designated memory partition.

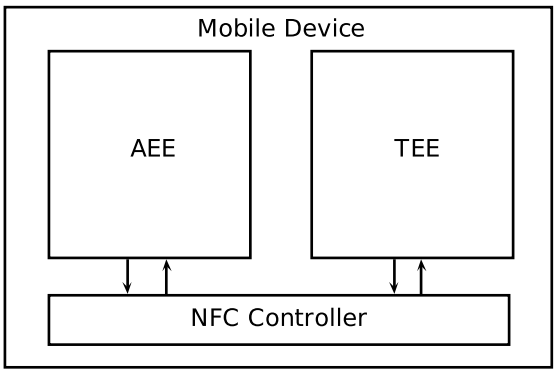

Figure 1.1: Mobile NFC environment

1.6.2 Attacks on NFC

A communication protocol that is used for payments, with very little security by design, is extremely interesting for the hacker community. Also the number of hacks in the RFID technology [20, 21], like unauthorized reading, falsification of content, eavesdropping, and attacks on back-end systems, adds to the interest of finding possible exploits with NFC.

Hoepman and Siljee [6] describe eight issues in the NFC security landscape, which are:

1. All NFC devices are readers and writers;
2. All NFC devices can emulate a tag;
3. The range of NFC is not enforced;
4. NFC is a gateway to the attached device;
5. Multiple NFC applications are run on a single device;
6. NFC lacks a security standard;
7. The possibility to unintentionally connect to an NFC device;
8. Privacy of the user is not protected, because data send over NFC is not encrypted.

The authors also propose solutions for the security issues. They start with solutions that affect both RFID and NFC:

- The device should be stored in a Faraday Cage when not in use. This will prevent unauthorized access to the NFC device.

- Only read-only information should be put on the card (Tag ID). It is also necessary to sign the read-only data on issuing to assure the authenticity of the tag.
- For privacy reasons a tag ID should be randomized to prevent tracking of the user.
- Communication between devices should be encrypted to increase the privacy of the user and reduce eavesdropping.
- The tag reader should authenticate tags and validate the content on the tags.

The authors also provide four NFC-specific solutions:

- Use the user interface of the device to ask for user confirmation.
- Use the computing power of the NFC device for encryption.
- Use the secure element of a NFC device to safely execute applications and store data.

Card emulation

Roland [5] investigated the security of software card emulation in NFC-enabled mobile phones from the manufacturer Research In Motion. The card emulation allows RFID tags to be stored on the NFC device and it also enables peer-to-peer communication with older NFC devices without peer-to-peer support. At first, card emulation was only available on the secure element that was only available for large trusted developers. Later, card emulation without a secure element was also enabled for other developers to promote the NFC technology. This choice made it hard to secure sensitive data on a NFC device. It also grants malicious users the possibility to perform card emulation and relay attacks. In a relay attack a NFC device is used to relay the information stored on a tag or RFID card onto a card reader. In other words, the NFC device is pretending to be the tag.

Relay attack

The possibility of a relay attack with NFC was further investigated by Francis et al. [4]. They developed a practical relay attack on contactless transactions with NFC-enabled smartphones. The attack requires close proximity to the RFID card of the victim with a NFC-enabled smartphone, the proxy reader. Another NFC-enabled smartphone, the proxy token, is placed near the reader. In the attack, the proxy reader reads the information on the tag. The information is then sent from the proxy reader to the proxy token using a wireless communication protocol, like Bluetooth, UMTS, or WiFi. The proxy-token identifies itself to the card reader as the tag using the information received from the proxy reader. If needed, the proxy token and the proxy reader will relay any additional information between the tag and the card reader. The card reader will accept the proxy token as the tag. This relay attack circumvents any decryption of encrypted communication between the tag and the card reader, since it is able to just relay the encrypted

information. The only challenge in this relay attack are the timing issues. Relaying the RFID card information over a Bluetooth connection and the processing in the two NFC devices adds a couple of extra seconds. The authors discovered this was not an issue, since they could delay the transaction up to 35 seconds without encountering any time-outs. They also propose some countermeasures to prevent such relay attacks. For the contactless platform (or reader), they propose to put strict timing restrictions set by the reader, so there is insufficient time between responses for performing such relay attacks. Furthermore, they propose to implement distance bounding, where the maximum round-trip-time of a response is restricted by the maximum physical distance between the reader and the card. For the mobile platform, the authors propose that a mobile device should provide proof of its physical location in the response by sending, for example, its GPS coordinates. Another proposed countermeasure is to provide better authentication mechanisms in the NFC communication. For example, a check if the reader is communicating with the same device during the entire transaction.

Lee [22] turned the relay attack from Francis et al. [4] into a simple application for the mobile operating system Android. Besides the proxy-reader and proxy-token modes, the application has two other modes, the skimming mode and the spending mode. In the skimming mode, the application acts like a payment terminal, so it can fake a transaction with a payment tag. In the spending mode, the application can replay the faked transaction, with the payment tag, to a payment terminal.

Fuzzing

Miller [23] performed a fuzzing test on the NFC implementation in Android and Nokia Mobile OS Meego. In a fuzzing test small changes and insertions are made to the input messages sent to a program or device. Miller limited the fuzzing test to the Mifare UltraLight, Mifare DESfire and the P2P mode (LLCP protocol) of NFC. A total of 52 362 test cases were performed on Android and 34 852 on Meego. The test cases can be divided in low-level fuzzing tests and high level fuzzing tests. On the low-level, the focus was on the modification of NDEF messages and tag emulation of the Mifare UltraLight and the Mifare DESfire. Besides causing a lot of exceptions in the NFC application, Miller found nothing interesting on low-level fuzzing. With high level fuzzing, the author tried to exploit known bugs in the software. In Android, he found the OS requires the phone to be unlocked and its screen to be on before NFC is enabled. However, when the phone is unlocked, no user interaction is required to launch applications (e.g., the browser) using NFC. This makes it possible to send the user to a malicious website, when he is using NFC, for example by placing a malicious tag near a payment terminal. Nokia Meego had a more severe bug. In Meego it is possible to initiate a Bluetooth pairing using NFC, which by default, requires no user interaction. This allows a malicious user to access the victims device from a longer distance.

NFC as attack gateway

Besides using NFC as target, it can also be used as a gateway to the NFC device. Mulliner [24] showed an attack where NFC was used to install a worm on a smartphone. In order to achieve this the stickers on smartposter were replaced with a sticker containing a URL linking to a malicious application. By making use of new lines and spaces, Mulliner was able to show the victim the legitimated title and URL of the original tag, while the original URL did not fit on the screen any more. The unsuspecting victim would just install the malicious application, thinking it came from a trusted source and allowing the attacker to access the mobile phone. The author also showed that this same attack could be used to make user call or text to premium numbers by sending sms-type or tel-type NDEF messages over NFC.

The possibilities of this attack were once again shown by McClure [25]. In this case, a phishing attack was performed using malicious tags. In his experiment the NFC tags on smart posters from the Red Cross were replaced by malicious tags. The tag contained a URL to a fake donation page from the attackers, which looked like the real donation page of the Red Cross. While people believed they donated to the Red Cross, they actually donated money to the attackers.

A more practical approach was shown by Bargaonkar [26]. Bargaonkar presented a vulnerability that made it possible to execute malicious commands on Android version 4.0.4 and lower. The victim would scan a NFC tag, which would send him/her to a malicious URL. This malicious URL contains a so-called USSD code, which is a code that can execute applications or actions via the phones dialer application. These codes allowed the attacker to reset the phone to its factory defaults or permanently block the SIM card. This could be performed without requiring any interaction from the victim.

1.6.3 Risk Assessment

In everything we do, we are subjected to certain risks. However, the actual definition of risk is quite hard to specify. Economists define risk as the difference between predefined objectives and actual performance, whereas the Oxford dictionary defines it as "a situation involving exposure to danger".

In information security (IS), risk can be defined as "the potential that a given threat will exploit vulnerabilities of an asset or group of assets and thereby cause harm to the organization" [27]. It is also often depicted as a simple equation:

$$Risk = Likelihood\ of\ an\ incident \times Impact\ of\ an\ incident \tag{1.1}$$

A risk emerges from a vulnerability, which is a weakness in the asset (or system) that may be exploited by a threat. An example of a vulnerability is the lack of access control in an application containing sensitive information. A threat is defined as a potential cause of an incident that may result in harm to the system or organization. Threats consists of three parts: a target asset (or system), a threat agent, and the actual action of the threat agent initiating the threat.

A threat in IS influences the main qualities of information security: Confidentiality, Integrity and Availability or simply CIA. Confidentiality is the degree to which access to information or functionality is restricted to the right user. Integrity is the degree to which data or functionality is correct. Availability is the degree to which data and functionality is available for users at the right moment.

The threat and vulnerability are part of the *likelihood of an incident*, which is the probability a certain vulnerability is exploited by a threat. The ETSI TISPAN [27] defines 5 domains that contribute to the likelihood of an attack or incident:

1. The time required to prepare and perform the attack;
2. The expertise required by the attacker to perform the attack;
3. The (specific) knowledge about the system required to perform an attack;
4. The opportunity window needed to perform the attack;
5. The need for special equipment required to perform an attack.

The *impact of an incident* is the severity of the result of a security incident caused by a threat. The impact involves, for example, loss of data, reputation damage, costs due downtime or incurred by the investigation, financial loss (fraud and missing payments) and costs of required infrastructure changes.

The threats, vulnerabilities and risks for a specific information system are evaluated in a risk assessment. The main purpose of a risk assessment is to identify which risks are too high and need to be mitigated [28]. This can be done by one of three types of risk assessments. The first is the checklist which is a simple list for identifying the only most common risks. The second is a qualitative risk assessment where the risks of an information system are calculated. Last, is the quantitative risk assessment where risks are turned into measurable criteria such as money. The quantitative risk assessment is the most detailed type of risk assessment.

Like mentioned before, a risk assessment helps decision makers identify which risks need to be mitigated. This is possible by implementing countermeasures that reduce the likelihood or impact and thereby the risk of a security incident. However, when the risks are reduced with a countermeasure, the cost of the information system increases with the cost of the countermeasure. A countermeasure might affect published standards, implementation in the user community, disruption in operation, regulatory compliance and market acceptance [27]. Given these costs, there exists a trade-off between a reduction in risk on one side and the cost of the countermeasure on the other side. This is usually resolved by quantifying both the cost of the countermeasure and the expected cost of the risk and performing a cost-benefit analysis. The risk that is not reduced by a countermeasure is called the residual risk.

Besides the ETSI TISPAN TVRA, there are risk assessment methods available that are often used to assess security risks. Most of these method, like OCTAVE [29], have a qualitative nature,

since they require less effort to perform. However, there also other quantitative methods available for security risk assessments. One of those methods is the Common Vulnerability Scoring System developed by the Forum of Incident Response and Security Teams (FIRST) [30]. This method divides the evaluation criteria over three different metric groups. The first group are the base metrics, which are the basic criteria of a vulnerability that do not change over time or between user environments. The base metrics group consists of the access vector, the access complexity, authentication, confidentiality impact, integrity impact and availability impact. The second group are the temporal metrics, which are the criteria of a vulnerability that unlike the base metrics do change over time. The temporal metrics group consists of exploitability, remediation level and report confidence. The last group consists of the environmental metrics, which are an optional set of criteria that asses how the user environment is affected by a vulnerability. The environmental metric consists of the collateral damage potential, the target distribution and the security requirements for the integrity, availability and confidentiality. The scores for the different groups are calculated using a set of equations. In the calculation of the scores, the base metrics and the temporal metrics have a dependency, whereas the environmental score is calculate separately from the first two metric groups. Although, this risk assessment method provides a quantitative way of calculating the risks, it does not provide a way to classify the risks in different risk classes. Furthermore, the method is focused on the risk assessment of established IT systems, rather than an evaluation of the security risks in new technologies. These two properties of the method makes it less suitable for the proposed model and the performed case study in this thesis.

1.6.4 Multi-criteria Decision Analysis

Multi-criteria decision analysis (MCDA) is a technique of evaluating how to rank, sort or classify different alternatives based on a set of criteria [31]. When the criteria and alternatives are considered in the decision making process, it is called multi-criteria decision making (MCDM). We use MCDM every day, for example to decide what to eat for dinner. The possible choices are evaluated on price, taste and easiness of cooking. For this example, complex models are not necessary; the human mind is perfectly able to choose the best alternative without models. However, if the decisions become more complex, with larger numbers of criteria or alternatives, MCDA models are quite helpful.

An example of a more complex decision is a company that wants to build a new office. For an office multiple criteria are important, like the fixed costs, the average distance from the customers, the accessibility, and the maintenance costs. The company selects multiple possible building sites for the new office and expressed their preferences for different criteria. By averaging the preferences given by the different decision makers in the company, the importance weights of the different decision criteria can be obtained. By entering the criteria, the building sites and the weights into a multi-criteria decision model, the outcome of the model can be used to support the decision making in the company.

An information security risk assessment is also an application for decision analysis. The

calculation of risk, as described in Section 1.6.3, is also subjected to a number of criteria for likelihood and impact. Multi-criteria decision analysis can be used to rank the different risks (i.e., alternatives) in order to see which risks need to be reduced.

Yang et al. [32] applied their proposed multi-criteria decision method VIKORRUG on the information security risk domain. VIKORRUG is based on VIKOR, a multi-criteria decision aiding method for discrete choice problems with a group of decision makers. VIKOR is designed to solve decision problems with non-commensurable and conflicting criteria. In VIKOR a trade-off is made between maximizing the average performance for the group (i.e., the majority) of alternatives and minimizing the maximum loss for an individual alternative. In the VIKORRUG method (VIKOR for Ranking Unimproved Gap) as proposed by the authors of [32], it is possible assign a specific set of criteria for each alternative instead of all the criteria for each alternative in VIKOR. Another difference between VIKOR and VIKORRUG is on the focus of the methods. Whereas VIKOR is used to returns the best ranking for alternatives, the VIKORRUG is used to show which of the gaps in decisions need to be improved. They demonstrate the use on their approach on information security risk domain. They asked multiple security expert to give the probability of a certain security breach, the consequence of a security breach and the weight of the criteria. The authors used this information to identify which control objectives, like access control and asset management, need to be improved for different projects using VIKORRUG, in order to minimize the gaps between decision makers. The main difference between the method of Yang et al. and the model proposed in this thesis, is that it ranks controls that need to be improved in existing assets or project rather than classifying the security risks.

Wang and Elhag [33] proposed a fuzzy approach for the MCDA technique *Technique for Order of Preference by Similarity to Ideal Solution* (TOPSIS). TOPSIS is a technique for order preference that maximizes the benefit of criteria and minimizes the cost of criteria. The technique ranks alternatives according to their distance to the best and the anti-ideal solutions. The authors introduced their own fuzzy TOPSIS method for MCDM, since existing methods still contain crisp elements or the results were exaggerated. Wang and Elhag use alpha-level sets to calculate the defuzzified values. The alpha level set is an interval of the membership function that is determined for each predefined α-level. An α-level represents a certain degree of membership at a point in the membership function. For example, an interval at an α-level of 0.2 extracts a lower bound value and an upper bound value from the original function at the points where the membership is 0.2. The fuzzy TOPSIS approach in [33] consists of 6 steps:

1. Normalize the fuzzy decision matrix with the predefined weights;
2. Determine the ideal solution and the anti-ideal solution;
3. Calculate the alpha-level sets by setting different α-levels;
4. Calculate the fuzzy closeness of each alternative by solving an NLP program;
5. Defuzzify the fuzzy relative closeness;

6. Rank the alternatives using their defuzzified relative closeness.

The authors apply their method to a bridge risk assessment to determine their maintenance priorities. The risk is calculated using fuzzy numbers in the $likelihood \times impact$ risk formula and used as criteria in this model. They gathered data for five bridges and determined the order in which the bridges require maintenance. The main difference between this method and the proposed model is that this method ranks different projects rather than classifying the risk for a single project or technology.

Chapter 2

Methods

In this chapter the methods are explained that are part of the proposed model. This chapter starts by describing the risk assessment. Section 2.2 explains how a set of ordinal rankings can be aggregated into a single ranking. Section 2.3 explains the employed method for expert elicitation. The chapter concludes with a description of two multi-criteria decision analysis methods that will be used in the proposed model.

2.1 Risk assessment method

In this thesis the European Telecommunications Standards Institute (ETSI) Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN) Threat, Risk and Vulnerability Analysis (TVRA) [27] is used, which is based on the ISO 18045 standard [34]. This standard shall be used to perform an initial risk analysis in Appendix C and will be the basis for the risk analysis approach proposed in this thesis.

2.1.1 Likelihood

The likelihood, according to the ETSI TISPAN TVRA, is build up of five criteria that are required to exploit a vulnerability. A brief explanation of each of the criteria is given below.

1. *Elapsed time* is the time it takes the attacker to identify and exploit a vulnerability in the system. The elapsed time is measured in weeks from less than a day (0 weeks) up to 6 months (26 weeks). The probability of an attacker spending more than 6 months on one exploit is very low, so every exploit with an elapsed time of more than 6 months is considered impossible.

2. *Specialist expertise* refers to the technical knowledge of the attacker. This involves generic knowledge on often used protocols, operating systems and attack methods. The knowledge of the attacker can be classified in one of three categories.

- A layman does not have any technical knowledge and has no expertise about a certain protocol, operating system or attack method;
- A proficient attacker has some technical knowledge and is familiar with some systems;
- Experts have expertise on the software, hardware and protocols used in the system. They are also able to identify new attack methods.

3. *Knowledge of the system* involves the specific knowledge the attacker has about the system. There are four types of information an attacker can have about the system.

   - Public information about the system. This type of information can be gathered through the Internet or other media;
   - Restricted information about the system. This type information is for example information shared between partner organizations under a non-disclosure agreement;
   - Sensitive information about the system. This type of information is only available for people who are concerned with the direct operations of the system within an organization;
   - Critical information about the system. This type of information is only known to key individuals in the development or operations and is controlled under a need to know basis.

4. *Window of opportunity* refers to the time windows it is possible to access the system and exploit a vulnerability. Exploiting the vulnerability may require long continuous access to the system, which increases the chance of detection. Since time is also an important factor in the window of opportunity, this criteria is strongly linked to the elapsed time criteria. Five types of access can be identified:

   - Unlimited access means the attacker has continuous access to the system and there is no risk of being detected;
   - Easy access means the access required for the attack is less than a day and there is little risk of getting detected;
   - Moderate access means the access required for the attack is less than a month and there is a moderate risk of getting detected;
   - Difficult access means the access required for the attack is at least a month and there is a high risk of getting caught;
   - No access means the window of opportunity is too small to perform the attack. The reason for this can be that time needed for the attack is too long (e.g. the asset is no longer exploitable or not sensitive anymore) or can not access the system often enough.

5. *Equipment availability* refers to the availability of hardware, software and other equipment needed to perform the attack. Three types of equipment availability can be identified:

    - Standard equipment is freely available for the attacker. The equipment can be part of the system itself (e.g. a debugger) or can be obtained through open channels, like Internet downloads;
    - Specialized equipment is not freely available, but can be acquired with some effort. The development of some custom attack tools and acquiring moderate amounts of equipment can be classified as specialized equipment;
    - Bespoke equipment is equipment that needs to be specially or produced for which the distribution is controlled or restricted;

As one might notice, these criteria are not independent of each other. For example, if more specialist expertise is available it is possible that the elapsed time will decrease. This results in that an attack scenario is assessed with one specific type of attack in mind, often the most likely attacker.

In order to aggregate the different criteria a numerical value is assigned to the categories of each criteria. A security expert performing the risk analysis chooses a category for each criterion in a specific attack scenario. The sum of the expert choices can be lookup up in a conversion table to determine the final likelihood.

2.1.2 Impact

Whereas the likelihood criteria mainly focus on the likelihood of an attack scenario, the impact is used to assess the impact of a fraud scenario that arises from an attack. Unfortunately, the ISO 18045 [34] or the ETSI Threat, Risk, Vulnerability Analysis [27] do not contain multiple criteria to assess the impact, as is the case for the likelihood of an attack. The ETSI TVRA uses the asset impact to evaluate the impact of a fraud scenario, where an asset is defined as "anything that has value to the organization, its business operations and its continuity" [27]. The asset impact has three levels:

1. **Low** The owner of the asset is not really harmed and the possible damage is low;
2. **Medium** The attack harms customers and providers. The attack can not be neglected;
3. **High** The core of the business is under attack and the damage is severe.

Finally, the impact value and the aggregated likelihood value are combined to give a risk rating of low, medium or high.

2.1.3 Risk Calculation

Section 2.1.1 and 2.1.2 show the qualitative values to rank the likelihood and impact. The model discussed in this thesis requires quantitative values to give an assessment. Luckily, both the ETSI TISPAN TVRA [27] and the ISO18045 [34] also assigned quantitative values to the different choices of the evaluation criteria. In case of the likelihood the quantitative value increases from high to low, as shown in Table 2.1. For example, for the criterion "knowledge of the system",

Table 2.1: Criteria quantitative values

Criterion	**Lowest value**	**Highest value**
Elapsed Time	0 weeks (0)	26 weeks(26)
Specialist Expertise	Layman (0)	Expert (5)
Knowledge of the system	Public (0)	Critical (10)
Window of opportunity	Unlimited (0)	None (Not exploitable - assigned value of 26)
Equipment availability	Standard (0)	Multiple Bespoke (9)

public information is assigned a value of 0 and critical information receives as value of 11. The values of the evaluation criteria for a specific scenario are then summed up to determine the required attack potential of the attacker, as shown in Table 2.2. This rating has five categories

Table 2.2: Conversion from quantitative values to attack potential

Values	**Attack potential**
$0-2$	No rating
$3-6$	Basic
$7-14$	Moderate
$15-26$	High
≥ 26	Beyond high

no rating, basic, moderate, high and beyond high, where no rating is reserved for attacks that can be performed by anyone, and beyond high means the attack is virtually impossible. The required attack potential can be converted to a likelihood rating: no rating and basic receive a likelihood of likely, moderate is classified as possible and the attack potential of high and beyond high are classified as unlikely. The likelihood is assigned values from 1 for unlikely to 3 for likely.

The impact has only 3 possible values and no other criteria. It is also assigned the values 1 (low) to 3 (high) as happens with the likelihood. By multiplying the impact values and the likelihood values one gets the risk values. The possible results for this multiplication and the meaning is described in Table 2.3.

A good example how the risk calculation is performed, can be found in Appendix C.

Table 2.3: Conversion table for risk quantitative risk values to qualitative risk values

Values	**Risk**	**Description**
1 or 2	Minor	No essential assets are concerned. An attack is unlikely and has a low impact on the users.
3 or 4	Major	A significant number of users and some major assets are threatened by this risk. An attack with significant impact is possible.
6 to 9	Critical	Many users and assets are threatened by this risk. It has a high impact and an attack is likely to happen.

2.2 Ordinal ranking aggregation

An ordinal ranking can be aggregated using a distance-based consensus method. In this method the preferences of experts are aggregated to a set of preferences with a minimum distance to the original preferences of the experts. According to Cook [35], a distance function is subjected to six axioms related to social choice properties. These axioms are based on a distance function $d(A, B)$ between expert's priorities and are defined as follows:

- *Axiom 1:* $d(A, B) \geq 0$
- *Axiom 2:* $d(A, B) = d_{cs}(B, A)$
- *Axiom 3:* $d(A, B) + d(B, C) \leq d(A, C)$
- *Axiom 4:* $d(A, B) = d(A', B')$, where A' and B' are A and B subjected to the same mutation in distance. (Invariance axiom)
- *Axiom 5:* If an extra alternative with result A^* and B^* is added on place $n + 1$, which is the same as the alternative on place n with result A and B, then $d(A, B) = d(A^*, B^*)$.
- *Axiom 6:* The minimum positive distance is 1.

These axioms can be applied in a rank-based distance method as proposed by Cook et al. [35, 36]. A rank-based distance also takes into account the degree of disagreement and makes use of two vectors, $(P^+(j))_k$ and $(P^-(j))_k$, where n is the number of alternatives:

$$(P^+(j))_k = \begin{cases} 1, \text{if alternative } k \text{ is ranked in a lower position than alternative } j; \\ 0, \text{otherwise,} \end{cases} \tag{2.1}$$

$$(P^-(j))_k = \begin{cases} 1, \text{if alternative } k \text{ is ranked in a higher position than alternative } j; \\ 0, \text{otherwise,} \end{cases}$$

These are used to calculate the distance function $d(A, B)$ as:

$$d(A,B) = n(n-1) - \sum_{j=1}^{n}[\langle P_A^+(j), P_B^+(j)\rangle + \langle P_A^-(j), P_B^-(j)\rangle] \tag{2.2}$$

2.3 Expert weighting method

The goal of an elicitation is to collect and extract knowledge from an expert. The gathering of knowledge is a simple task, when dealing with only one expert. If more experts are involved, the information provided by the experts needs be weighted carefully before it can be interpreted. Cooke [37, 38, 39] proposed a method to solve this problem. This method aims to quantify the uncertainty in the elicitation and aggregate and weigh the opinions of multiple experts. Some calibration questions on the elicited domain are added to the elicitation. The experts are asked to answer the question by providing a 5, 50 and 95 percentile of the answer. These questions are used to determine the insight on the elicitation subject and certainty of the expert. The result of the question is evaluated by a calibration score and an information score, which together make up the weight of an expert. The calibration score determines how accurate the expert estimated the answer to the calibration question compared to the actual answer. In order for the experts answers in the elicitation to be evaluated, the calibration score should be higher than a predefined threshold α.

For each question an expert has answered it is possible to extract the probabilities of 4 inter-quartiles: $p_1 = 0.05$, for the inter-quartile below 5 percentile $[0, 5]$; $p_2 = 0.45$, for the inter-quartile between the 5 percentile and the 50 percentile $(5, 50]$; $p_3 = 0.45$, for the inter-quartile between the 50 percentile and the 95 percentile $(50, 95]$; and $p_4 = 0.05$, for the inter-quartile above the 95 percentile $(95, 100]$. This means each answer the expert gave, is covered by the inter-quartiles in probability vector $p = \{0.05; 0.45; 0.45; 0.05\}$.

The sample distribution for an expert e is shown in Equation 2.3, where N are the number of calibration questions, the vector $x_1, ..., x_k, ..., x_N$ contains the correct answers to the calibration questions and $q_{k,e,p}$ is the value at percentile p given by expert e for question k.

$$\begin{aligned} s_1(e) &= \frac{\{k|x_k \leq q_{k,e,5}\}}{N} \\ s_2(e) &= \frac{\{k|q_{k,e,5} < x_k \leq q_{k,e,50}\}}{N} \\ s_3(e) &= \frac{\{k|q_{k,e,50} < x_k \leq q_{k,e,95}\}}{N} \\ s_4(e) &= \frac{\{k|q_{k,e,95} < x_k\}}{N} \\ s(e) &= \{s_1, s_2, s_3, s_4\} \end{aligned} \tag{2.3}$$

The sample distribution s and the probability vector p can be combined to calculate the relative

information (or Kullback-Leiber divergence [40]) of the expert as:

$$I(s(e)|p) = \sum_{i=1,...,4} s_i \ln(s_i/p_i) \tag{2.4}$$

The calibration score is shown in Equation 2.5, where N is the number of knowledge questions, r is the relevant quantity value of the vector $x_1, ..., x_k, ..., x_N$, and H_e is the hypothesis that states "*the inter-quartile interval containing the true value for each variable is drawn independently from probability factor p*" [41].

$$\textit{calibration score} = Prob\{2N \cdot I(s(e)|p) \geq r|H_e\} \tag{2.5}$$

The calibration score is essentially a chi-square test with 3 degrees of freedom, also known as a likelihood ratio test. The calibration score is a value between 0.0 and 1.0.

The information score is the concentration of the distribution or the average relative information with respect to the background. To calculate the information score the full distribution of the expert has to be known, including the 0^{th} percentile and the 100^{th} percentile. In order to determine these percentiles, the minimum, $q_L = \min\{r, q_{5,1}, ..., q_{5,e}\}$ and the maximum $q_U = \max\{r, q_{95,1}, ..., q_{95,e}\}$ need to be calculated over all experts e. Next, the q_0 and q_{100} can be calculated using Equation 2.6, where k is an intrinsic range, that allows the analyst to make the experts more informative by stretching the expert's distribution.

$$\begin{aligned} q_0 &= q_5 - (k/100)(q_U - q_L) \\ q_{100} &= q_{95} + (k/100)(q_U - q_L) \end{aligned} \tag{2.6}$$

The information score is shown in Equation 2.7, where $f_{e,i}$ is the expert e probability for question i and g_i is the background probability density over I.

$$\textit{information score} = (1/N) \sum_{i \in N} I(f_{e,i}|g_i) \tag{2.7}$$

The calculation for $I(f_{e,i}|g_i)$ is similar to that of $I(s(e)|p)$, only in this case the expert answers are compared to uniform distributions around the actual calibration answers. The range of the information score depends on the realizations and the values given by the experts, therefore it can exceed 1.0.

Given the calibration score and the information score the weight for expert e can be calculated using Equation 2.8, where $Ind_\alpha(x)$ is 0 if $x < \alpha$ and 1 otherwise. Since the expert weight consists of two independent scores the value needs to be normalized by the number of experts.

$$w_e = Ind_\alpha(\textit{calibration score}) \times \textit{calibration score} \times \textit{information score} \tag{2.8}$$

Given the expert weights w_e and the inputs of the experts $f_{e,i}$ an aggregated decision maker

realization DM_i can be calculated for each question i with:

$$DM_i = \frac{\sum_e w_e \cdot f_{e,i}}{\sum_e w_e} \tag{2.9}$$

2.4 Multi-criteria decision analysis

2.4.1 ELECTRE

In multi-criteria decision analysis different families of methods are available. One such family is the ELECTRE one, which is an acronym for ELimination Et Choix Traduisant la REalité or ELimination and Choice Expressing REality. The main focus of methods of this family is on determining the outranking relation of two alternatives a_x and a_y out of a set of alternatives $(a_1,\ldots,a_i,\ldots,a_m)$ given a set of criteria $(g_1,\ldots,g_j,\ldots,g_n)$. This comparison between a_x and a_y can have multiple different outcomes:

- $a_x S a_y$ means a_x outranks a_y or a_x is at least as good as a_y;
- $a_x P a_y$ means a_x is preferred to a_y or a_x is better than a_y ($a_x S a_y$ and NOT $a_y S a_x$);
- $a_x I a_y$ means a_x is indifferent to a_y;
- $a_x R a_y$ means a_x is incomparable to a_y, there exists a strong opposition on some criteria between a_x and a_y (veto).

In ELECTRE I [42] crisp outranking relations are used, which are determined by a concordance (or agreement) and discordance (or disagreement) conditions. If both these conditions hold, then a_x outranks a_y. The concordance condition states that the sum of weights for all outranking criteria g_j are larger than a predetermined concordance threshold s for a_x to outrank a_y [43], as shown in Equation 2.10.

$$\sum_{j:g_j(a_x)\geq g_j(a_y)} w_j \geq s \tag{2.10}$$

The discordance condition states that for all criteria the disagreement between a_x and a_y should be equal or less than a predetermined veto threshold v for a_x to outrank a_y, as shown in Equation 2.11.

$$\max_{j:g_j(a_x)<g_j(a_y)} (g_j(a_y) - g_j(a_x)) \leq v \tag{2.11}$$

An algorithm is used to determine the necessary pair-wise comparisons between alternatives. From these comparisons the best alternative can be determined.

Since the introduction of the first ELECTRE method in the sixties by Roy [42], many improvements have been made in newer versions of the ELECTRE method. One of the latest methods is ELECTRE-TRI [44]. This method uses outranking for ordinal classification. ELECTRE-TRI assigns a set of alternatives to certain categories $\{C_1,\ldots,C_h,\ldots,C_k\}$ given a set of criteria. The

different categories are separated by different predefined profiles $\{b_1, \ldots, b_h, \ldots, b_{k-1}\}$, which have a value for each criterion g_j. In ELECTRE-TRI each alternative is compared to these profiles rather than to other alternatives, as is shown in Figure 2.1.

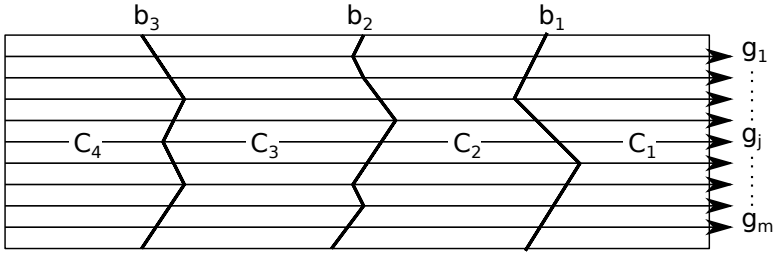

Figure 2.1: An example of the coherence between the classes, criteria and class borders.

ELECTRE-TRI [44] makes use of valued outranking. With valued outranking the concordance and discordance are determined with more complex equations than in ELECTRE I, where simple binary conditions are used to determine the concordance and discordance. The concordance index has two important parameters for each criterion g_j. The first is the indifference threshold q_j defines an area for which an alternative and a profile are considered to be indifferent. The second is the preference threshold p_j defines an area for which the class border is preferred to the alternative. The concordance index $c_j(a, b_h)$ can be expressed using Equation 2.12. The equation provides a linear transition between indifference and preference.

$$c_j(a, b_h) = \begin{cases} 1, \text{ if } g_j(b_h) - g_j(a) \leq q_j(b_h) \\ 0, \text{ if } g_j(b_h) - g_j(a) \geq p_j(b_h) \\ \frac{p_j(b_h) + g_j(a) - g_j(b_h)}{p_j(b_h) - q_j(b_h)}, \text{ otherwise.} \end{cases} \tag{2.12}$$

The discordance index adds veto threshold v_j for each criterion g_j that is a threshold for when a veto should be raised for an alternative on a certain criterion. The discordance index can be calculated using Equation 2.13.

$$d_j(a, b_h) = \begin{cases} 1, \text{ if } g_j(a) > g_j(b_h) + v_j(b_h) \\ 0, \text{ if } g_j(a) \leq g_j(b_h) + p_j(b_h) \\ \frac{g_j(b_h) - g_j(a) - p_j(b_h)}{v_j(b_h) - p_j(b_h)}, \text{ otherwise.} \end{cases} \tag{2.13}$$

The total concordance for all criteria of alternative a can be calculated using Equation 2.14, where k_j is the weight of criterion j.

$$c(a, b_h) = \frac{\sum_{j \in \{1,\ldots,n\}} k_j c_j(a, b_h)}{\sum_{j \in \{1,\ldots,n\}} k_j} \tag{2.14}$$

The relation between alternative a and class border b_h can be calculated using credibility index σ, as shown in Equation 2.15. The credibility index should be larger than a cut-off value λ that

is used to determine the relation between an alternative and a category.

$$\sigma(a, b_h) = c(a, b_h) \prod_{j \in \{1,\ldots,n\}: d_j(a,b_h) > c(a,b_h)} \frac{1 - d_j(a, b_h)}{1 - c(a, b_h)} \tag{2.15}$$

The following situations can exist after calculating Equation 2.15:

- $\sigma(a, b_h) \geq \lambda$ and $\sigma(b_h, a) \geq \lambda$, then a is indifferent to b_h;
- $\sigma(a, b_h) \geq \lambda$ and $\sigma(b_h, a) < \lambda$, then a is preferred to b_h;
- $\sigma(a, b_h) < \lambda$ and $\sigma(b_h, a) \geq \lambda$, then b_h is preferred to a;
- $\sigma(a, b_h) < \lambda$ and $\sigma(b_h, a) < \lambda$, then a is incomparable to b_h.

The assignment of the alternatives to a class is determined using one of the two assignment rules. The alternatives can be ranked using a pessimistic and an optimistic view. The pessimistic view iterates from the last class border b_p to the first class border b_1. The first class border b_h where aSb_h, alternative a is assigned to C_{h+1}. The optimistic iterates from the first class border b_1 to the last class border b_p. Alternative a is assigned to class C_h when b_h is the first class border where $(b_h P a)$. The results for ELECTRE-TRI can easily be calculated by performing the previously described operations to determine the best fit to a certain category for each alternative.

ELECTRE-TRI is also applied in the domain of risk assessment. Merad et al. [45]. applied the method to identify risk-zones with possible collapses in an old mining area. Before it was implemented, experts used old mining maps and risk analysis method to determine risk zones. The authors used this input for the criteria and categories in the model. In a sensitivity analysis, they showed the model returned stable results.

2.4.2 SMAA-TRI

Stochastic Multi-criteria Acceptability Analysis Tree (SMAA-TRI), as proposed by Tervonen et al. [46], is a multi-criteria decision analysis method based on ELECTRE-TRI. It offers a way to perform a parameter sensitivity analysis on the ELECTRE-TRI algorithm. In order to achieve this, three changes are made to the original algorithm. Firstly, uncertain profiles can be represented by stochastic variables in a joint density function. The joint density function is added to make sure all profiles are satisfied and profiles do not overlap. Second, the λ cutting level can be represented by a range between $[0.5, 1.0]$. Lastly, it allows for uncertain weight spaces or even missing weights. The algorithm consists of a Monte-Carlo simulation. In each Monte-Carlo iteration the ELECTRE-TRI algorithm is executed with different samples of weights, cutting levels and profiles. After completing all iterations, the algorithm calculates a category acceptability index for each alternative based on the outcome of Monte Carlo simulation. This shows the probabilities of an alternative belonging to a certain class, which provides more information and is easier to interpret than a crisp outcome with a sensitivity analysis. The SMAA-TRI [47]

was also applied to an environmental risk assessment of nano-materials. This case study shows the extra value of category acceptability indices, when dealing with imprecise inputs. The results showed the input criteria, obtained from experts, were not precise enough the give a stable recommendation.

Chapter 3

Methodology

Section 1.2 presented the main research question and five sub-questions for evaluating security of NFC payment applications for smartphones. In order to answer these questions a security evaluation needs to be conducted. Normally a security evaluation is performed by two or three persons. This might cause bias advise on the security risks, since only limited perspective is taken into account. For this reason, a model is proposed in this chapter that aggregates the view of multiple security experts to get a more unbiased and realistic evaluation. A diagram of the proposed model is shown in Figure 3.1. The proposed model consists of multiple methods that are either proposed in this chapter or extracted from the literature and explained in Chapter 2. The proposed model should be executed by a security analyst on behalf of a client. In this chapter each of the five steps in the proposed model is described in a separate section. In the Section 3.1, the different inputs for the proposed model will be introduced. Next, the gathering of information from the security experts is explained. The third part describes the transformation of the expert views, so they can be handled by the multi-criteria decision analysis tool, which is explained in Section 3.4. The last section of this chapter explains how the results of the multi-criteria decision analysis tool can be used to make a recommendation on the security of a system.

3.1 Step 1: Input

For the model proposed in this thesis four inputs are required: a set of attack scenarios, a set of fraud scenarios, a security evaluation method, a set of security experts and a set of calibration questions.

3.1.1 Attack and fraud scenarios

The proposed model requires a set of attack scenarios for the system, technology or asset under evaluation that represent the security landscape of the subject under evaluation. These attack

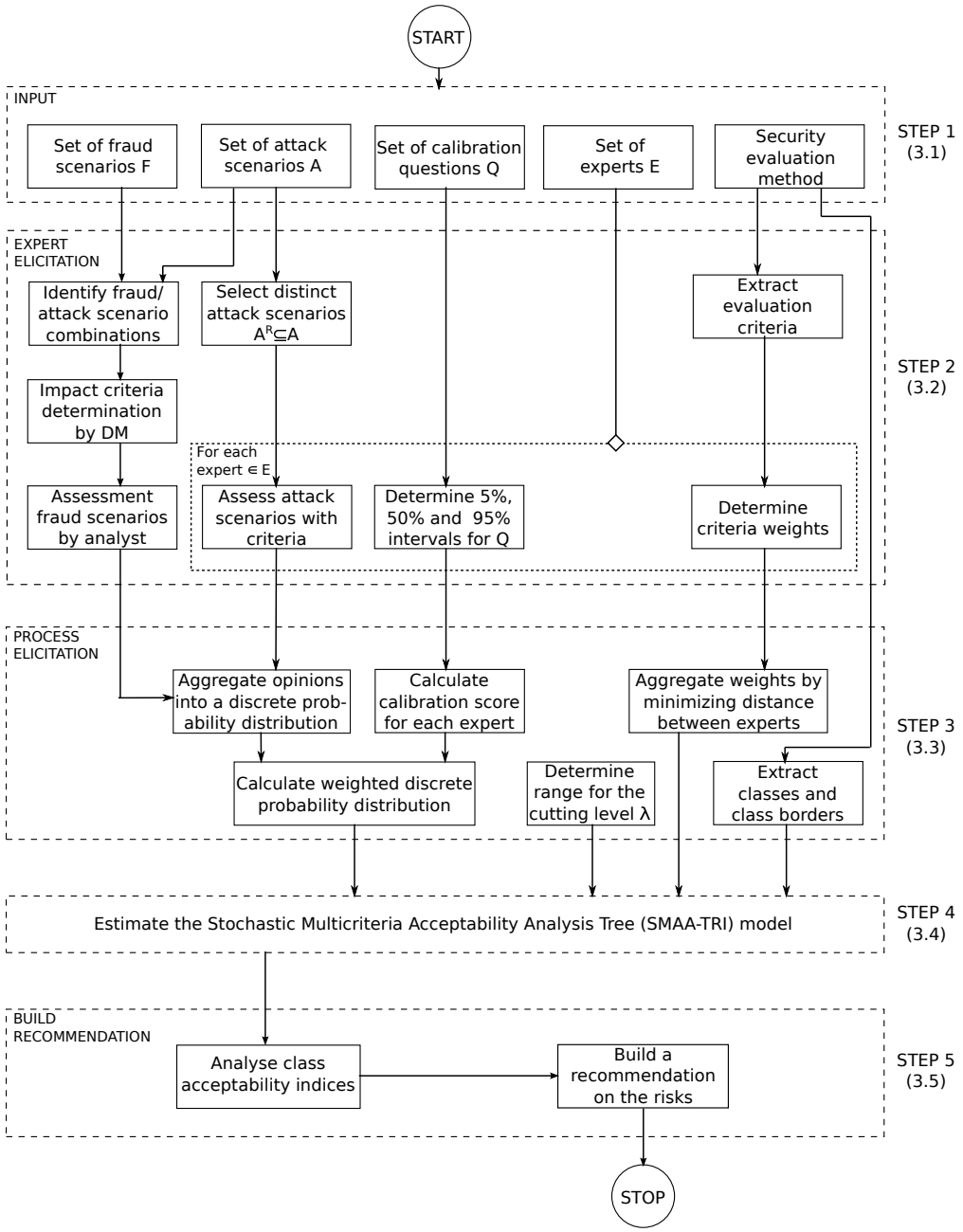

Figure 3.1: A detailed overview of the proposed process for robust risk assessment.

scenarios describe which steps need to be taken in order to exploit a certain vulnerability in the evaluated subject. This is required to determine the likelihood of a security risk. In the proposed model, the determination of the likelihood will be performed by multiple experts. The attack scenarios need to be defined for the evaluated subject. When defining the attack scenarios, there

are two requirements. Firstly, the attack must be accurately described. Leaving out too much details causes a major dissent among the elicited experts, because every expert is interpreting the attack scenario differently. Secondly, the attack scenarios need to have a theoretical or practical foundation. The feasibility of the attack needs to be supported by an experiment or by the literature [27].

The proposed model also requires a set of fraud scenarios that is used to determine the impact. These scenarios describe how one or more attack scenarios can be used to the benefit of the attacker and might harm the owner or user of system, technology or asset under evaluation. Likewise, every attack scenario leads to one or more fraud scenarios. A security expert will perform the impact assessment using criteria provided by client taking the risk. In order for the security expert to make a fair assessment, the fraud scenarios also need to be accurately defined. The second requirement of the attack scenarios does not hold for the fraud scenarios. Fraud scenarios cannot simply be extracted from an experiment as with the attack scenarios. However, they can be extracted from newspapers or other news media, when they are already exploited. The scenarios can also be derived from reasoning on the attack scenarios applying experience in the related fields.

3.1.2 Security evaluation method

In order to offer a meaningful recommendation as a result of the evaluation process, it makes sense to use an already existing risk assessment method as a basis for the proposed model. These methods are developed over multiple years by experts in the security field and are already proven concepts for evaluating security risks. Furthermore, the use of these methods is often necessary in order to be compliant with certain security standards required in some industries. Nevertheless, there are also some requirements for the risk assessment method to be used in the model proposed in this thesis. Firstly, it needs to evaluate attack scenarios with a set of predefined criteria that are the same for each scenario, as otherwise comparing the scenarios with each other becomes impossible. Secondly, the proposed model requires the method to have a quantitative foundation, since the actual values will be used in the computational risk estimation. The ETSI TISPAN TVRA [27] used in the case study satisfies these requirements. The structure of this security evaluation method is already explained in Section 2.1. The implementation of other quantitative risk assessment methods in the proposed model is out of the scope of this thesis.

3.1.3 Experts

The experts are people who have experience with the system, the technology or the asset under evaluation and people with experience in information security. These experts will be asked to evaluate the security of system, technology or asset under evaluation using the attack scenarios and security evaluation. The number of experts depends on the availability of experts and the way the elicitation is performed. It might happen, in case of a new technologies, the number of

available experts with knowledge of the subject is limited. Therefore, it might not be possible to get enough experts to take part in the elicitation. With a small number of experts, it is better to perform a regular risk analysis, since the proposed model does not add much value with just a few experts. Furthermore, when an elicitation is performed offline rather than online, a smaller number of experts is required. The reason for this is, that the response rate in an offline elicitation is higher, and one can expect the opinions to be closer together, when experts are able to discuss the scenarios with each other.

3.1.4 Calibration questions

The calibration questions are used to quantify the uncertainty in the experts by weighting them for the insights on the subject and certainty of the answers. In these questions the experts are asked to estimate the answer for a certain question. Typically, the experts are asked to give an 5, 50, and 95 percentiles to a question with a numerical answer. These questions are in the domain of the evaluated subject and security, so experts from both domains are able to give a good estimation of the questions asked.

3.2 Step 2: Expert elicitation

3.2.1 Attack scenario elicitation

The data on attack scenarios for the proposed model, as shown in Figure 3.1, is gathered from security experts using a elicitation. This elicitation step can be divided in two parts: the creation of the elicitation and eliciting experts. Some of described inputs in Section 3.1 require pre-processing before they can be used in the elicitation. An expert elicitation cannot take too long since the expert's time is valuable. Therefore, only distinct attack scenarios from the set attack scenarios are used as input for the elicitation. This guarantees most of the security landscape is covered. One should note that in some cases, one detail in the scenario can have a significant change on the likelihood of an attack scenario. Accordingly, one should find a careful trade-off between the elicitation time and coverage of the security domain.

In order to implement the security evaluation method in elicitation, the likelihood evaluation criteria need to be extracted from security evaluation method. This involves the criteria and the possible values for each of the criteria. For example, in case of the risk assessment method used in the case study, the criteria are elapsed time, specialist expertise, knowledge of the system, window of opportunity and equipment availability.

The elicitation consists of three parts. The elicitation starts by presenting the extracted criteria to the experts, since not all experts are familiar with the used security evaluation standard. Then, the experts are asked to give an ordinal ranking for how much each criterion in general contributes to an attack. Next, the selected set of attack scenarios is presented to the experts. In this part, the experts are asked to evaluate each attack scenario with extracted criteria (ques-

tions) and criteria values (answers). In the last part of the elicitation, the experts are asked to answer the set of calibration questions. Note that it is best to ask the experts for a lower bound, median and upper bound for each of the questions, since this is more clear for most experts than asking for the percentiles [37].

3.2.2 Fraud scenario elicitation

It is impossible for the experts to make an estimation of the impact, so the elicited experts are only asked to give their opinion on the likelihood of the attack scenarios and not on the impact of the fraud scenarios. The impact of a risk strongly depends on the client taking the risk. For example, a possible loss of 10,000 euro has a far larger impact on a small family business than on a multinational. The experts do not have information on the client implementing the NFC payment application. Even when a profile of a (fictional) company is provided in the elicitation, each expert will interpret the profile in a different way. This does not correspond to a real-world application of information security risk assessment, where the impact criteria is often assessed in collaboration with the client. As mentioned in Section 3.1.1, the fraud scenarios are therefore assessed by the security analyst who determines the impact classes together with the client. The way the impact classes are determined is left open to the choice of the client and the security analyst, as long the result corresponds with the used risk assessment method. However, it is recommended a quantified impact classes are used, because the impact has a considerable effect on the outcome of the proposed model. Before the security analyst starts the assessment of the fraud scenarios, the possible combinations between fraud and attack scenarios need to be mapped. One should note that, it is not strictly necessary to link attack scenarios to fraud scenarios, but it supports the security analyst in the assessment.

3.3 Step 3: Process elicitation

In this step the results of the elicitation are processed, so it can be handled by the multi-criteria decision analysis method. Firstly, the method of Cooke, as explained in Section 2.3, is used to calibrate the experts. This method proved itself in the calibration of experts in expert elicitations over the past twenty years in various domains. Over 25 instances can be found in the literature that use this method for risk analysis, ranging from the risks in the application of vaccines to security risks [38, 39, 48]. The answers of the experts on the calibration questions will be used as input for the method. In most cases, the method of Cooke is used to determine unknown quantitative values using an expert elicitation, where the expert weights are used to determine these values. However, we only use the experts weights that are produced in this process and use those to weigh the assessments of experts. Secondly, the method of Cook [35], as explained in Section 2.2, is used to determine the criteria ranking. This method can extract a consensus ranking from a set of ordinal rankings. This offers an advantage towards consensus methods that use pairwise comparison or weights, since these require either more questions in the elicitation

or require more effort from the expert to fill in. The criteria rankings of the different experts are used as input for the proposed model. This input is then used to determine the criteria ranking that is closest to the criteria ranking of experts, resulting in a distance-based consensus ranking. Lastly, the views of the experts are aggregated, as described in Section 3.3.1.

3.3.1 Expert view aggregation

For each criteria of each attack scenario in the elicitation, the experts gave their opinion. In most cases, the experts did not all chose the same criterion values. In order to respect the inputs from all the experts, the opinions need to be aggregated in a probability distribution. Since the criterion values are discrete from nature, a discrete probability distribution is used. The discrete probabilities can be extracted by taking the number of answers for the criterion values for a certain criteria. The probability is derived by the number of experts from the set of experts E that chose a particular criterion value ($\sum_{e \in E} c_e$) divided by the total number of experts that gave their opinion for that criterion of an attack scenario ($|E|$), as shown in Equation 3.1. Note that c_e represents a binary value that is 1 if the expert has chosen criterion value x and 0 otherwise. The attached values x in the discrete probability distribution are the quantitative values that are attached to the criterion values as defined in the risk evaluation method.

$$P(x) = \frac{\sum_{e \in E} c_e}{|E|} \tag{3.1}$$

This last equation suggests, the experts are assigned equal weight. However, having applied the expert calibration, this is not the case. The expert opinions have to be compensated with the expert weight to make a fair assessment. This is shown in Equation 3.2, where w_e is the weight for a certain expert.

$$P(x) = \frac{\sum_{e \in E} w_e * c_e}{\sum_{e \in E} w_e} \tag{3.2}$$

This probability distribution is determined for each criterion of each attack scenario.

Before the probability distributions can be utilized in the proposed model, the impact values need to be incorporated. Doing this will create risk scenarios for every possible combination between the attack and the fraud scenarios. If an attack scenario can lead to two fraud scenarios, these will become two separate risk scenarios. However, for the final classifications only the scenario with the highest risk is taken into account. The impact values are combined with the attack scenarios by changing the values for x. This magnitude of this change depends on the risk calculation between impact and likelihood defined in the security evaluation method. For example, if the security evaluation method defines the values for an attack scenario need to be divided by 2 to get a risk scenario with low impact, all the x values in the corresponding probability distribution for that attack scenario are divided by two.

3.4 Step 4: Multi-criteria decision analysis

The proposed model uses SMAA-TRI, described in Section 2.4.2, as multi-criteria decision analysis method, since it offers ordinal classification, can handle imprecise values and makes use of valued outranking. These three properties make it a suitable multi-criteria decision analysis method for risk analysis. The ordinal classification allows for scenarios to be divided over risk categories. The imprecise values, such as discrete probability distributions and ranges, can compensate for the uncertainty that is associated with risk assessments. The value outranking makes sure a quantitative risk assessment can be implemented into the method. In the model proposed in this thesis, the discrete probability distributions and the criteria ranking from Section 3.3 are used as input for the SMAA-TRI method. Furthermore, the risk scenarios will be represented by the alternatives, then evaluation criteria by the criteria and the class borders will be extracted from the risk assessment method in the SMAA-TRI method.

3.5 Step 5: Build recommendation

After running the SMAA-TRI method, it will return the results of the risk assessment. The SMAA-TRI method, as described in Section 2.4.2, returns category acceptability indices rather than crisp classifications. In order to give a good recommendation to the client based on these indices, one needs first to interpret the results.

Multiple recommendations one can be given based on the partition in the category acceptability index for a certain attack scenario.

Distinct classification In case of a risk analysis, if one class has an acceptability of 0.8 or higher, it is safe to accept this class as the final risk classification, since insufficient evidence is available that proves the risk belongs to another risk class. The value of 0.8 is chosen, because for this value almost every expert opinion points towards a single class. One should also note that the probability of unanimity is quite small when dealing with independent samples. This specific class can be reported back to the client as the result of the risk assessment of the attack scenario.

Minor dissent among experts In case the majority of the class acceptability is divided among two adjacent classes and one class has an acceptability of 0.6 or higher, one can speak of a minor dissent among experts. The value of 0.6 is chosen, because at this point there still is a clear majority of the experts that chose for a certain class. One should report back to the client that the assessment is between two classes with a preference towards the higher class for this attack scenario, because of the precautionary principle in security risk assessments [49, 50].

More research required In case the majority of the class acceptability is divided among three adjacent classes with an ascending or descending class acceptability, it safe to say more

research into the likelihood is required in order to give a good classification. In this case there is also no clear majority in one of the classes. As long as the results are inconclusive, the best strategy is to report the high risk class with a significant acceptability to the client as classification.

Confusing scenario When the class acceptability index does not correspond to any of the previous recommendations and the class acceptability is not concentrated around a specific class, it is safe to consider the attack scenario was not clear. It could be possible the implementation of the attack scenario strongly depended on not specified details or specific knowledge required to assess the attack scenario was not available for all the experts. Hence, it is not possible to report an classification to the client. The best strategy is to modify the scenario and redo the assessment for this scenario.

Note that only in case of a correct classification or a minor dissent among experts, a final classification can be reported to the client.

3.6 Applications of the model

In the case study of the model, the model is applied to security risk assessment. However the model proposed in this chapter is not limited to this kind of risk assessments and could also be applied to different kinds of risk assessments, as long they have quantitative foundation and are able to classify the risks in different risk classes. The SMAA-TRI as well as the ELECTRE-TRI are both applied to different kind of risk assessments. In Tervonen et al. [47], the authors mention the values for the risk assessment for materials presented in the paper are determined using only the authors' expert judgment. By eliciting different experts in the nano-material domain using the proposed model, one might get an even more robust risk assessment than the one presented by Tervonen et al. One should note that the criteria are extracted from the literature rather than a predefined risk evaluation method and that the fraud scenarios are not used for this case. The mining hazards problem of Merad et al. [45] could also benefit from the model. The authors already make use of expert committee to determine the data and to validate the results, but only use individual expert opinions for ranking the criteria and do not weigh the different expert opinions.

Besides the type of risk assessment and the risk evaluation standard, the ordinal distance-based method for the criteria can also be replaced with weights or weight ranges. However, one should use a method for aggregating the expert ranking of the criteria that provides an empirically proved consensus to ensure a robust risk assessment. Furthermore, the used values for the different recommendations in step 5 of the proposed model can also be changed. Although, the current values for the recommendation are selected to match the precautionary principle in risk assessments, one can change these values to match the perceived risk of the client.

Chapter 4

Case study: NFC Payments

In this chapter, a case study for NFC payment applications for smartphones is implemented into the proposed model described in Section 3 . This chapter starts with an overview of the experiments performed to gain insight in the NFC security landscape on a popular mobile operating system, and the different attack and fraud scenarios in Section 4.1. In Section 4.2, the setup of the expert elicitation for this case is presented. Section 4.3 discusses the implementation of the methods used to process the results of the elicitation. This chapter ends by explaining how the SMAA-TRI method is implemented for the NFC payments case.

4.1 Attack and fraud scenarios

4.1.1 Attack experiments

In order to gain insight in the different attack scenarios and the likelihood of possible attacks, there were also some attack experiments performed for this thesis. These experiments focused on the relay attack, the secure element and placing malicious tags near terminals. The experiments were performed using two Samsung Galaxy Nexus smartphones, a SCM SDI010 RFID reader, HID Omnikey 5321 RFID reader and a Touchatag NFC reader.

Relay attack

In order for a relay attack to succeed, at least one of the (attack) devices involved should support card emulation. By default, card emulation mode is not enabled in Android. One of the advantages of Android being open source is the availability of after-market versions of the mobile operating system, the so-called custom ROMs. In one of the popular custom ROMs, CyanogenMod, a patch was released that enables card emulation for Android. This patch was written by Yeager [51] in order to support his company's NFC wallet application, SimplyTapp. This patch enabled the possibility of a relay attack. The original patch was committed to the

CyanogenMod source in January 2012. The implementation was changed in March of the same year, because of conflicts with Google's payment application Google Wallet.

The experiments on the relay attack were based on the work of Lee [22]. As discussed in Chapter 1, Lee developed an Android application to demonstrate a relay attack on the Android operating system. Lee's app needed to be installed on two Android smartphones. It relayed the data from an RFID card via the first Android smartphone over WiFi to the second Android, which behaved like the RFID card when sending the data to a payment terminal. The drawback of Lee's app is, the attack's part of the app only works with the original card emulation patch and stopped working after the modified implementation from March.

Some research into the source code of CyanogenMod suggested it was still possible to use the card emulation mode. In order to achieve this, one had to use the Foreground Dispatch function in Android's app development framework. With some modifications to the publicly available source code of Lee's app, it was possible to perform a relay attack with all versions of CyanogenMod 9 and 10 released between January 2012 and the moment of writing (December 2012). As of February 2013, these modifications are part of the original release made available by Lee.

The implementation of the card emulation patch in CyanogenMod is however somewhat limited. In card emulation mode, the smartphone notifies the reader it is an emulated MiFare Classic card. Most readers will only send data back, if the presented card is supported, so with most public transport and access control cards this will not work. However, most payment terminals from MasterCard and Visa do support this card type, because Google Wallet sends out the same notification when it starts communicating with the terminal. One may conclude that only certain NFC applications on Android are vulnerable for a relay attack.

Secure element

The secure element interface in the Android mobile operating system is provided by the NFC chip of semiconductor manufacturer NXP. The chip provides an embedded secure element and an interface for a secure element on the UICC (SIM card), the latter of which is disabled by the OS. The secure element was introduced in Android 2.3.4 together with Google Wallet and was relatively easy to gain access to. In version 4.0.4 of the mobile operating system the access control was changed. From this version on the secure element only supported apps that are signed with a certificate and are white-listed on the system-level. On a regular Android smartphone the only way to get an application white-listed is by cooperating with the vendor of the phone. However, on a rooted phone it is possible to gain access to the white list and even edit it. This enables custom apps to gain access to the secure element.

For this thesis, the app of Elenkov [52, 53] was used to explore the secure element of one of the Galaxy Nexus phones. The secure element contains multiple so-called applets. These applets are maintained by Global Platform and use the Global Platform Card Specification for all internal commands. Most of the commands for making changes to an applet (adding and deleting cards)

require authentication with a Card Manager. The Card Manager manages the different applets on the secure element and allows for a secure channel between the applets and the app on the phone. The authentication requires for a pair of triple Data Encryption Standard (3DES) keys, which also provides encryption and data integrity. It is not possible to brute force the keys, since the Card Manager terminates itself after ten failed authentication attempts. One of the applets in the secure element is the MiFare4Mobile applet, which provides emulation of Mifare Classic card that can be used in customer loyalty programs. This applet is free, but comes with a non-disclosure agreement from NXP.

One of the Galaxy Nexus smartphones had version 1.5-R79-v5 of Google Wallet installed. Google Wallet also stores part of its information on the secure element. Using the application of Elenkov, it is possible to view the different cards stored in the secure element. This is mainly limited to prepaid cards, since actual credit cards are stored on the servers at Google rather than on the phone. The app also allowed to view historic transaction data stored in the SE. Until the latest version of Google Wallet, it was even possible to act like the Wallet app and relay data to another Android smartphone, as shown by [5].

Malicious Tags

The previous two experiments mainly focused on gaining access to NFC and the secure element. However, it is not unthinkable an attacker wants to gain access to the phone over NFC. As described in Section 1.6, the use of malicious tags is a serious threat to NFC. When an attacker places such a tag on top of terminal, he could expect a significant amount of victims, since all people using the terminal will have NFC enabled.

A sticker was fitted with a malicious URL that would reboot any phone with Android 4.0.4 or lower. This attack was based on the Samsung USSD hack as presented by Borgaonkar [26] as described in Section 1.6.2. The sticker was placed onto the reader, after which the reader was enabled.

The experiment was performed on three different readers: a SCM SDI010 RFID reader, a HID Omnikey 5321 and a Touchatag NFC reader. The first two readers were too powerful. The phone picked up on the signal of the reader before it got a chance to connect with the sticker. As long as the reader was active, the phone would completely ignore the sticker on the reader. The last reader stopped working, when both the phone and the sticker were in its field simultaneously.

4.1.2 Attack scenarios

Ten different attack scenarios are defined in this section, each of which consists of a vulnerability and an attack path. Although it is possible to define more attack scenarios by combining different vulnerabilities and attack, it was decided not to do this. This is because, the focus of this section is on not on vulnerabilities, but on attacks. Additionally, the number of attack scenarios needs also be limited for time restrictions on the elicitation. Every attack scenario is based on known

attacks as described in Section 1.6, the performed experiments, or common sense. Furthermore, from every attack one or more fraud scenarios can arise. The possible fraud scenarios for each attack scenario are defined in Table 4.1, which is part of Section 4.1.3.

1. Relay attack

Vulnerability In some smartphones, NFC is always on, even if the smartphone is not in use. This makes it possible to perform a transaction with the payment application on the phone, which makes the smartphone vulnerable for a so-called relay attack.

Attack scenario In a relay attack, there are two attackers. One attacker has a relay device and the other attacker has a proxy device. These devices can be a smartphone or another NFC-enabled device and are connected with each other over Internet. The relay device is held close to the victim's smartphone in a crowded place, like in public transport during rush hour. The proxy device is used to perform an NFC payment at a payment terminal. The communication between the victim's phone and the payment terminal is relayed over the proxy and relay. The relay attack is described by Francis et al. [4] and shown in Figure 4.1.

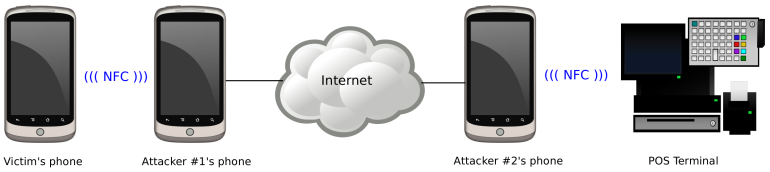

Figure 4.1: A relay attack

2. Relay attack using malicious app

Vulnerability On most modern smartphones, it is possible to gain superuser privileges (also known as rooting or jail-breaking) with a simple procedure. Many smartphone users do this to gain full control over their smartphone and use it, for example, to perform extended backups or run customized versions of the mobile operating system. Unfortunately, it also circumvents some of the security features of the phone, like the sandbox for applications. The secure element, in which NFC payment application resides, is also protected by these security features. Subsequently, the secure element is more vulnerable on a rooted smartphone.

Attack scenario An attacker tricks the victim in installing a malicious app by offering an interesting feature or hack. The victim thinks he grants the app access rights for the feature. The app uses the access rights to perform the feature, but meanwhile grants itself access to the secure element of the smartphone. The app notifies the attacker over the Internet, it gained

access to the secure element. The attacker can now perform a payment using the payment details of the victim, which are relayed from secure element on the victim's phone to the NFC-enabled smartphone of the attacker.

3. Eavesdropping using malicious app

Vulnerability In some implementations of NFC payment applications the account data is not saved on the secure element, but on the servers of the payment platform provider. In this case the user receives a call back of the transaction from the payment platform provider over an Internet connection, rather than over NFC. If rooted phone is used in such a transaction, the communication between the server and the phone is vulnerable for eavesdropping.

Attack scenario Just as in the previous attack, the victim is tricked into downloading a malicious application by offering an interesting feature or hack. The victim grants the application access to its phone and the app performs the feature. The victim invokes an transaction with a payment terminal and receives a valid call back from the payment platform provider. Unlike in the previous scenario, the application is now used to gather account and transaction data during transactions. This data is sent to the server of the attacker.

4. Eavesdropping

Vulnerability Most current NFC payment applications for smartphones make use of the Europay, MasterCard and Visa (EMV) Contactless Specifications for Payment Systems standard. This standard is a framework for all types of contactless payments, like NFC payment apps and RFID-enabled credit cards. In this protocol, the communication between the payment terminal and the NFC-enabled smartphones is not encrypted. This makes it vulnerable for eavesdropping.

Attack scenario The attacker can buy or make equipment that listens on the 13.56 MHz frequency of near-field communication. The attacker places the equipment near a payment terminal. If a victim uses the payment terminal for a payment, the attack can listen in to the transaction.

5. Eavesdropping by exploiting the terminal

Vulnerability Terminals are essentially computers that allow for communication outside the system environment. In the case of payment terminals, these communication channels consist of wide area network (e.g. Internet) connection with the payment service provider and a possible connection with a NFC-enabled smartphone using near-field communication. This makes the terminal susceptible for malicious inputs.

Attack scenario An attacker creates a malicious NFC device containing an exploit for the payment terminal. The attacker injects the malicious code into the terminal during a transaction. The exploit gathers information during real transactions with NFC devices. The attacker returns to the payment terminal and gathers information using his NFC device. In case the terminal is connected to the Internet, the attacker could also retrieve information over the Internet.

6. Modify transactions by exploiting the terminal

Vulnerability As described in the previous vulnerability, terminals have multiple communication channels, which can be utilized as gateway to the payment terminal. This makes the terminal not only susceptible for eavesdropping, but also for manipulation using malicious input.

Attack scenario An attacker creates a malicious NFC device containing an exploit for the payment terminal. The attacker injects the malicious code into the terminal during a transaction. The exploit allows the attacker not only gain insight into transaction date, but also allows the attacker access to the processing of the transaction. This means the attacker is able to manipulate transactions, by modifying, creating or blocking them. The attacker might be able to control this over the Internet.

7. Malicious terminal

Vulnerability Not all terminals can always be trusted. An attacker can act like a merchant and be in possession of a NFC payment terminal. The attacker is free to do as he wishes with the terminal.

Attack scenario The attacker obtains an NFC payment terminal. Thereafter, the attacker modifies the payment terminal to obtain transaction information and manipulating transactions. By having physical access to the legitimate terminal, it is possible for the attacker to swap the legitimate terminal for a fraudulent terminal. The payment terminal is then used in a legitimate transaction.

8. Denial of service using a zero-day vulnerability

Vulnerability It is possible for a zero-day vulnerability to exists in the mobile operating system running on the NFC-enabled smartphone. A zero-day vulnerability can cause a privilege escalation in the mobile OS. Since the secure element in the smartphone might rely on the security features of the mobile OS as well, it could be vulnerable as well.

Attack scenario The attacker finds a zero-day vulnerability in the mobile operating system and creates an exploit to use the vulnerability to its advantage. The attacker uses a malicious tag near a payment terminal to send the victim to a website that executes the exploit on the smartphone of the victim. With the exploitation of the privilege escalation, the attacker gains

access to the secure element on the smartphone. It then tries authenticate itself with the secure element, using bogus authentication details. After multiple failed tries, the secure element will permanently lockout all users.

9. Denial of service of the terminal

Vulnerability NFC payments are also interesting for vending machines, since this involves small payments. In case of an integrated payment terminal in a vending machine, the terminal is left unattended.

Attack scenario An attacker places a jammer near a payment terminal. This jammer sends out a powerful signal on the NFC frequency band. When a victim tries to make a payment, the jammer corrupts or even blocks the data sent between the terminal and the victims smartphone.

10. Theft

Vulnerability The smartphone used in transactions is in possession of the user. In most smartphones it is left up to the user to enable security settings in the smartphone. If no security is set, the phone can be used by anyone to perform transactions.

Attack scenario The attacker steals the smartphone of the victim. The victim not immediately notices the smartphone is missing. The attacker uses the victims phone to perform NFC payment transactions.

4.1.3 Fraud scenarios and impact classification

The attack scenarios described in the previous section can be used for different types of fraud scenarios, as shown in Table 4.1. While attack scenarios mainly focus on the likelihood of an attack, fraud scenarios are concerned with the impact of attacks.

These fraud scenarios are defined with a payment provider as main stakeholder. A payment platform provider is an entity that offers the NFC payment platform to the merchant and the customer. It offers the customer a NFC payment app and the merchant a NFC payment terminal. The fraud scenarios in this section mainly focus on the direct consequences of the execution of a fraud scenario. One should note that indirect consequences, like reputation damage, might also harm payment platform provider. We identified four different fraud scenarios for the payment platform provider.

In our case study, the impact for the fraud scenarios is not determined in collaboration with a client. We use an impact indication for the implementation of NFC payments of a specific payment provider that serves 10 million users. This payment provider will remain anonymous due to confidentiality restrictions. The payment provider gave an indication for the impact of financial damage and the impact of reputation damage, as can be seen in Table 4.2 For the

Table 4.1: Attack/fraud scenario matrix

	Privacy infringement	Single malicious tran.	Multiple malicious tran.	Disable service
Relay attack	X	X		
Relay attack with malicious app.			X	
Eavesdropping with malicious app.	X			
Eavesdropping	X			
Eavesdropping with terminal exploit	X			
Modify transactions with terminal exploit		X		
Malicious terminal			X	
Denial of Service with zero-day exploit				X
Denial of Service on terminal				X
Theft	X		X	

financial impact, all fraud scenarios with financial damage below 1 million euro are considered to have a low impact. The incidents that cause a damage between 1 million and 10 million are deemed to have a medium impact. In case the financial damage exceeds 10 million euro, the fraud scenario has a high impact. The reputation damage is subjected to a more qualitative scale. The fraud scenario is considered to have low impact, if the incidents are only published in web articles and the payment provider can prevent further escalation. A medium impact for reputation damage is assigned, if the incidents are published in newspapers and it leads to a decrease in users and merchants. In case the fraud scenario leads to a significant loss of clients and the payment provider is forced to perform a large migration or discontinue the service, the reputation damage is deemed high. If the reputation damage and financial impact return a different value, the highest value is used to rate the impact of the fraud scenario. This definition of the impact is applied to fraud scenarios in order to determine the impact for this one payment provider.

Privacy infringement

When an attacker has access to the transaction and payment information, there exists the possibility of privacy infringement. In this case the attacker has access to information the victim wants to keep to itself. It is for instance possible for the attacker to see what the victim bought, where he/she has been and the amount of money he/she has. Moreover, access to transaction and payment information also violates the privacy laws in most countries. In the Netherlands, access to someone's payment and transaction details is in violation with article 10 of the constitution [54].

Table 4.2: Impact classification

Impact	**Financial damage**	**Reputation damage**
Low	$<$ 1 million	Incidents are only published in web articles and the payment provider is able to prevent further escalation.
Medium	1 - 10 million	Incidents are published in newspapers and it has negative effect on the number of users.
High	$>$ 10 million	The aforementioned matters escalate to such a level that the payment provider is forced into a large software migration or discontinuation of the NFC payment service.

The financial impact for privacy infringement consists of a possible fine and modification of the software. In the Netherlands, a maximum fine of 4500 euro is determined for intentional privacy infringement by an organization [55]. In order to prevent future privacy infringement, it might be necessary to make changes to the NFC payment application. Since only changes to the mobile application are required and not to the back-end of the payment platform, the costs for theses modifications are relatively low. This modification of the software and the fine will not exceed 1 million euro, so the financial impact is deemed low. The reputation damage is a different story. Nowadays, privacy infringement incidents receive a lot of attention in the regular media, resulting in reputation damage for the company in question. Furthermore, such an incident might also cause users to switch to another payment method, since enough alternatives are available. These two facts cause a medium impact of reputation damage and therefore this fraud scenarios is classified to have a medium impact.

Single malicious transaction

In case of a relay attack, the attacker is granted limited access to the payment details of the victim. Although, the attacker might not able to read the payment details, he is able to perform a payment with the payment details of the victim. So, it is possible for the attacker to rob the victim and obtain goods. Added to this, it is not common to ask someone for identification before a transaction in shops, so the attacker has essentially free rein. Another variation of this fraud scenario is a unwanted transaction on a malicious terminal. This already happens with mobile payment terminals for debit card payments in the Netherlands. For example, the victim is thought to believe his first transaction failed and he needs to make another one, which is sent to the account of the attacker. Another possibility is that the terminal shows a lower amount than is actually processed, causing the victim to transfer large amounts to the attacker.

The financial impact for the single malicious transaction is low, because the payment provider

uses a restriction on the maximum transaction value, 25 euro, on every transaction. This restriction makes it also less interesting for criminals to exploit an attack path that leads to a single malicious transaction. Therefore, the total damage will not exceed 1 million euro. The reputation damage is also low, since related fraud scenarios, such as skimming or e-banking, also occur and payment providers accept these risks, not holding consumer liable. The overall impact for this scenario is therefore considered low.

Multiple malicious transactions

In case of a malicious applications or zero-day exploit, it is possible for an attacker to gain access to payment details for a longer period of time. This means the attacker can make multiple transactions with the same payment details. Since the combined value of multiple transaction is likely to be higher than those of a single transaction, this has possibly a larger impact on the victim.

The case for multiple malicious transactions is a different story. As described in Section 4.1.3, this scenario involves malware that transfers payment details from the victim to the attacker. A professional criminal organization may eventually cause 1% of the 10 million users to be infected with their malware. The maximum transaction value is set by the payment provider at three transactions of 25 euro each. It is expected that the attacker needs to make one big strike, as some users will notice within a day that their transaction log is incorrect and might contact their payment provider. Next, the payment provider will block payments towards the attackers. An attack can gather an estimated maximum 7.5 million (= 1% of million users infected $\times$ 3 transactions $\times$ 25 euro per transaction) exploiting this fraud scenario, so the direct financial damage has a medium financial impact. The costs of software modification are considered low, similarly to the privacy infringement scenario. The reputation damage is deemed a medium impact, since this attack harms many users in a short period of time, it will be reported in newspapers. Furthermore, it might cause users to use alternative payment methods, i.e. a loss of users. All in all, this fraud scenario has a medium impact.

Disable service

The payment service can be disabled at the merchant-end and on the customers-end. On the customers-end, it is inconvenient for the victim, if the NFC payment application cannot be used, since another payment method needs to be used. If in addition, the service is blocked by locking the user out of the SE, the secure element of the phone might be rendered unusable. On the merchant-end (e.g. jamming), disabling the NFC payment service might even cause customers to choose for another merchant with working NFC payment system. This can have a negative impact on the merchants revenue.

The disable service fraud scenario consists of a case for the merchant and the user. The merchant case is limited to a single merchant, therefore the reputation damage and the financial impact for the payment provider are low. For the user case, it is also anticipated that 1% of

the 10 million users would have a malware-infected smartphone. Based on the available data for a similar payment method, the Chipknip [56], it is estimated every user pays on average 0.60 euro a week with the NFC payment application. It is expected it takes two weeks to patch the vulnerability the malware is exploiting and to release an update through the app store, resulting in 120 000 euro (= 1% of 10 million users × 0.60 euro of lost transactions per user a week × 2 weeks) in missed transactions, so the financial impact is low. The experience for the user is anticipated to be same as any other disruption in existing payment systems; it will cause some annoyance, but it will probably not lead to a permanent loss of clients. Therefore the reputation damage is deemed low. Given the low impact for both the user and the merchant case, this fraud scenarios has a low impact.

4.2 Expert elicitation setup

An important part of a robust risk assessment is the fact the opinion of multiple experts in a particular field is considered, as explained in Section 3.2. Considering multiple opinions will reduce the uncertainty involved in a risk assessment and takes different views on the risks into account. This will give a more robust risk assessment than the one performed in Appendix C. In order to obtain the opinions for the risk assessment an expert elicitation was performed.

A group of 39 experts was invited to participate in the expert elicitation and had expertise in NFC, RFID, mobile payments and security in general. The group of experts consisted of TNO employees from the ICT Security expertise group, authors of related work on NFC security, NFC developers and security experts. 27 of 39 experts selected for the elicitation actually responded. Only 19 of the experts completed all the 57 questions of the elicitation. The elicitation was performed using an online questionnaire for two reasons. First, an online questionnaire is a good way to guarantee objectivity since the analyst has less influence on the expert's opinion than in an interview. Second, experts with experience with NFC and security are hard to find. So, some of the experts are located in different parts of the world. Paying them a personal visit for an interview would be too expensive. The experts were asked to review the criteria ranking, assess the attack scenarios and answer a set of calibration questions. These three sets of questions elicit all the necessary inputs from the experts for the remaining part of the model. The elicitation contained 57 questions in total, which can be found in Appendix A. In the criteria ranking the experts were asked to give an ordinal ranking of the likelihood criteria described in Section 2.1. For the experts that were not familiar with the criteria in ETSI TISPAN TVRA or ISO 18045, a description for each of the evaluation criteria was available. The criteria ranking phase of the elicitation contained only one closed-end ordinal ranking question. The attack scenarios phase contained a brief version of the ten attack scenarios described in Section 4.1.2. Furthermore, an example of the possible fraud scenarios arising was given for each specific attack scenario. As mentioned before, the fraud scenarios were not part of the assessment. During this phase, there was also a glossary available for some technical definitions. The experts were asked to assess these

attack scenarios using the criteria of the ISO 18045 standard. This phase contained 50 questions, five closed-end questions for each attack scenario, corresponding to the five evaluation criteria. The criteria used in the elicitation were extracted from the ISO 18045 standard. The ISO 18045 standard adds an extra option to two of the criteria compared to the ETSI TISPAN TVRA [27]; the option multiple experts for the specialist expertise criterion and the option multiple bespoke for equipment availability. This ISO standard is part of a larger security evaluation framework consisting of multiple ISO standards, so it was expected more experts would be familiar with this standard than with the ETSI TISPAN TVRA. However, the ISO 18045 lacked methods for impact and risk calculation. For this reason, the ETSI TISPAN TVRA is used in the rest of this thesis, since this did contain the necessary methods for impact and risk calculation. In case an expert selected one of these extra options from the ISO 18045, the option was changed to the second last option (expert or bespoke) while processing of the results in order to comply with the ETSI TISPAN TVRA. The calibration phase consisted of 6 calibration questions, which are used to determine the expert weights as described in Chapter 3. In these questions the experts were asked to provide a lower bound, a median and an upper bound for their answers. The questions are based on the research performed for this thesis and the literature. In Table 4.3, the calibration questions with the corresponding answers are listed.

Table 4.3: Calibration question used in the expert elicitation

Question	**Answer**
How many security measures are defined in the NFC ISO standards (ISO 16353, ISO 18092, ISO 21481 and ISO 28361) [1, 2, 10, 11]?	0
What is the expected worldwide mobile payment transaction value (in billion of dollars) in 2016? [3]	617
How many failed authentication attempts does it take before a user is permanently locked out of the secure element on a Galaxy Nexus running Android 4.0.4? [52]	10
What is the average added delay (in ms), when command is relayed over Wi-Fi during a relay attack?	150
How many weeks would it take a computer science student to set up a relay attack?	2
What percentage of vulnerabilities in computer systems and networks are easy to exploit, requiring only moderate computer skills? [48]	70

4.3 Processing elicitation results

4.3.1 Criteria ranking

The experts gave an ordinal ranking for the five evaluation criteria of the risk assessment; elapsed time, specialist expertise, knowledge of the system, window of opportunity and equipment availability. As explained in Section 2.2, these ordinal rankings are aggregated using the distance-

based ordinal consensus method of Cook [36, 35]. In order to apply the method of Cook on the criteria, a tool was written in the Java programming language [57]. This tool takes the results of the elicitation as input, as presented in Appendix B.1. It then tries to find the optimal ranking with the highest distance-based consensus by comparing different rankings to the rankings of the individual experts. It does this by taking all possible permutations of the criteria ranking ($5! = 120$ *permutations*) and using Equation 2.2 to calculate the distance between each expert ranking and each permutation. Once it has found the optimal ranking, it returns the optimal distance, optimal ranking and the distance between the optimal ranking and the ranking of the individual experts. If more than one optimal ranking are found, the ranking with the lowest standard deviation is shown.

4.3.2 Expert weights

The classical method of Cooke [38, 39] was used to determine the expert weights based on the calibration questions, as explained in Section 2.3. In order to process the results, the software package Excalibur v1.0 [58] was used. This software package was developed by the team of Cooke and is able to perform all the necessary operations to determine the expert weights. It takes the calibration questions ($x_1, ..., x_k, ..., x_N$, as shown in Table 4.3), the expert answers ($q_{k,e,p}$, as shown in Appendix B.3) and the method parameters (p, k and $Ind_\alpha(x)$) as input. It applies the set of equations presented in Section 2.3 to the input data to calculate the result. This software package was chosen, because it offered the functionality needed for the expert weights and it was developed by the original author of the method. This application requires several parameters and options for calculating the expert weights. First, an option to run a decision maker optimization is available. When running this optimization, the experts are aggregated into one single performance score. Since, the proposed model requires a weight value for every expert, the decision maker optimization is not calculated. The second parameter is the calibration power value, which determines the number of samples used to determine the weights. This value is not explained in the literature [41], but the documentation of the program shows this value needs to be set at 1.0 in order to be compliant with the literature. The third parameter is the significance level ($Ind_\alpha(x)$), which is set at 0.0 since this returns all the exact expert weights. The last parameter is the intrinsic range (k). It determines how much the expert can be off from the exact answer to be still considered as correct. This parameter set at 10%, which represents a uniform distribution with a range from -10% to $+10\%$ around the actual answer that is considered correct. The results of the program show the information and calibration score and the (normalized) expert weight for each expert. Besides the results of the method, it also offers a robustness analysis for the expert opinions and for the different questions. Furthermore, we assign experts that did not fill in the questionnaire completely a weight of 0.0. And having done so, the opinion of these eight experts on the attack scenarios will not be taken into account.

4.3.3 Expert view aggregation

Before the elicitation results can be used as input for the SMAA-TRI method, the individual results first need to be aggregated in a discrete probability distribution, as described in Section 3.3.1. As input the expert weights obtained in the previous section are used a combined with the individual answers of the experts, as presented in Appendix B.2 for each question of the risk assessment part of the elicitation. This results in a set of discrete probability distributions, as presented in Appendix B.4. The x values from the discrete probability distributions are extracted from the risk calculation in the security evaluation method, as described in Section 2.1.3. However, these values only contain the values for likelihood and not for the impact. When only the likelihood values are used, one can not draw conclusions on the risks of the evaluated subject. In the risk calculation of the security evaluation method, the risk is calculated by multiplying the assigned values for the likelihood and impact classes. This risk calculation method returns the results shown in Table 4.4. If the risk calculation is deduced to the individual criteria, an

Table 4.4: Results of the ETSI TISPAN TVRA risk calculation

Impact \ **Likelihood**	**Unlikely**	**Possible**	**Likely**
Low	Low	Low	Medium
Medium	Low	Medium	High
High	Medium	High	High

inverse proportional relationship between the impact and likelihood values exists. However, it is not simply possible to divide the likelihood values by the impact values, since some likelihood criteria values have a quantitative value of 0. Nevertheless, one can have the same effect by adding or subtracting the difference between the corresponding lower and upper class borders of the likelihood values form the SMAA-TRI method from the likelihood values. Given this information a formula can be derived to calculate a x risk value containing the impact and likelihood. This formula is shown in Equation 4.1, where the impact class i and likelihood class l can be determined by checking to which class the original x value should be assigned given the class borders. The classes for impact range from low (1) to high (3) and for likelihood the classes range from unlikely (1) to likely (3).

$$x_{risk} = x_{likelihood} + \frac{l-i}{|l-i|} b* \tag{4.1}$$

where $b_0 = 0$ and

$$b* = \begin{cases} |b_{l-i} - b_{(l-i)-1}| & \text{if } |l-i| > 0 \\ 0 & \text{if } |l-i| = 0 \end{cases} \tag{4.2}$$

The impact values of Section 4.1.3 are used for the impact. In order to make sure the right combinations of attack and fraud scenarios are turned into risks, the fraud/attack scenario matrix

in Table 4.1 is used.

4.4 SMAA-TRI

As described in Chapter 3, the proposed model for the robust risk assessment is based on the stochastic multi-criteria acceptability analysis tree (SMAA-TRI) method. This method can outrank the different attack scenarios into different classes based on the evaluation criteria, class borders and a criteria ranking. The evaluation criteria is represented by discrete probability distributions from the weighted expert elicitation answers. The criteria ranking is the distance-based consensus ranking, as explained in Section 2.2.

The stochastic multi-criteria acceptability analysis tree (SMAA-TRI) is performed in the Java program JSMAA 1.0.2 [59]. This program is able to process SMAA-TRI, of which the used implementation corresponds with the literature [46, 47]. This program was chosen, because it was freely available (released under GPL license) and it was developed by the author of the related literature [46, 47].

On the first encounter with the software, JSMAA only supported exact measurements, uniform distributions, Gaussian distributions, log normal distributions, logit normal distributions and beta distributions to represent the evaluation criteria. As explained in Section 3.3, the results of the expert elicitation are aggregated into discrete probability distributions. For this reason, support for a discrete probability distribution was implemented into the JSMAA software by the author of this thesis. As of January 2013, this addition is publicly available in version 1.0 of the JSMAA software.

By default, the JSMAA software requires a manual input for all values. Since the elicitation returned a large amount of variables, it would take hours to enter the data into the program. Added to this, if a parameter change was applied to the expert weights, one has to enter the data all over again. Luckily, the software is able to save and open SMAA-TRI cases in a specially formatted XML file. A small Java program [60] was developed by the author of this thesis that was able to solve the manual input problem. This tool would take the criteria ranking, the expert weights and the elicitation results and output an XML file that could be read by the JSMAA software.

4.4.1 Parameters

Class borders

An important part of the SMAA-TRI method is the class borders. These borders determine if a certain criteria should be classified as low, medium or high likelihood. The class borders can be extracted from the security evaluation method by examining the risk calculation methods, as explained in Section 2.1.3 for the method used in this thesis. The class borders in the used method are not that obvious, since the criteria also have a ratio between them. For example,

the values of the elapsed time criterion range from 0 to 26, while expertise criterion ranges from 0 to 8. With this knowledge, three methods for extracting the class borders can be identified:

- By making the class division using the conversion table (see Table 2.2) and dividing it by the number of criteria. This will give a class border between Likely and Possible of 2 and a class border of 4 between Possible and Unlikely. However, this method does not really correspond to the assigned values for the individual criteria and the ratio between them;
- Dividing the maximum value for each class in the equal parts. In this case, each criterion will have different class borders. This method neglects the values from the conversion table and the ratio between the criteria, which results in a very pessimistic classification on the risks;
- Averaging the values of the previous method, so each criterion will use the same class borders. This will give a class border between Likely and Possible of 3 and a class border of 7 between Possible and Unlikely. Using this method will respect the values assigned to the criteria and the values from the conversion tables. It also respects the ratio between the criteria in a lesser extent, since only the size of the Unlikely class corresponds to the relation between the criteria. The only downside is that some criteria can not be assigned to the likely class.

The first method is used in the proposed model, since it covers all the criteria have values in the different classes and does not give an over-conservative classification of the risks.

SMAA-TRI thresholds

The preference threshold is set on an exact measurement with a value of 2.0 for all evaluation criteria. Due to the use of discrete probabilities and fact that for this case only integer criteria values are used, this threshold can only influence the results if the used value is larger than 1. Furthermore, the literature [46, 47] shows that a preference threshold between 5% and 20% of the maximum criteria value is commonly used for uncertain variables. These statements also apply to the indifference threshold, only here values between 5% and 10% of the maximum criteria value are more common, so an exact measurement of 1.0 is used for all evaluation criteria.

For the *lambda*-cutting level a uniform distribution with a range between 0.65 and 0.85 is used. This is the weighted overall concordance a scenario should have over a certain class to outrank this class. This range is chosen, because it is also used in the literature [46, 47].

The optimistic assignment rule is used to run the method. Bouyssou and Marchant [61, 62] describe the optimistic assignment rule as compensatory approach. In a compensatory approach, it is possible an alternative can be assigned to another class, if the value for one of its criteria changes within the class boundaries, while in a non-compensatory approach (pessimistic rule), this is not possible. Since the used security evaluation framework uses the summed criteria values to determine the likelihood, it is also possible for an attack scenario to be assigned to

another class when one criteria changes. Although in most cases for risk assessments [46, 47] the pessimistic assignment rule is used, given the definition of Bouyssou and Marchant and the security evaluation framework, the optimistic assignment rule is the best fit for the problem.

Chapter 5

Results

This chapter starts with Section 5.1, which will go into the processed results of the expert elicitation. Section 5.2 will go into the results of the model, when only the likelihood of the scenarios is taken into account. In the last section of this chapter, the risk classifications using the proposed model will be presented.

5.1 Elicitation results

Criteria weights The experts gave an ordinal ranking for the five evaluation criteria of the risk assessment; elapsed time, specialist expertise, knowledge of the system, window of opportunity and equipment availability. Due to the modular nature and the aggregation steps taken in the modeling the eight partially completed assessments are taken into account in the criteria weighting, but not in the attack scenario assessment. The results of the criteria ranking by experts can be found in Section B.1. When the rankings of the 26 experts were used as input for the application, it returned an optimal distance value of 318 and the optimal ranking, as shown in Table 5.1. An interesting fact for this combined ranking from the experts is that it

Table 5.1: Optimal criteria ranking

Rank	**Criterion**
1	Specialist expertise
2	Elapsed time
3	Window of opportunity
4	Knowledge of the system
5	Equipment availability

differs from the implicit criteria ranking in the ETSI TISPAN TVRA method on some criteria. In this method, specialist expertise has the lowest quantitative criteria values assigned from the criteria, whereas it receives the first rank from the experts. The rest of criteria is quite close

to the implicit ranking of the used security evaluation method. In here elapsed time is rated as highest, followed by window of opportunity, knowledge of the system and equipment availability.

Expert weights The experts were presented with 6 calibration questions in the expert elicitation for which they were asked to give a lower bound, a median and an upper bound. A detailed overview of the calibration answers given by each expert can be found in Section B.3. The experts that have completed the elicitation are used as input for the weight determination method. This gave expert weights in a range from 9.006×10^{-6} to 0.2880. Note that some weights are 0.0, this are experts that did not complete the survey and skipped the calibration questions. The (partly) completed assessments of the attack scenarios provided by these experts are therefore not taken into account in the model. The weight for each expert, as calculated with the method of Cooke, is shown in Table 5.2. Note that due to a rounding error in the Excalibur software, the weights do not exactly add up to 1.0. However, the error is quite small, so it does not have a significant effect on the final results.

Table 5.2: The expert weights calculated with the Classical method of Cooke.

Expert	**Weight**	**Expert**	**Weight**
1	0,00462000	14	0,00010560
2	0,00000901	15	0,00000000
3	0,00000000	16	0,12200000
4	0,28850000	17	0,00008797
5	0,00000000	18	0,03494000
6	0,00567300	19	0,03571000
7	0,00086970	20	0,00000000
8	0,11710000	21	0,03014000
9	0,01200000	22	0,00647100
10	0,00000000	23	0,00009114
11	0,00000000	24	0,00019090
12	0,03243000	25	0,00000000
13	0,00000000	26	0,15610000
		27	0,15300000

5.2 Likelihood acceptability

The results of the elicitation, as presented in Section 5.1, are used as input for the SMAA-TRI multi-criteria decision analysis method. As described in Chapter 3, the SMAA-TRI model gives class acceptability indices as result. For this section, the model is executed with only the likelihood values, as shown in the preference matrices in Appendix B.4. Since these values were elicited from experts, they are the most interesting to investigate. The likelihood class acceptability indices for this case are presented in Table 5.3 and Figure 5.1.

Table 5.3: Class acceptability for the NFC attack scenarios

Scenario	**Likely**	**Possible**	**Unlikely**
Relay Attack	0.1664	0.5598	0.2738
Relay Attack with malicious app	0.01	0.5102	0.4798
Eavesdropping with malicious app	0.0022	0.4627	0.5351
Eavesdropping	0.0739	0.2384	0.6877
Eavesdropping with a terminal exploit	0.0	0.1321	0.8679
Modify transactions with a terminal exploit	0.0001	0.1121	0.8878
Malicious terminal	0.0314	0.4698	0.4988
Denial of service with a zero-day exploit	0.0186	0.3024	0.679
Denial of service on the terminal	0.438	0.5023	0.0597
Theft	0.9487	0.0508	0.0005

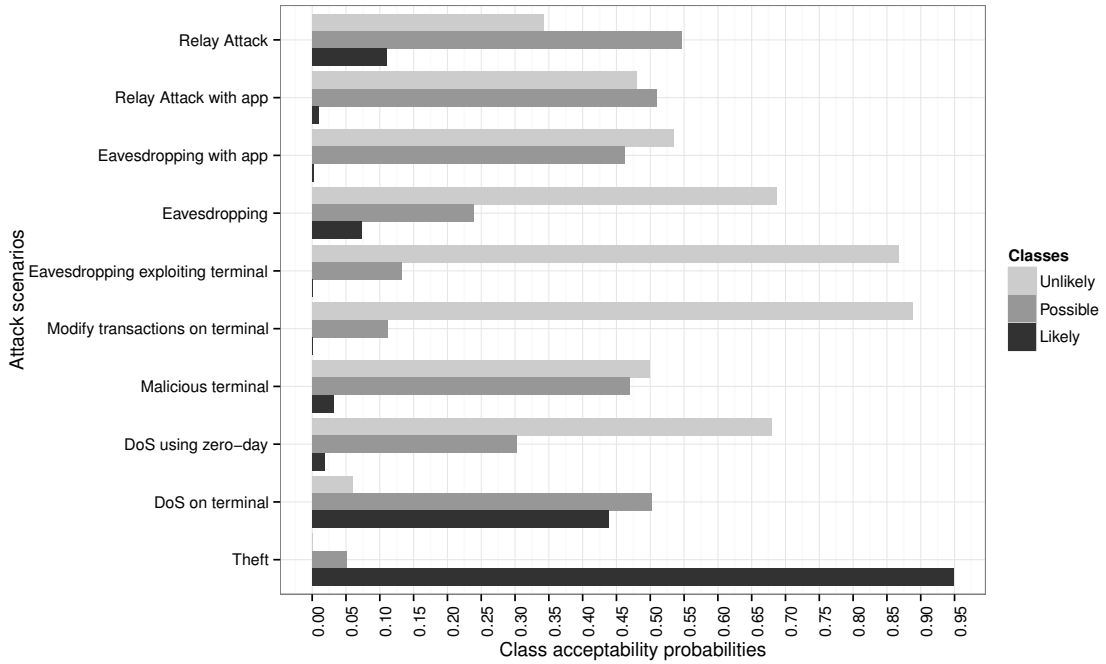

Figure 5.1: Class acceptability for the NFC attack scenarios

The most notable fact about the results is that all ten scenarios have a majority of the probability assigned to one class. For the ten scenarios, three classes have a distinct classification. The experts deemed the scenarios for eavesdropping by exploiting the terminal and the modify transactions on the terminal as a unlikely scenario. The theft scenarios has even a more distinct classification. Almost every expert opinion pointed towards a likely attack scenario. Furthermore, two scenarios are dealing with a minor dissent among the experts. The denial of service using a zero-day exploit has a clear majority in the possible class, but it also has a significant probability assigned to the unlikely class. For the eavesdropping scenario the majority of the probability is

assigned to the unlikely class, but also has a probability of 0.325 of being a possible scenario. The five remaining scenarios do not have a clear majority of the probability assigned in a single class. The eavesdropping with a malicious application, malicious terminal and the relay attack with a malicious app scenarios are all divided between a likelihood of possible and unlikely. The denial of service on the terminal has most of its class acceptability assigned to the likely and possible class. The relay attack scenario is divided among all three classes, with the majority in possible. In order to give an classification for these scenarios, more research is required.

5.3 Risk acceptability

In order to determine the risks for the NFC payment application case, both the likelihood and the impact need to be taken into account. For this section, the model is executed by using both the results from the expert elicitation and the impact assessment. This returns the class acceptability indices for 12 risk scenarios as shown in Table 5.4 and Figure 5.2. Note that in Figure 5.2 abbreviations are used for the different fraud scenarios.

Table 5.4: Class acceptability for the NFC security risks

Scenario	**High**	**Medium**	**Low**
Relay Attack Privacy Infringement	0.1108	0.551	0.3382
Relay Attack Single malicious transaction	0.0	0.0	1.0
Relay Attack with a malicious app Multiple malicious transaction	0.0115	0.5055	0.483
Eavesdropping with a malicious app Privacy Infringement	0.002	0.4687	0.5293
Eavesdropping Privacy Infringement	0.0748	0.2469	0.6783
Eavesdropping with a terminal exploit Privacy Infringement	0.0	0.134	0.866
Modify transactions with a terminal exploit Single malicious transaction	0.0	0.0	1.0
Malicious terminal Multiple malicious transaction	0.0325	0.4588	0.5087
Denial of service with a zero-day exploit Disable service	0.0	0.0	1.0
Denial of service on the terminal Disable service	0.0	0.0	1.0
Theft Privacy Infringement	0.9472	0.0518	0.001
Theft Multiple malicious transaction	0.9476	0.0519	0.0005

Most noticeable are four risk scenarios that received a class acceptability of 1.0. The denial

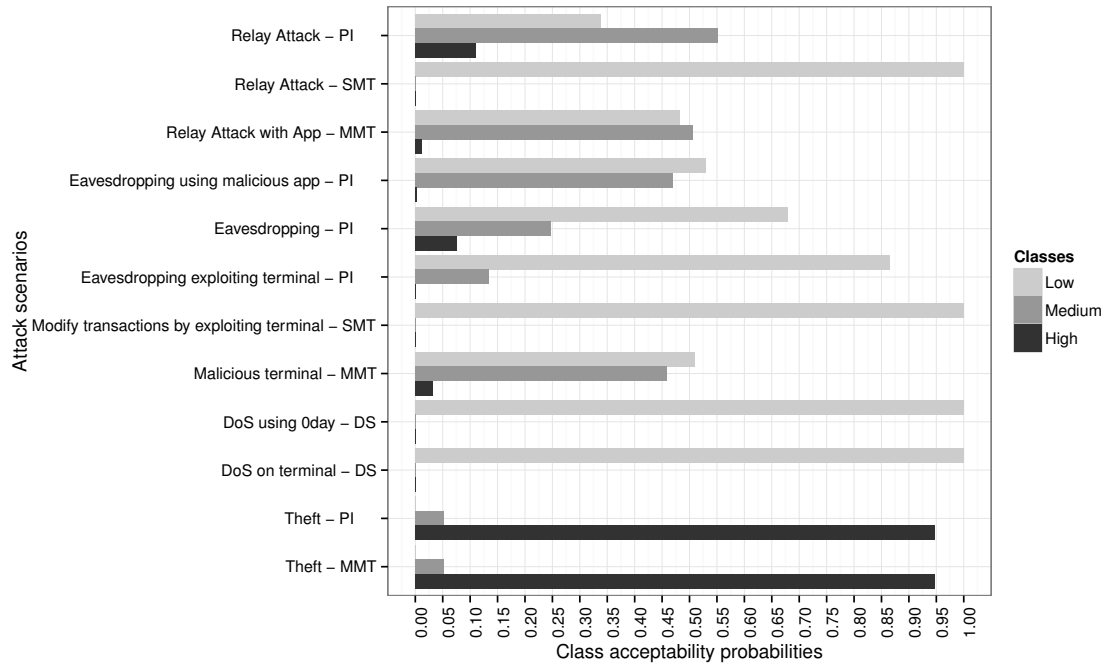

Figure 5.2: Class acceptability for the NFC security risks

of service on terminal with the disable service, the denial of service using a zero-day exploit with disable service, modify transactions on the terminal with single malicious transaction and relay attack with single malicious transaction are also deemed to have low risk. The theft scenarios in combination with privacy infringement and single malicious transaction are deemed high. Eavesdropping by exploiting the terminal with privacy infringement was also deemed to have low risk, but not with a probability of 1.0. Besides these seven risk scenarios with a distinct classification, privacy infringement by eavesdropping has a minor dissent among experts. This scenario has a majority of its probability assigned to the low risk class, but also has a significant probability of just being a medium risk scenario. Furthermore, five scenarios required more research before a classification could be given. The eavesdropping using an malicious app with the privacy infringement, the relay attack using a malicious app with multiple malicious transactions and the malicious terminal with multiple malicious transactions scenarios do not have a clear majority in a single class. While these three risk scenarios are mainly divided between medium risk and low risk, it is recommended to perform more research regarding these risks. Since it is common in risk assessment to go with the worst case scenario when in uncertainty [49, 50], these risk scenarios should be handled as medium risk scenarios, until proven otherwise. This is also the case for the relay attack with privacy infringement scenario. This scenario is divided among all three classes and should thereby be classified as high risk.

When an attack scenarios has multiple fraud scenarios attached, it is common to only use the

fraud scenario with the highest impact value in the final classification of an attack scenario. This means that for the relay attack only the risk scenario with privacy infringement is used for the final risk classification. For the theft scenarios, it does not matter which value is reported back to the client, since they received the same classification. The final classifications for all risks are presented in Table 5.5.

Table 5.5: Risk classifications for the NFC attack scenarios

Risk	**Class**
Theft	High
Relay Attack	High
Relay attack with malicious app	Medium
Eavesdropping with malicious app	Medium
Malicious terminal	Medium
Eavesdropping	Low
Eavesdropping with terminal exploit	Low
Denial of service on the terminal	Low
Denial of service with a zero-day exploit	Low
Modify transactions with a terminal exploit	Low

Chapter 6

Conclusion

This thesis tried to answer the main research question *What are the risks in payment applications for NFC-enabled smartphones?*. The research started by investigating the literature for already existing vulnerabilities in the security of near-field communication in general. This presented a set of interesting attacks. By performing a set of experiments with two NFC-enabled smartphones, it was possible to validate and even improve the attacks described in the literature. However, performing these attacks alone was not sufficient to answer the research question of this thesis. In order to answer the question, an expert-based quantitative risk analysis model had to be developed. So, ten attack scenarios and four fraud scenarios were extracted from the literature and the performed experiments. The fraud scenarios were assessed to determine the impact of an attack with information acquired from an anonymous payment provider. A group of 27 experts from the security, NFC and RFID fields were elicited on the likelihood of the attack scenarios using evaluation criteria from an existing security evaluation standard. In order to compensate for the fact not all the experts had expertise in all three fields, some calibration question were added to the elicitation. In these quantitative question the experts were asked to answer with a lower bound, median and upper bound. The experts were also asked to rank the evaluation criteria, to get an idea of the relative importance of the criteria. Thereafter, the opinions of the individual experts needed to be aggregated. The calibration questions were used to calculate a weight for each experts based on the knowledge and informativeness in the answers given by the experts. The assessment of the attack scenarios were condensed into weighted discrete probability distributions for each criterion for all scenarios using the calculated experts weights. Furthermore, the assigned quantitative values for each of the evaluation criteria was compensated with the impact to be able to calculate the risk. The rankings of the evaluation criteria were aggregated using a distance-based ordinal consensus ranking method. Next, the aggregated elicitation and assessment results were used as input for multi-criteria decision analysis method. This method is SMAA-TRI and can classify the different scenarios in different risk classes based on imprecise input values, such as probability distributions. The SMAA-TRI method provides

class acceptability indices as output rather than a crisp classification, showing the stability and probability of the different classifications are to the different risk classes.

6.1 Research questions

The research question of this thesis '*What are the risks in payment applications for NFC-enabled smartphones?*' was split into four sub questions. The first sub question is '*What are vulnerabilities in using NFC-enabled smartphones for payment applications?*'. This is answered by the literature in Chapter 1 and the performed NFC attack experiments in Chapter 4. These chapters showed that vulnerabilities exist in all facets of NFC payments. First of all, the NFC smartphone is not only a payment device, but also a all-round media device. This makes it vulnerable for malware in terms of malicious applications and other direct attacks, such as relay attacks. Second, the communication between the phone and the point-of-sale (PoS) terminal is vulnerable to eavesdropping, since in most devices this communication this is not encrypted. Although a standard for the encryption of the NFC communication channel [12] was recently released, this standard is not yet implemented in popular NFC-enabled smartphones. Furthermore, the point-of-sale terminals have multiple external communication channels, such as a card reader, a NFC reader and a secured internet connection to the payment provider. This makes the modern PoS terminals also vulnerable for malware, which also poses a threat to NFC payments. Lastly, NFC-enabled smartphones have to deal with the same physical threats as other payments methods, such as theft and skimming.

The second question is '*What is the likelihood these vulnerabilities will be exploited?*'. In order to answer this question, ten attack scenarios were presented in Chapter 4 based on the identified vulnerabilities. The likelihood for these attack scenarios was determined through expert elicitation. The results of the elicitation were processed and used as an input for a valued outranking method. Three scenarios received a distinct classification using this method. It classified theft as a likely attack scenario. Eavesdropping by exploiting the terminal and the modifications of transaction on the terminal were deemed unlikely according to the model. Furthermore, the method showed that two scenarios have to deal with a minor dissent among the experts. The denial of service using a zero-day exploit shows the highest probability to be unlikely, with a probability of 0.3 of being possible. Eavesdropping is classified as unlikely, yet it also has a probability of 0.24 to be possible. The remaining five attack scenarios require further research, before a clear classification can be given. Since, these attack scenarios do not have a strong probability in one of the likelihood classes, one should go with the worst class with a significant non-zero probability. So, the relay attack has probability indices assigned to all three classes. It has the majority of the probability in the possible class, but it also has significant probability for the likely class, therefore this scenario is classified as likely. The malicious terminal, the eavesdropping with a malicious application and the relay attack using a malicious app scenarios could be possible as well as unlikely, but are classified as possible

until further research is available. Last, the denial of service on the terminal has probability assigned to both the possible and the likely class, but is classified as likely until further research is available. In conclusion, three scenarios could be classified as likely, three as possible and four as unlikely.

The third sub question is '*What is the impact when these vulnerabilities are exploited?*' In order to answer this question, four fraud scenarios where drafted in Chapter 4 based on possible ways to exploit certain attack scenarios. The fraud scenarios could be linked to one or more attack scenarios. The impact was assessed using information from an anonymous payment service provider. The impact for the privacy infringement fraud scenario was deemed medium. The main reason for this was the possible reputation damage by the negative publicity a privacy infringement incident would get in the media. A single malicious transaction is expected to have a low impact, since the attacker is only able to gain small amounts from NFC payment users. Furthermore, payment providers usually take the risk of similar methods of attack, such as skimming and e-banking fraud. Therefore, the reputation damage of a single malicious transaction is also low. The threat of multiple malicious transactions has a medium impact. It is anticipated the attackers might gain an estimated 7.5 million euro in one large strike. Such a large strike can also catch the attention of the general media and have a medium impact at the reputation damage. These facts cause this fraud scenario to be classified as having an overall medium impact. Disabling the NFC payment service through a denial of service attack has a low impact. In case a jammer is used, the damage is contained to one merchant and has therefore a low impact. If malware is used to block access to the secure element on the phone, this causes an estimated maximum of 60 000 euro a week in financial damage. It is also expected this will not cause a loss of clients. In conclusion two of the fraud scenarios are deemed to have a medium impact and the two other fraud scenarios to have a low impact.

The last question is '*What is the view of experts on the risks connected to these vulnerabilities?*'. Since expert opinions were used to determine the risks for payment applications on NFC-enabled smartphones, this question also answers the research question '*What are the risks in payment applications for NFC-enabled smartphones?*'. The impact values were combined with the results of the expert elicitation, which were used as an input for the SMAA-TRI method to determine security risks. This analysis showed that seven risk scenarios have a distinct classification. Privacy infringement by theft and multiple malicious payments by theft are deemed as high risk. Disabled service by denial of service on the terminal, disabled service by denial of service using a zero-day exploit, single malicious transactions by modifying transaction by exploiting the terminal, and single malicious transactions by relay attacks are deemed as low risk with a probability of 1.0. Privacy infringement by eavesdropping on a exploited terminal is also deemed as a low risk. Furthermore, privacy infringement by eavesdropping has a minor dissent among the experts, having a majority of its probability assigned to the low risk class, but also a significant probability assigned to medium risk. The remaining four risk scenarios require more research in order to give a clear classification. The multiple malicious transactions

by using a malicious terminal scenario, multiple malicious transactions by a relay attack using a malicious app and privacy infringement by eavesdropping using a malicious application were divided between medium and low risk. For the sake of the uncertainty involved in these scenario, the scenarios should be considered as medium risk until proven otherwise. Privacy infringement by a relay attack has a similar case, where it is divided among all three risk classes. This risk scenario should therefore be considered as high risk until further research is available.

In conclusion, the risk scenarios can be converted into risks by only taking the scenarios classified with the highest risk value for the attack scenarios with multiple fraud scenarios. This turns the twelve risk scenarios into ten risks, five of which are classified as low risks. Three out of the ten risks are classified as medium risk. The remaining two risks are classified as high risk, where for one of the risk scenarios more research is required into the likelihood to give a clear classification. This shows that five risks require countermeasures to be implemented before a mass consumer release of NFC payment applications can take place.

6.2 NFC payments

This thesis looked into the risks of NFC payment applications on smartphones using an expert elicitation. Besides filling out the elicitation, some experts commented on the likelihood of the attack scenarios. One of the experts commented on the likelihood of the relay attack. This expert pointed out, that all the currently used payment specification were not vulnerable for relay attacks. In most payment specifications, random data is generated by the terminal during the payment processing, making a relay attack impossible. Only an old version of Visa Paywave specification is still vulnerable to this attack, but this isn't used anymore. Also, the literature [6] shows a simple countermeasure that can be taken against possible relay attacks. For example, by implementing timing restrictions on the responses from the phone to terminal, so not enough time is available to relay the data over another communication channel. Another expert commented on the modify transactions and eavesdropping by exploiting the terminal. This expert indicated the infecting a terminal with malware using an NFC device or RFID token is pretty much impossible. Infecting the terminal using other means of injection, like an Internet connection, is far more likely.

Although these countermeasures exist, it is possible a new NFC payment provider does not implement the necessary countermeasures. The NFC payment market is still very competitive without any clear standard. So, it is to be expected that more NFC providers arise who choose for a quick-to-market approach in order to secure a part of the NFC payments market, leaving security as a lower priority for them.

All in all, this thesis provides a fair overview of the different attack and fraud scenarios and corresponding security risks regarding NFC payment applications on smartphones for NFC developers and payment providers. Especially, the insight into Android card emulation patch and the fix to NFC proxy might contribute to a better practical understanding of the security

issues in the NFC card emulation mode in general.

6.3 Security Evaluation Model

In this thesis a model was proposed to obtain to a expert-based risk assessment under uncertainty. This model used expert elicitation to gather information on the likelihoods of the attack scenarios. The results of the elicitation were aggregated with impact information into weighted discrete probability distributions. The aggregated inputs were used in a valued outranking model in order to determine the overall security risk level.

This model has some advantages over a regular risk assessment that is performed by two or three security and domain experts in a meeting. Firstly, the usage of an electronic expert elicitation offers a security analyst to reach more experts. For this thesis, experts from seven different countries were contacted to contribute in the elicitation. When dealing with new technologies this offers an advantage. As the available knowledge on the security risks is limited, an elicitation on a larger group reduces the uncertainty. Also, a large group of experts with different opinions makes the risk assessment less biases. In a regular risk assessment, one is often limited by geographical constraints for the choice of experts, since this performed at a single location. Second, for risk assessments on new technologies, it might be hard to find experts that have the knowledge of the evaluated subject. If experts are found, the knowledge might be limited to a specific sub domain of the evaluated subject. The calibration questions in the expert elicitation can fix this by testing and evaluating the certainty and insight in the evaluated subject of the different experts. In a regular risk assessment, there are no initial weights assigned to the opinions of experts. This might cause a certain and knowledgeable expert not to be heard, because he lacks communication skills to persuade a more communicative expert. Lastly, the model gives a probability that a risk scenario belongs to a certain risk class instead of crisp classifications. This gives a security analyst more insight in the stability of the risk classifications. The results in Section 6.1 contained risk scenarios that required more research and did not have a clear classification. In a regular risk assessment, these scenarios were probably assigned to the lower majority class. By using the class acceptability indices, the analyst would notice the significant probability the scenario belonged to a higher risk class.

Besides these advantages of using the model proposed in this thesis for risk assessments, also some disadvantages exist. Firstly, it takes significantly longer to perform a risk analysis using the model than it would in a regular risk assessment, as performed in Appendix C. A regular risk assessment usually cost two to three people a couple of days, while only the elicitation of the experts for this thesis took a couple of weeks. The processing and interpreting the results of the expert elicitation took another two weeks in man hours, while the risk calculation in the ETSI TISPAN TVRA could be executed within an hour. Secondly, it is hard to represent the original risk assessment method into the model. In the case of the ETSI TISPAN TVRA, the class borders specified in this security evaluation method could not simply be copied into

the model, since the likelihood was determined by the sum of the chosen criteria values. It took some trial-and-error procedures to find the model parameters that best fitted the original risk assessment method. This is not only a time-consuming operation, but also might cause a loss of information on the likelihood calculation during the process, as shown in Section 4.4.1. Furthermore, the foundation of the model is less transparent than in a regular risk analysis, where one can simply show the risk calculations. In the model, far more calculations are made to get to a classification for a risks. This makes it harder for a security analyst to justify and explain to the client why a scenario has a certain risk classification. Especially when the client is not familiar with the methods used in the proposed model. Lastly, one of the experts pointed out that the discussion among experienced experts in a regular risk assessment offered some extra value to the risk assessment. Recent studies [63] in group decision-making indeed show that decision-making with two persons outperforms decision-making by a single person. However, the model uses the expert weights and the SMAA-TRI method to compensate for uncertainty in the experts opinions.

One should also note the importance of the calibration questions in the model. One should define these questions with great care, since they have a large influence on the outcome of the model. If the questions are not detailed enough, misunderstandings can arise, which has a negative effect on the score of an expert. For example, one of the experts pointed out the questions on number of authentication tries before lockout on the secure element was wrong, since this number is different among phones and Android versions. Luckily, this was one of the first experts, so it was still possible to correct this question. One should also define about eight to twelve calibration questions. The six question used in this thesis were a bit too few, since some of the experts that were expected to be less knowledgeable on the subject, were still able to obtain good scores by guessing.

All in all, the proposed model especially shows to have added value in risk assessments with high uncertainty and where expertise on the evaluated subject is hard to find. The fact that different methods for quantification of uncertainty in a risk assessment are used, makes the proposed model suitable for this. New technologies, like NFC payments on smartphones, are good example of this. However, for risk assessments on more common subjects, like web services, the disadvantages, such as the required time and effort to elicit the model parameters, might outweigh the advantages.

6.4 Future research

Payments using near-field communication are still a new way of making payments compared to the established methods. This makes it interesting for hackers and security specialists to find new exploits. Almost every year, new attacks on NFC are presented at the large hacker conventions, such as Blackhat and DefCon. This shows research in finding new attacks for NFC is still necessary. But also existing attack scenarios still require further research, as shown in

this thesis. Especially the eavesdropping with a malicious application, the malicious terminal and both the relay attack scenarios require some additional research into the likelihood of these attacks, before a classification can be given. It is also interesting to design countermeasures for the already known attacks on NFC payment application and, more importantly, to create awareness regarding the existing attacks and countermeasures among existing and future payment providers. These subjects for future research hopefully hardens the NFC payment security landscape.

With regard to the proposed security evaluation model, there are also some subjects for further research. The impact has a large influence on the calculation of the different risks However, the impact calculation in the model is performed without the use of multiple expert opinions, making it bias and sensitive for uncertainty. It makes sense to also use multiple security experts for the impact estimation for the fraud scenarios, so it will also return a discrete probability distribution. Incorporating this into the model will make the risk assessment less bias and gives a more insight into the risks involved. Furthermore, since only one security evaluation method was used, it was not possible to adequately validate the model presented. By incorporating different security evaluation methods into the proposed model could possibly give more insight in to the validity of the proposed model. Also, in the calibration questions is asked for estimations in different domains, so it might happen a knowledgeable expert from a certain is close to the correct answer on the questions in his domain and the expert is way off on questions outside of his domain. In this case, an expert receives a lower weight than he or she deserves. For further research, it is interesting to look at the possibility to take only the best answers of expert into account, so experts in a specific domain get the expert weight they deserve. Lastly, it might be possible to reflect the relation between the criteria and the risk calculation in the ETSI TISPAN and the proposed model more accurately by inferring the class borders from assignment examples [44] or using partial values functions [64]. By improving on those four points for future research will hopefully give a risk assessment model that can be used in quantitative risk assessment with high intrinsic uncertainties.

Bibliography

[1] International Organization for Standardization/International Electrotechnical Commission, ISO/IEC 18092 Information technology — Telecommunications and information exchange between systems — Near Field Communication — Interface and Protocol (NFCIP-1), ISO/IEC 18092:2004(E) (2004).

[2] International Organization for Standardization/International Electrotechnical Commission, ISO/IEC 21481 Information technology – Telecommunications and information exchange between systems – Near Field Communication Interface and Protocol -2 (NFCIP-2), ISO/IEC 21481:2012 (2012).

[3] S. Shen, Forecast: Mobile payment, worldwide, 2009-2016, Tech. rep., Gartner (2012).

[4] L. Francis, G. Hancke, K. Mayes, K. Markantonakis, Practical NFC peer-to-peer relay attack using mobile phones, in: Sixth International Conference on Radio Frequency Identification: Security and Privacy Issues (RFIDSec 2010), Springer-Verlag, 2010, pp. 35–49. doi:10.1007/978-3-642-16822-2_4.

[5] M. Roland, Software card emulation in NFC-enabled mobile phones: Great advantage or security nightmare, in: Fourth International Workshop on Security and Privacy in Spontaneous Interaction and Mobile Phone Use (IWSSI/SPMU 2012), 2012, p. 6, in press.

[6] J.-H. Hoepman, J. Siljee, Beyond RFID: the NFC security Landscape (2007) 15.

[7] Makezine, DEFCON RFID world record attempt, http://www.webcitation.org/6BzqiuSpX, retrieved 12 September 2012 (2005).

[8] International Organization for Standardization/International Electrotechnical Commission, ISO/IEC 14443 Identification cards – Contactless integrated circuit cards – Proximity cards, ISO/IEC 14443 (2001).

[9] NFC Forum, NFC Data Exchange Format (NDEF) Technical Specification (2006).

[10] International Organization for Standardization/International Electrotechnical Commission, ISO/IEC 28361 Information technology — Telecommunications and information exchange

between systems — Near Field Communication Wired Interface (NFC-WI), ISO/IEC 28361:2007 (2007).

[11] International Organization for Standardization/International Electrotechnical Commission, ISO/IEC 16353 Information technology — Telecommunications and information exchange between systems — Front-end configuration command for NFC-WI (NFC-FEC), ISO/IEC 16353 (2011).

[12] European Computer Manufacturers Association, ECMA-385 NFC-SEC: NFCIP-1 Security Services and Protocol, ECMA-385 (2010).

[13] Nokia, Nokia 6131 NFC phone taps into mobile payment, ticketing and local sharing, http://www.webcitation.org/6BzqpnrIN, retrieved 11 September 2012 (2007).

[14] NFC News, DCM to deploy NFC posters in UK cinemas, http://www.webcitation.org/6Daeh3oMS, retrieved 11 September 2012 (2012).

[15] M. Reijnders, Mobiele telefoon wordt toegangskaartje bij Roda JC, Webwereld, http://www.webcitation.org/6BzqxEQ0U, retrieved 11 September 2012 (2005).

[16] Payter, De grootste proef in europa, http://www.webcitation.org/6Bzqy4Yy3, retrieved 11 September 2012 (2010).

[17] M. Keferl, Near-field communication is shifting marketing in japan, Ad Age, http://www.webcitation.org/6BzqyzYBh, retrieved 25 september 2012 (2012).

[18] Google, Google wallet, http://www.webcitation.org/6Bzr98XZO, retrieved 11 September 2012 (2011).

[19] NFC Forum, Essentials for successful NFC mobile ecosystems, http://www.webcitation.org/6BzrFp3s4, retrieved 12 September 2012 (2009).

[20] A. Juels, RFID security and privacy: a research survey, IEEE Journal on Selected Areas in Communications 24 (2) (2006) 381–394. doi:10.1109/JSAC.2005.861395.

[21] J. Libbenga, Dutch transit card crippled by multihacks, The Register, http://www.webcitation.org/6BzrLWhwO, retrieved 14 September 2012 (2008).

[22] E. Lee, NFC Hacking: The Easy Way, in: DEFCON 20, 2012.

[23] C. Miller, Exploring the NFC attack surface, in: Blackhat 2012 whitepaper, 2012, p. 44.

[24] C. Mulliner, Vulnerability analysis and attacks on nfc-enabled mobile phones, Fourth International Conference on Availability, Reliability and Security (ARES 2009) (2009) 695–700doi:10.1109/ARES.2009.46.

[25] U. Ries, Phishing via NFC, The H Security, http://www.webcitation.org/6BzrM8Qmp, retrieved 26 September 2012 (2012).

[26] R. Borgaonkar, USSD/Android Dailer vulnerability, ekoparty Security Conference 8th edition, http://www.webcitation.org/6DW71H3uK (June 2012).

[27] European Telecommunications Standards Institute (ETSI), Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); Methods and protocols; Part 1: Method and proforma for Threat, Risk, Vulnerability Analysis (2011).

[28] P. Overbeek, E. R. Lindgreen, M. Spruit, Informatiebeveiliging onder controle, 2nd Edition, Pearson Education Benelux, 2005, ISBN 90-430-0692-0.

[29] C. J. Alberts, A. J. Dorofee, Operationally Critical Threat, Asset, and Vulnerability Evaluation Criteria, Version 2.0, Software Engineering Institute (2001).

[30] P. Mell, K. Scarfone, S. Romanosky, A Complete Guide to the Common Vulnerability Scoring System Version 2.0, Forum of Incident Response and Security Teams (FIRST), 2nd Edition (2007).

[31] V. Belton, T. J. Stewart, Multiple Criteria Decision Analysis, Springer, 2001, ISBN 978-0-7923-7505-0.

[32] Y.-P. O. Yang, H.-M. Shieh, J.-D. Leu, G.-H. Tzeng, A VIKOR-based multiple criteria decision method for improving information security risk, International Journal of Information Technology & Decision Making 8 (02) (2009) 267–287. doi:10.1142/S0219622009003375.

[33] Y.-M. Wang, T. M. Elhag, Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment, Expert Systems with Applications 31 (2) (2006) 309–319. doi:10.1016/j.eswa.2005.09.040.

[34] International Organization for Standardization/International Electrotechnical Commission, ISO/IEC 18045 Information technology — Security techniques — Methodology for IT security evaluation, ISO/IEC 18045:2008(E) (2008).

[35] W. D. Cook, Distance-based and ad hoc consensus models in ordinal preference ranking, European Journal of Operational Research 172 (2) (2006) 369–385. doi:10.1016/j.ejor.2005.03.048.

[36] W. Cook, M. Kress, Relationships between l^1 metrics on linear ranking spaces, SIAM Journal on Applied Mathematics 44 (1) (1984) 209–220. doi:10.1137/0144016.

[37] R. M. Cooke, K. Shrader, Experts in Uncertainty : Opinion and Subjective Probability in Science: Opinion and Subjective Probability in Science, Environmental ethics and science policy, Oxford University Press, USA, 1991.

[38] R. Cooke, L. Goossens, Procedures guide for structural expert judgement in accident consequence modelling, Radiation Protection Dosimetry 90 (3) (2000) 303–309.

[39] T. Bedford, R. Cooke, Probabilistic risk analysis: foundations and methods, Cambridge University Press, 2001. doi:10.1017/CBO9780511813597.

[40] S. Kullback, R. Leibler, On information and sufficiency., Annals of Mathematical Statistics 22 (1951) 79–86. doi:10.1214/aoms/1177729694.

[41] R. M. Cooke, L. L. Goossens, TU Delft expert judgment data base, Reliability Engineering & System Safety 93 (5) (2008) 657–674. doi:10.1016/j.ress.2007.03.005.

[42] B. Roy, Classement et choix en présence de points de vue multiples: La méthode ELECTRE, Revue Francaise d'Informatique et de Recherche Opérationnelle 8 (1968) 57–75.

[43] L. Dias, C. Antunes, ELECTRE models (a brief summary), VRTUOSI, http://www.webcitation.org/6BzrU0Yf6, retrieved 9 October 2012 (2010).

[44] V. Mousseau, R. Slowinski, Inferring an ELECTRE TRI model from assignment examples, Journal of Global Optimization 12 (2) (1998) 157–174. doi:10.1023/A:1008210427517.

[45] M. Merad, T. Verdel, B. Roy, S. Kouniali, Use of multi-criteria decision-aids for risk zoning and management of large area subjected to mining-induced hazards, Tunnelling and Underground Space Technology 19 (2) (2004) 125–138. doi:10.1016/S0886-7798(03)00106-8.

[46] T. Tervonen, J. R. Figueira, R. Lahdelma, J. A. Dias, P. Salminen, A stochastic method for robustness analysis in sorting problems, European Journal of Operational Research 192 (1) (2009) 236–242. doi:10.1016/j.ejor.2007.09.008.

[47] T. Tervonen, I. Linkov, J. Figueira, J. Steevens, M. Chappell, M. Merad, Risk-based classification system of nanomaterials, Journal of Nanoparticle Research 11 (2009) 757–766. doi:10.1007/s11051-008-9546-1.

[48] J. J. Ryan, T. A. Mazzuchi, D. J. Ryan, J. L. de la Cruz, R. Cooke, Quantifying information security risks using expert judgment elicitation, Computers & Operations Research 39 (4) (2012) 774–784. doi:10.1016/j.cor.2010.11.013.

[49] L. Sun, R. P. Srivastava, T. J. Mock, An information systems security risk assessment model under dempster-shafer theory of belief functions, Journal of Management Information Systems 22 (4) (2006) 109–142.

[50] M. D. Rogers, Scientific and technological uncertainty, the precautionary principle, scenarios and risk management, Journal of Risk Research 4 (1) (2001) 1–15. doi:10.1080/136698701455997.

[51] D. Yeager, Card emulation patch, CyanogenMod Code Review, http://www.webcitation.org/6Ci6WgpXM, retrieved 6 december 2012 (2012).

[52] N. Elenkov, Android secure element execution environment, http://www.webcitation.org/6Ci6sBaJF, retrieved 6 december 2012 (2012).

[53] N. Elenkov, Exploring google wallet using the secure element interface, http://www.webcitation.org/6Ci6x6ZO2, retrieved 6 December 2012 (2012).

[54] Dutch Ministry of Security and Justice, Wet bescherming persoonsgegevens (2001).

[55] Dutch Ministry of Security and Justice, Wet bescherming persoonsgegevens, chapter 10, article 66 (2001).

[56] Currence, Gemiddeld chipknipbetaalbedrag per maand per jaar, http://www.webcitation.org/6EOvHIfM2, retrieved 13 February 2013 (2013).

[57] R. Vermaas, Expert consensus application, https://github.com/rayz90/thesis-expertconsensus (2013).

[58] R. Cooke, EXpert CALIBRation version 1.0 Pro (EXCALIBUR), http://www.webcitation.org/6DaeNbRqI, retrieved 10 January 2013 (2007).

[59] T. Tervonen, JSMAA: open source software for SMAA computations, International Journal of Systems Science [to appear] (2012) 1–13. doi:10.1080/00207721.2012.659706.

[60] R. Vermaas, Result processor application, https://github.com/rayz90/thesis-resultprocessor (2013).

[61] D. Bouyssou, T. Marchant, An axiomatic approach to noncompensatory sorting methods in MCDM, I: The case of two categories, European Journal of Operational Research 178 (1) (2007) 217–245. doi:10.1016/j.ejor.2006.01.027.

[62] D. Bouyssou, T. Marchant, An axiomatic approach to noncompensatory sorting methods in MCDM, II: More than two categories, European Journal of Operational Research 178 (1) (2007) 246–276. doi:10.1016/j.ejor.2006.01.033.

[63] A. Koriat, When are two heads better than one and why?, Science 336 (6079) (2012) 360–362. doi:10.1126/science.1216549.

[64] M. Kadzinski, T. Tervonen, Stochastic ordinal regression for multiple criteria sorting problems, Decision Support Systems [to appear]. doi:10.1016/j.dss.2012.12.030.

[65] A. Greenberg, Shopping for zero-days: A price list for hackers' secret software exploits, Forbes, http://www.webcitation.org/6DUekiRi7, retrieved 7 January 2013 (2012).

Appendix A

Expert elicitation survey

Raymond Vermaas - 322126

March 20, 2013

Expert elicitation on the security risks of mobile payment applications

Introduction

I am currently writing my master thesis on the security risks of mobile payment applications using near-field communication (NFC) at the Dutch organization for applied scientific research TNO for my master Economics & Informatics at the Erasmus University Rotterdam. A part of my thesis is a risk analysis for different attacks based on the opinion of experts in the NFC security domain. Since you received an invitation to this elicitation, I consider you an expert, because you contributed to the security of NFC, mobile payments or another domain linked to the security of NFC.

The elicitation consists of three parts and takes about 15 to 20 minutes in total. In the first part you are asked to rank different criteria that are used for assessing the attack scenarios. In the second part you are confronted with 10 attack scenarios, which you need rank based on the criteria from the first part of the elicitation. In the last part you are asked to fill in a couple of calibration questions. These questions will be used to calibrate the model that aggregates the inputs from different experts.

I hope you are willing to contribute to my thesis.

Kind regards,
Raymond Vermaas

Criteria ranking

The criteria used for evaluating the likelihood of an attack are extracted from the ISO 18045 standard. The criteria are defined as follows:

- a) Time taken to identify and exploit (Elapsed Time);
- b) Specialist technical expertise required (Specialist Expertise);
- c) Knowledge of the system design and operation (Knowledge of the system);
- d) Window of opportunity;
- e) IT hardware/software or other equipment required for exploitation.

Click here for an extended description of the criteria

***1. Please rank these five criteria from most contributing to the likelihood of an attack (1) to least contributing to the likelihood of an attack (5).**

	Specialist Expertise
	Elapsed Time
	Knowledge of the system
	Window of opportunity
	Equipment availability

Attack scenarios

You now will be confronted with ten attack scenarios that need to be evaluated using the five criteria from the previous section of this elicitation. Each scenario starts with vulnerability and is followed by the attack steps of the attack scenario. The attack scenario also contains possible fraud scenarios. These fraud scenarios are not up for evaluation,

Expert elicitation on the security risks of mobile payment applications

but just give an example in which fraud scenarios the attack can be used.

Tip: by clicking on the questions, you will get the extended information on the criterion.

1. Relay Attack

Vulnerability
In some smartphones, NFC is always on, even if the smartphone is not in use. This makes it possible to perform a transaction with the payment application on the phone, which makes the smartphone vulnerable for a so-called relay attack.

Attack scenario
In a relay attack, there are two attackers. One attacker has a relay device and the other attacker has a proxy device. These devices can be a smartphone or another NFC-enabled device and are connected with each other over Internet. The relay device is held close to the victim's smartphone in a crowded place, like in public transport during rush hour. The proxy device is used to perform an NFC payment at a payment terminal. The communication between the victim's phone and the payment terminal is relayed over the proxy and relay.

Possible fraud scenarios
This exchange makes it possible for the attacker to perform a payment using the victim's card, in which case the attacker receives free goods and the victim is robbed. Since the data sent between the victim's smartphone and the payment terminal may not be encrypted, it could also be possible to obtain transaction information, like account numbers. Although this scenario does not directly harm the payment platform provider, it can cause reputational damage, possible claims from victims and might even lead to a decline in users.

*2. Elapsed time in weeks

*3. Specialist expertise

- Layman
- Proficient
- Expert
- Multiple experts

*4. Knowledge of the system

- Public information
- Restricted information
- Sensitive information
- Critical information

*5. Window of opportunity

- Unlimited access
- Easy access
- Moderate access
- Difficult access
- No access

Expert elicitation on the security risks of mobile payment applications

*6. **Equipment availability**

- Standard equipment
- Specialized equipment
- Bespoke equipment
- Multiple bespoke equipment

2. Relay attack using malicious app

Vulnerability
Many smartphone users gained superuser privileges (also known as rooting or jail-breaking) on their phone. This allows more advanced users to gain full control over their phone, but it also circumvents some of the security features of the phone. The secure element, in which NFC payment application resides, is also protected by these security features.

Attack scenario
An attacker tricks the victim in installing a malicious app by offering an interesting feature or hack. The victim thinks he grants the app access rights for the feature. The app uses the access rights to perform the feature, but meanwhile grants itself access to the secure element of the smartphone. The app notifies the attacker, it gained access to the secure element. The attacker can now perform a payment using the payment details of the victim, which are relayed from secure element on the victim's phone to the NFC-enabled smartphone of the attacker.

Possible fraud scenarios
With access to the secure element, the attacker can perform a relay attack, as described in the previous attack scenario. The only difference is the victims phone acting both as target and as relay.

*7. **Elapsed time in weeks**

*8. **Specialist expertise**

- Layman
- Proficient
- Expert
- Multiple experts

*9. **Knowledge of the system**

- Public information
- Restricted information
- Sensitive information
- Critical information

Expert elicitation on the security risks of mobile payment applications

***10. Window of opportunity**

- Unlimited access
- Easy access
- Moderate access
- Difficult access
- No access

***11. Equipment availability**

- Standard equipment
- Specialized equipment
- Bespoke equipment
- Multiple bespoke equipment

3. Eavesdropping using malicious app

Vulnerability
In some implementations of NFC payment applications the account data is not saved on the secure element, but on the servers of the payment platform provider. In case of a rooted phone, the communication between the server and the phone is vulnerable for eavesdropping.

Attack scenario
An attacker tricks the victim in installing a malicious app by offering an interesting feature or hack. The victim thinks he grants the app access rights for the feature. The app uses the access rights to perform the feature, but meanwhile wraps itself around the payment application. The app sends data to the attackers server, such as account data and transaction details.

Possible fraud scenarios
With direct access to the payment app, the attacker can extract private information, such as account data and transaction data. It might also be possible, to gain access to the authentication credentials the payment app uses to authenticate itself with the server of the payment platform provider. By sending these credentials to the attacker, they can be used to make fraudulent payments.

***12. Elapsed time in weeks**

***13. Specialist expertise**

- Layman
- Proficient
- Expert
- Multiple experts

Expert elicitation on the security risks of mobile payment applications

***14. Knowledge of the system**

- Public information
- Restricted information
- Sensitive information
- Critical information

***15. Window of opportunity**

- Unlimited access
- Easy access
- Moderate access
- Difficult access
- No access

***16. Equipment availability**

- Standard equipment
- Specialized equipment
- Bespoke equipment
- Multiple bespoke equipment

4. Eavesdropping

Vulnerability
Most current NFC payment applications for smartphones make use of the EMV contactless payments protocol. In this protocol, the communication between the payment terminal and the NFC enabled smartphones is not encrypted.

Attack scenario
The attacker places specialized eavesdrop equipment near a payment terminal. The equipment listens in during a transaction between the payment terminal and a NFC-enabled smartphone.

Possible fraud scenarios
Since the transaction data may not be encrypted, it is possible for the attacker to extract private information, like account numbers. This information can be used for identity theft or fraudulent payments using the account data.

***17. Elapsed time in weeks**

***18. Specialist expertise**

- Layman
- Proficient
- Expert
- Multiple experts

Expert elicitation on the security risks of mobile payment applications

***19. Knowledge of the system**

- Public information
- Restricted information
- Sensitive information
- Critical information

***20. Window of opportunity**

- Unlimited access
- Easy access
- Moderate access
- Difficult access
- No access

***21. Equipment availability**

- Standard equipment
- Specialized equipment
- Bespoke equipment
- Multiple bespoke equipment

5. Eavesdropping by exploiting the terminal

Vulnerability
Terminals are essentially computers that allow for communication outside the system environment. In the case of payment terminals, these communication channels consist of a WAN connection with the payment service provider and a possible connection with a NFC-enabled smartphone using near-field communication. This makes the terminal susceptible for malicious inputs.

Attack scenario
An attacker creates a malicious NFC device containing an exploit for the payment terminal. The attacker injects the malicious code into the terminal during a transaction. The exploit gathers information during real transactions with NFC devices. The attacker returns to the payment terminal and gathers information. In case the terminal is connected to the Internet, the attacker could also retrieve information over the Internet.

Possible fraud scenarios
With this exploit, it is possible to gather information during transactions without the victim's knowledge, which causes privacy harm and possible identity theft.

***22. Elapsed time in weeks**

Expert elicitation on the security risks of mobile payment applications

***23. Specialist expertise**

- Layman
- Proficient
- Expert
- Multiple experts

***24. Knowledge of the system**

- Public information
- Restricted information
- Sensitive information
- Critical information

***25. Window of opportunity**

- Unlimited access
- Easy access
- Moderate access
- Difficult access
- No access

***26. Equipment availability**

- Standard equipment
- Specialized equipment
- Bespoke equipment
- Multiple bespoke equipment

6. Modify transactions by exploiting the terminal

Vulnerability
Terminals are essentially computers that allow for communication outside the system environment. In the case of payment terminals, these communication channels consist of a WAN connection with the payment service provider and a possible connection with a NFC-enabled smartphone using near-field communication. This makes the terminal susceptible for malicious inputs.

Attack scenario
An attacker creates a malicious NFC device containing an exploit for the payment terminal. The attacker injects the malicious code into the terminal during a transaction. With the exploit in place, the attacker might be able to modify, delete or create transactions. In case the terminal is connected to the Internet, the attacker might be able to trigger these transactions over the Internet.

Possible fraud scenarios
It might also be possible for the attacker to inject false payments, delete real payments or to modify real payments. This can result in fraudulent transactions for both the merchant and legitimate users of the NFC payment application.

Expert elicitation on the security risks of mobile payment applications

***27. Elapsed time in weeks**

***28. Specialist expertise**

- Layman
- Proficient
- Expert
- Multiple experts

***29. Knowledge of the system**

- Public information
- Restricted information
- Sensitive information
- Critical information

***30. Window of opportunity**

- Unlimited access
- Easy access
- Moderate access
- Difficult access
- No access

***31. Equipment availability**

- Standard equipment
- Specialized equipment
- Bespoke equipment
- Multiple bespoke equipment

7. Malicious terminal

Vulnerability
Not all terminals can always be trusted. An attacker can act like a merchant and be in possession of a NFC payment terminal. By having physical access to the terminal, it is possible for the attacker to modify the payment terminal and even to swap the original terminal for a fraudulent one.

Attack scenario
The attacker obtains an NFC payment terminal. The attacker modifies the payment terminal. The payment terminal is used in a legitimate transaction.

Possible fraud scenarios
The fraud scenarios are comparable to those of mobile debit card payment terminals used in the Netherlands, like showing a different amount to the user than is actually used in the transaction. An other example is that, the user is told the first transaction failed and he/she should perform the same transaction again.

Expert elicitation on the security risks of mobile payment applications

***32. Elapsed time in weeks**

***33. Specialist expertise**

- Layman
- Proficient
- Expert
- Multiple experts

***34. Knowledge of the system**

- Public information
- Restricted information
- Sensitive information
- Critical information

***35. Window of opportunity**

- Unlimited access
- Easy access
- Moderate access
- Difficult access
- No access

***36. Equipment availability**

- Standard equipment
- Specialized equipment
- Bespoke equipment
- Multiple bespoke equipment

8. Denial of service using zero-day vulnerability

Vulnerability
It is possible for a zero-day vulnerability to exists in the mobile operating system running on the NFC-enabled smartphone. A zero-day vulnerability can cause a privilege escalation in the mobile OS. Since the secure element in the smartphone might rely on the security features of the mobile OS as well, it could be vulnerable as well.

Attack scenario
The attacker finds a zero-day vulnerability in the mobile operating system and creates an exploit to use the vulnerability to its advantage. The attacker uses a malicious tag near a payment terminal to send the victim to a website that executes the exploit on the smartphone of the victim. With the exploitation of the privilege escalation, the attacker gains access to the secure element on the smartphone. It then tries authenticate itself with the secure element, using bogus authentication details. After multiple failed tries, the secure element will permanently lockout all users.

Possible fraud scenarios

A permanent lockout of the secure element renders the phone useless for any secure element-based application. It is essentially a permanent denial of service for the victim.

***37. Elapsed time in weeks**

***38. Specialist expertise**

- Layman
- Proficient
- Expert
- Multiple experts

***39. Knowledge of the system**

- Public information
- Restricted information
- Sensitive information
- Critical information

***40. Window of opportunity**

- Unlimited access
- Easy access
- Moderate access
- Difficult access
- No access

***41. Equipment availability**

- Standard equipment
- Specialized equipment
- Bespoke equipment
- Multiple bespoke equipment

9. Denial of service of the terminal

Vulnerability
NFC payments are also interesting for vending machines, since this involves small payments. In case of an integrated payment terminal in a vending machine, the terminal is left unattended.

Attack scenario
An attacker places a jammer near a payment terminal. When a victim tries to make a payment, the jammer corrupts or even blocks the data sent between the terminal and the victims smartphone.

Possible fraud scenarios
The attacker essentially preforms a denial of service attack, which hurts the revenue of the merchant and the payment platform provider.

Expert elicitation on the security risks of mobile payment applications

***42. Elapsed time in weeks**

***43. Specialist expertise**

- Layman
- Proficient
- Expert
- Multiple experts

***44. Knowledge of the system**

- Public information
- Restricted information
- Sensitive information
- Critical information

***45. Window of opportunity**

- Unlimited access
- Easy access
- Moderate access
- Difficult access
- No access

***46. Equipment availability**

- Standard equipment
- Specialized equipment
- Bespoke equipment
- Multiple bespoke equipment

10. Theft

Vulnerability
The smartphone used in transactions is in possession of the user. In most smartphones it is left up to the user to enable security settings in the smartphone. If no security is set, the phone can be used by anyone to perform transactions.

Attack scenario
The attacker steals the smartphone of the victim. The victim not immediately notices the smartphone is missing. The attacker uses the victims phone to perform NFC payment transactions.

Possible fraud scenarios
The attacker can perform financial transactions after having stolen the phone. It might take some time before the victim notices the phone is missing and blocks the account attached to the NFC payment application. Although, the security of the phone is the concern of the user, the payment platform provider can expect claims from victims, when

the phone is used in illegitimate transactions. It is also possible for the attacker to extract account information from transaction details to use for identity theft.

***47. Elapsed time in weeks**

***48. Specialist expertise**

- Layman
- Proficient
- Expert
- Multiple experts

***49. Knowledge of the system**

- Public information
- Restricted information
- Sensitive information
- Critical information

***50. Window of opportunity**

- Unlimited access
- Easy access
- Moderate access
- Difficult access
- No access

***51. Equipment availability**

- Standard equipment
- Specialized equipment
- Bespoke equipment
- Multiple bespoke equipment

Calibration questions

Below you find nine calibration questions. For each question, you are asked to provide a lower bound, a median and an upper bound estimate as answer, to express a credible interval that covers your opinion. Please, answer these questions without additional resources.

Note: The answers to these questions will NOT be used to determine knowledge, scholarship or intelligence. The answers will contribute to the quantification of uncertainty in the decision support model that aggregates the different opinions of experts and classifies the risks.

Expert elicitation on the security risks of mobile payment applications

***52. How many security measures are defined in the NFC ISO standards (ISO 16353, ISO 18092, ISO 21481 and ISO 28361)?**

lower bound

median

upper bound

***53. What is the expected worldwide mobile payment transaction value (in million of dollars) in 2016?**

lower bound

median

upper bound

***54. How many failed authentication attempts does it take before a user is permanently locked out of the secure element on a Galaxy Nexus running Android 4.0.4?**

lower bound

median

upper bound

***55. What is the average added delay (in ms), when command is relayed over Wi-Fi during a relay attack?**

lower bound

median

upper bound

***56. How many weeks would it take a computer science student to set up a relay attack?**

lower bound

median

upper bound

***57. What percentage of vulnerabilities in computer systems and networks are easy to exploit, requiring only moderate computer skills?**

lower bound

median

upper bound

Thanks

You completed the expert elicitation. Thank you for your time.

Appendix B

Expert elicitation results

B.1 Criteria ranking

Table B.1 shows how many experts put a criterion at a certain rank. The rating of one of the experts was left out of this assessment, because in his opinion, he was not able to give a general ranking for the evaluation criteria.

Table B.1: Results of the criteria ranking

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
Elapsed Time	6	5	4	6	5
Specialist Expertise	8	6	4	2	6
Knowledge of the System	4	6	5	7	4
Window of Opportunity	6	5	7	5	3
Equipment availability	2	4	6	6	8

B.2 Attack scenario assessment

In Table B.2 an overview of the expert attack scenario assessments is provided. The results are displayed for the 27 experts. Note that some experts are left out, since they did not rate any of the attack scenarios. However, partial results are left in the table, but they are not used in the model, since these experts received a weight of 0.0 for not completing the calibration questions. The scenarios are assigned an ordinal number from 1 to 10, which corresponds to the numbering of the scenarios in Appendix A. On account of the criteria, the time column represents the elapsed time in weeks. The other criteria are assigned ordinal numbers for each of the answers of the specific criteria. The ordinal numbers start with 1 and have the same ordering as the criteria in Appendix A. Note that the criteria in the elicitation are extracted from the ISO 18045 [34] standard, which adds the option multiple experts for the specialist expertise criterion and the

option multiple bespoke for equipment availability compared to the ETSI TISPAN TVRA [27], as mentioned in Section 4.2.

Table B.2: An overview of the expert attack scenario assessments

Expert	**Scenario**	**Time**	**Expertise**	**Knowledge**	**Opportunity**	**Equipment**
1	*1*	5	2	2	3	1
	2	10	3	2	3	1
	3	10	3	2	2	1
	4	8	2	1	3	2
	5	52	3	3	1	2
	6	52	3	2	2	2
	7	52	3	2	2	2
	8	10	2	1	3	1
	9	5	1	1	5	1
	10	4	1	1	4	1
2	*1*	1	2	1	2	1
	2	1	2	1	2	2
	3	2	3	2	2	2
	4	1	2	1	2	1
	5	5	3	3	3	2
	6	8	3	3	4	2
	7	4	2	2	3	2
	8	3	2	2	2	1
	9	1	2	1	1	2
	10	0	1	1	3	1
4	*1*	20	2	1	2	1
	2	20	2	2	2	1
	3	8	2	2	2	1
	4	8	3	1	3	2
	5	24	3	2	2	1
	6	24	3	3	2	1
	7	12	2	2	1	1
	8	24	3	1	3	1
	9	4	1	1	2	2
	10	1	1	1	2	1

Continued on next page

Table B.2 – *Continued from previous page*

Expert	Scenario	Time	Expertise	Knowledge	Opportunity	Equipment
5	*1*	52	4	4	1	1
	2	52	4	4	1	1
	3	-	-	-	-	-
	4	-	-	-	-	-
	5	-	-	-	-	-
	6	-	-	-	-	-
	7	-	-	-	-	-
	8	-	-	-	-	-
	9	-	-	-	-	-
	10	-	-	-	-	-
6	*1*	1	2	1	3	1
	2	2	2	1	2	1
	3	2	2	1	1	1
	4	2	2	1	4	1
	5	3	2	2	4	2
	6	3	2	2	4	2
	7	4	2	2	4	2
	8	1	2	1	2	1
	9	1	2	1	1	2
	10	0	2	1	1	1
7	*1*	2	2	1	3	2
	2	4	3	2	4	2
	3	1	2	1	2	1
	4	2	2	1	3	2
	5	12	3	2	4	2
	6	12	3	2	4	2
	7	8	3	2	4	2
	8	12	3	1	3	1
	9	2	2	1	3	2
	10	1	1	1	3	1
8	*1*	2	1	1	4	1
	2	5	3	1	3	2
	3	5	3	2	3	2
	4	6	3	1	2	2
	5	52	3	3	5	2
	6	52	3	4	5	2

Continued on next page

Table B.2 – *Continued from previous page*

Expert	Scenario	Time	Expertise	Knowledge	Opportunity	Equipment
	7	4	2	1	3	2
	8	3	2	1	2	1
	9	3	2	1	1	2
	10	0	1	1	2	1
9	*1*	0	1	1	2	1
	2	4	3	4	2	1
	3	1	2	1	2	1
	4	2	3	1	1	3
	5	6	3	3	1	2
	6	6	4	4	1	2
	7	1	2	1	1	1
	8	4	3	1	2	1
	9	1	2	1	1	2
	10	0	1	1	1	1
10	*1*	0	1	1	3	1
	2	-	-	-	-	-
	3	-	-	-	-	-
	4	-	-	-	-	-
	5	-	-	-	-	-
	6	-	-	-	-	-
	7	-	-	-	-	-
	8	-	-	-	-	-
	9	-	-	-	-	-
	10	-	-	-	-	-
12	*1*	6	3	2	2	2
	2	8	3	2	2	1
	3	12	2	3	3	1
	4	6	3	1	2	2
	5	10	3	3	3	2
	6	12	3	3	4	2
	7	14	3	3	3	2
	8	6	2	1	2	1
	9	2	1	1	2	2
	10	1	1	1	2	1

Continued on next page

Table B.2 – *Continued from previous page*

Expert	Scenario	Time	Expertise	Knowledge	Opportunity	Equipment
14	*1*	24	3	2	1	2
	2	2	2	1	2	1
	3	1	2	1	1	1
	4	4	3	2	4	3
	5	12	3	3	4	3
	6	26	3	4	4	3
	7	16	3	3	3	2
	8	24	3	3	2	1
	9	1	1	1	1	1
	10	1	1	1	1	1
16	*1*	2	2	1	2	1
	2	3	2	1	2	1
	3	4	3	2	3	1
	4	6	3	2	2	2
	5	6	3	3	4	2
	6	6	3	2	4	2
	7	3	2	2	2	1
	8	8	3	3	4	1
	9	1	1	1	1	1
	10	0	1	1	2	1
17	*1*	20	3	1	3	2
	2	12	3	1	2	1
	3	12	3	2	2	1
	4	12	3	2	3	2
	5	20	3	3	4	3
	6	20	3	3	4	3
	7	20	3	2	3	3
	8	16	2	1	1	2
	9	4	2	1	1	1
	10	4	1	1	1	1
18	*1*	4	3	2	1	2
	2	10	3	3	1	1
	3	6	3	1	1	1
	4	16	4	3	1	2
	5	20	4	3	1	2
	6	26	4	4	1	2

Continued on next page

Table B.2 – *Continued from previous page*

Expert	Scenario	Time	Expertise	Knowledge	Opportunity	Equipment
	7	16	4	4	1	2
	8	12	4	3	1	1
	9	8	3	1	1	1
	10	4	1	1	1	1
19	*1*	20	3	2	2	2
	2	20	3	1	1	1
	3	12	4	3	4	2
	4	15	4	3	4	3
	5	24	4	4	4	2
	6	24	4	3	4	3
	7	20	4	1	1	3
	8	15	3	1	1	2
	9	2	1	1	1	2
	10	1	1	1	3	1
21	*1*	2	3	1	2	2
	2	2	3	1	2	1
	3	2	3	1	2	1
	4	2	2	1	3	1
	5	10	3	2	3	1
	6	10	3	1	3	1
	7	10	3	1	4	2
	8	10	3	1	2	2
	9	1	2	1	2	2
	10	0	1	1	2	1
22	*1*	2	2	2	2	1
	2	4	3	1	3	1
	3	4	3	2	3	1
	4	2	2	1	2	2
	5	8	3	2	3	1
	6	8	3	2	3	1
	7	12	3	2	3	1
	8	2	2	1	2	1
	9	2	2	1	2	2
	10	2	1	1	2	1

Continued on next page

Table B.2 – *Continued from previous page*

Expert	Scenario	Time	Expertise	Knowledge	Opportunity	Equipment
23	*1*	4	3	1	1	1
	2	8	3	2	1	1
	3	8	3	1	1	1
	4	12	3	2	2	1
	5	12	4	2	1	2
	6	16	4	3	3	2
	7	20	4	3	3	3
	8	12	3	2	1	1
	9	1	1	1	1	1
	10	1	1	1	2	1
24	*1*	4	3	1	3	2
	2	8	2	2	3	1
	3	4	2	1	2	1
	4	3	2	1	2	2
	5	8	3	2	3	2
	6	12	3	3	3	2
	7	8	2	1	2	1
	8	10	2	2	1	1
	9	2	1	1	2	1
	10	1	1	1	2	1
25	*1*	4	3	2	4	2
	2	-	-	-	-	-
	3	-	-	-	-	-
	4	-	-	-	-	-
	5	-	-	-	-	-
	6	-	-	-	-	-
	7	-	-	-	-	-
	8	-	-	-	-	-
	9	-	-	-	-	-
	10	-	-	-	-	-
26	*1*	3	3	1	3	2
	2	6	3	1	2	3
	3	10	3	2	2	3
	4	8	3	1	2	3
	5	20	3	3	3	4
	6	20	3	3	3	3

Continued on next page

Table B.2 – *Continued from previous page*

Expert	Scenario	Time	Expertise	Knowledge	Opportunity	Equipment
	7	20	3	1	3	3
	8	6	4	4	3	4
	9	2	2	1	2	2
	10	2	1	1	3	1
27	*1*	1	3	1	3	2
	2	2	3	2	2	1
	3	1	3	2	2	1
	4	0	1	1	2	1
	5	2	3	3	3	2
	6	2	3	3	3	2
	7	2	3	2	2	2
	8	1	2	1	1	1
	9	1	3	2	2	2
	10	0	1	1	2	1

B.3 Calibration questions

The calibration question with the corresponding answers can be found in Table B.3. The answers the experts gave to each question are presented in Table B.4.

Table B.3: Calibration question used in the expert elicitation

#	Question	Answer
1	How many security measures are defined in the NFC ISO standards (ISO 16353, ISO 18092, ISO 21481 and ISO 28361) [1, 2, 10, 11]?	0
2	What is the expected worldwide mobile payment transaction value (in billion of dollars) in 2016? [3]	617
3	How many failed authentication attempts does it take before a user is permanently locked out of the secure element on a Galaxy Nexus running Android 4.0.4? [52]	10
4	What is the average added delay (in ms), when command is relayed over Wi-Fi during a relay attack?	150
5	How many weeks would it take a computer science student to set up a relay attack?	2
6	What percentage of vulnerabilities in computer systems and networks are easy to exploit, requiring only moderate computer skills? [48]	70

Table B.4: The expert answers to the calibration question (empty responses are left out)

Question	**1**			**2**			**3**			**4**			**5**			**6**		
	5%	*50%*	*95%*	*5%*	*50%*	*95%*	*5%*	*50%*	*95%*	*5%*	*50%*	*95%*	*5%*	*50%*	*95%*	*5%*	*50%*	*95%*
Expert 1	20	40	100	100	3000	6000	3	10	20	5	10	40	1	3	10	5	10	20
Expert 2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Expert 4	0	0	0	37	370	3700	3	5	10	1	4	10	1	4	8	10	50	70
Expert 6	5	7	10	5000	7000	10000	3	7	10	50	70	200	0	1	2	10	2	4
Expert 7	0	3	6	10	20	30	0	5	10	10	20	30	0	1	2	5	10	15
Expert 8	0	0	0	500	600	700	10	67	9999	100	200	400	0	4	999	1	5	20
Expert 9	4	10	9999	250	500	1200	3	5	10	15	40	150	0	1	2	50	60	75
Expert 12	5	15	25	1	5	10	6	8	10	100	230	500	6	8	12	60	75	90
Expert 14	30	50	60	1	3	6	1	3	6	0	1	5	2	4	8	10	50	90
Expert 16	40	80	120	3	13	25	3	10	15	100	200	400	0	2	3	20	50	75
Expert 17	40	75	150	150	175	200	2	3	6	10	20	50	12	20	26	50	65	80
Expert 18	0	0	0	100	200	600	3	7	10	100	300	1000	2	16	10	2	5	15
Expert 19	120	220	320	5	7	12	4	5	10	2	4	10	1	2	4	60	70	90
Expert 21	50	100	300	50	100	300	1	5	10	5	15	30	1	2	3	60	75	90
Expert 22	0	0	0	4	10	25	5	10	50	50	200	500	1	2	4	10	20	40
Expert 23	36	52	76	2	4	8	3	5	7	200	360	600	4	8	12	30	40	70
Expert 24	0	0	0	0	0	0	5	50	80	0	0	0	0	0	0	0	0	0
Expert 26	10	40	100	10	200	1000	3	7	100	10	200	3000	1	2	6	1	2	5
Expert 27	10	100	10000	100	1000	100000	3	5	10	1	80	300	1	8	20	1	10	90

B.4 Preference matrices: Likelihood

The tables in this section show the discrete probability distributions for likelihood calculation.

Table B.5: Discrete likelihood probability distributions for the specialist expertise criterion

Scenario	Layman	Proficient	Expert
1	0.1291	0.4281	0.4428
2	0.0	0.4165	0.5835
3	0.0	0.3398	0.6602
4	0.153	0.048	0.799
5	0.0	0.0057	0.9943
6	0.0	0.0057	0.9943
7	0.0	0.5455	0.4545
8	0.0	0.3196	0.6804
9	0.4836	0.3285	0.1879
10	0.9943	0.0057	0.0

Table B.6: Discrete likelihood probability distributions for the knowledge of the system criterion.

Scenario	Public	Restricted	Sensitive	Critical
1	0.8857	0.1143	0.0	0.0
2	0.4734	0.4797	0.0349	0.012
3	0.084	0.8479	0.0681	0.0
4	0.8071	0.1223	0.0706	0.0
5	0.0	0.3319	0.0357	0.6324
6	0.0301	0.1396	0.1641	0.6662
7	0.3512	0.5812	0.0349	0.03270
8	0.6866	0.0003	0.157	0.1561
9	0.847	0.153	0.0	0.0
10	1.0	0.0	0.0	0.0

Table B.7: Discrete likelihood probability distributions for the window of opportunity criterion.

Scenario	Unlimited	Easy	Moderate	Difficult	None
1	0.0351	0.5272	0.3206	0.1171	0.0
2	0.0707	0.8000	0.1284	0.0009	0.0
3	0.0408	0.6455	0.2780	0.0357	0.0
4	0.0469	0.5874	0.3242	0.0415	0.0
5	0.0516	0.2885	0.3783	0.1645	0.1171
6	0.0469	0.2931	0.346	0.1969	0.1171
7	0.3711	0.2798	0.3124	0.0367	0.0
8	0.2240	0.2039	0.4501	0.1220	0.0
9	0.3277	0.6668	0.0009	0.0	0.0046
10	0.0528	0.7499	0.1927	0.0046	0.0

Table B.8: Discrete likelihood probability distributions for the equipment availability criterion.

Scenario	**Standard**	**Specialized**	**Bespoke**
1	0.5564	0.4436	0.0
2	0.7259	0.118	0.1561
3	0.6911	0.1528	0.1561
4	0.1889	0.6072	0.2039
5	0.3251	0.5186	0.1563
6	0.3251	0.4829	0.192
7	0.4291	0.3789	0.192
8	0.778	0.0659	0.1561
9	0.162	0.838	0.0
10	1.0	0.0	0.0

Table B.9: Discrete likelihood probability distributions for elapsed time in number of weeks.

Scenario	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**
1	0.012	0.1587	0.2766	0.1561	0.0352	0.0046	0.0324	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3243	0.0	0.0	0.0	0.0001	0.0	0.0
2	0.0	0.0	0.1889	0.122	0.0193	0.1171	0.1561	0.0	0.0327	0.0	0.0396	0.0	0.0001	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3242	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.166	0.0358	0.0	0.1287	0.1171	0.0349	0.0	0.2886	0.0	0.1607	0.0	0.0682	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.153	0.0	0.0552	0.0002	0.0001	0.0	0.2715	0.0	0.4492	0.0	0.0	0.0	0.0002	0.0	0.0	0.0357	0.0349	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.153	0.0057	0.0	0.0	0.134	0.0	0.0067	0.0	0.0626	0.0	0.0011	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1911	0.0	0.0	0.0	0.3242	0.0	0.1217
6	0.0	0.0	0.153	0.0057	0.0	0.0	0.134	0.0	0.0065	0.0	0.0301	0.0	0.0335	0.0	0.0	0.0	0.0001	0.0	0.0	0.0	0.1562	0.0	0.0	0.0	0.3242	0.0	0.1568
7	0.0	0.012	0.153	0.122	0.1228	0.0	0.0	0.0	0.0011	0.0	0.0301	0.0	0.295	0.0	0.0324	0.0	0.035	0.0	0.0	0.0	0.192	0.0	0.0	0.0	0.0	0.0	0.0046
8	0.0	0.1587	0.0065	0.1171	0.012	0.0	0.1885	0.0	0.122	0.0	0.0349	0.0	0.0359	0.0	0.0	0.0357	0.0001	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2886	0.0	0.0
9	0.0	0.323	0.2318	0.1171	0.2886	0.0046	0.0	0.0	0.0349	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	0.4399	0.3579	0.1626	0.0	0.0396	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Appendix C

Initial risk assessment

In this appendix an initial risk assessment is performed, without input from experts, by the author of this thesis. The attack and fraud scenarios from Section 4.1.2 and 4.1.3 are used as input for this risk assessment. The ETSI TISPAN TRVA [27] standard is used to make the initial risk assessment.

C.1 Likelihood

The likelihood is assessed using the attack scenarios from Section 4.1.2.

Relay attack Taking into account the performed experiments concerning the relay attack, it was possible to accurately perform an assessment of the likelihood of the relay attack described in the attack scenario. A relay attack gets the rating likely as shown in Table C.1, since apps, like NFCProxy, are freely available it becomes quite easy to perform a relay attack.

Table C.1: Likelihood of a relay attack

Elapsed time in weeks	3
Specialist expertise	Proficient
Knowledge of the system	Public information
Window of opportunity	Easy access
Equipment availability	Standard Equipment
Result:	**Likely**

Relay attack using malicious app A relay attack using a malicious app seems somewhat less likely. It requires more time to develop such an attack. Also, more restricted information about the secure element is required. Given this information, this attack receives the rating possible from the security standard, as shown in Table C.2.

Table C.2: Likelihood of a relay attack using malicious app

Elapsed time in weeks	8
Specialist expertise	Expert
Knowledge of the system	Restricted information
Window of opportunity	Easy access
Equipment availability	Standard Equipment
Result:	**Possible**

Eavesdropping using malicious app A smartphone running an NFC payment app is an active device. This makes possible to use the smartphone of victim in man-in-the-middle attack. Especially in a rooted phone were there exists the possibility of an malicious app gaining access to the operating system of the smartphone. Because of these factors, the eavesdropping using a malicious app attack scenarios was rated as possible, as shown in Table C.3.

Table C.3: Likelihood of eavesdropping using malicious app

Elapsed time in weeks	6
Specialist expertise	Expert
Knowledge of the system	Restricted information
Window of opportunity	Easy access
Equipment availability	Standard Equipment
Result:	**Possible**

Eavesdropping As stated in the eavesdropping attack scenario there is no secure connection between the NFC device and the NFC terminal, which adds to the likelihood of this attack. However, an attacker also needs to be quite close, to actually be able to eavesdrop on a transaction, which makes this attack somewhat less likely. As shown in Table C.4, with all factors taken in consideration, this attack scenarios is still rated as likely.

Table C.4: Likelihood of eavesdropping

Elapsed time in weeks	2
Specialist expertise	Proficient
Knowledge of the system	Public information
Window of opportunity	Moderate access
Equipment availability	Specialized Equipment
Result:	**Likely**

Eavesdropping by exploiting the terminal Finding exploits for a payment terminal is not that easy. It takes time and one needs to have quite an understanding of the inner workings of a specific terminal. Table C.5 shows that the likelihood for this attack scenario is scaled at unlikely.

Table C.5: Likelihood of eavesdropping by exploiting the terminal

Elapsed time in weeks	13
Specialist expertise	Expert
Knowledge of the system	Restricted information
Window of opportunity	Moderate access
Equipment availability	Specialized Equipment
Result:	**Unlikely**

Modify transactions by exploiting the terminal This attack scenario is strongly related to the previous attack scenario. However, this scenario is even less likely, because modifying transactions is more visible for the victim than simply eavesdropping. As can be seen in Table C.6, this scenario is also rated as unlikely.

Table C.6: Likelihood of modifying transactions by exploiting the terminal

Elapsed time in weeks	16
Specialist expertise	Expert
Knowledge of the system	Restricted information
Window of opportunity	Moderate access
Equipment availability	Specialized Equipment
Result:	**Unlikely**

Malicious terminal If an attacker has physical access to a terminal, by for example posing as merchant, it becomes more easy to use the terminal as an attack platform. Therefore, this scenario ends up with a rating of possible, as shown in Table C.7.

Table C.7: Likelihood of a malicious terminal

Elapsed time in weeks	6
Specialist expertise	Proficient
Knowledge of the system	Public information
Window of opportunity	Easy access
Equipment availability	Specialized Equipment
Result:	**Possible**

Denial of service using a zero-day vulnerability Acquiring a zero-day for a mobile OS is quite easy. According to Forbes [65] you can acquire a zero-day exploit for between $30 000 and $60 000. Another problem is that most user neglect to get the latest software for their phone. Therefore, this scenario ends up with a rating of likely, as shown in Table C.8.

Denial of service of the terminal The attacker only needs to place a jammer near a NFC terminal, which makes it very easy to perform this attack. Consequentially, this attack is rated

Table C.8: Likelihood of a denial of service using a zero-day vulnerability

Elapsed time in weeks	6
Specialist expertise	Proficient
Knowledge of the system	Public information
Window of opportunity	Easy access
Equipment availability	Standard Equipment
Result:	**Likely**

as likely, as shown in Table C.9.

Table C.9: Likelihood of a denial of service of the terminal

Elapsed time in weeks	1
Specialist expertise	Layman
Knowledge of the system	Public information
Window of opportunity	Easy access
Equipment availability	Standard Equipment
Result:	**Likely**

Theft Just as with any payment method for consumers, theft is a very likely attack scenario. It requires little effort and expertise and is therefore rated as likely in Table C.10.

Table C.10: Likelihood of theft

Elapsed time in weeks	0
Specialist expertise	Layman
Knowledge of the system	Public information
Window of opportunity	Easy access
Equipment availability	Standard Equipment
Result:	**Likely**

C.2 Impact

The impact is hard to estimate without background of a specific company implementing NFC payment applications, as is described in Section 3.2.2. In able to still give estimate, it is considered a large multinational company, such as a telecommunication operator or a bank, which are typically the kind of organizations that take up the implementation of the NFC infrastructure. To determine the impact we consider the fraud scenarios described in Section 4.1.3.

Privacy infringement In case of privacy infringement, there is almost no direct financial impact. However, the vulnerabilities that led to possible privacy infringement are often the responsibility of the of the company implementing NFC for payment applications. Therefore,

the reputation damage and the loss of clients can contribute to a bigger in impact, when it happens on a large scale. Still, when compared to the other described fraud scenarios, privacy infringement has a low impact for the NFC payment provider.

Single malicious transaction This kind of fraud occurs frequently in the current payment systems, especially those fitted with a magnet strip. In these cases, the payment provider usually takes the loss of the victim, which strongly reduces the reputation damage and loss of clients. This is usually not a reason for clients to switch to another method of payment. Also, the direct financial loss for the payment provider is not that big. Therefore, the impact of a single malicious transaction is estimated as low.

Multiple malicious transaction Whereas single malicious transactions are considered to have a low impact, multiple malicious are considered to have a medium impact. It is often harder to prove for victims, that it were in fact malicious payments, causing reputation damage and loss of clients. Also, the financial losses are higher, than with a single payment.

Disable service Disabling the payment service has an impact on the number of transactions made and accounts for financial loss. However, considering the described attack scenarios, this can only happen on a small scale. In only focuses on a small number of users or merchants, which accounts for a low impact.

C.3 Risk

By using the original attack/fraud scenario matrix, the assessment of the attack and fraud scenarios can be combined into a risk matrix. This risk matrix, as shown in Table C.11, defines the risk in three categories: high risk (red), medium risk (yellow) and low risk (green).

There is one attack/fraud combinations that poses a high risk. This involves the theft attack scenario in combination with multiple malicious transactions fraud scenario. According to this risk assessment, this is a risk scenario that poses a serious threat to the NFC payment applications and could seriously damage the NFC payment business. Therefore, this risk requires countermeasures that reduce likelihood or impact of this risk.

Eight attack/fraud combinations are classified as medium risk. Although these risks are not critical, the risks still involve significant large group of clients and can account for moderate losses. The risks also require appropriate countermeasures, only with less priority than the risks classified as high.

Last, three attack/fraud combinations that pose only a low risk for an NFC payment provider. These risks involve only small groups of clients or are not very likely to occur. Countermeasures are not really necessary, when a risk is classified as low.

Table C.11: Risk matrix

	Privacy infringement	Single malicious tran.	Multiple malicious tran.	Disable service
Relay attack	M	M		
Relay attack with malicious app.			M	
Eavesdropping with malicious app.	M			
Eavesdropping	M			
Eavesdropping with terminal exploit	L			
Modify transactions with terminal exploit		L		
Malicious terminal			L	
Denial of Service with zero-day exploit				M
Denial of Service on terminal				M
Theft	M		H	