Prevention of NFC Relay Attack using Time Based Two Factor Authentication

Swanand Sawant
School of Computer Science
George Mason University
Fairfax, Virginia 22030
Email: ssawant@gmu.edu

Sriman Yadagiri
School of Computer Science
George Mason University
Fairfax, Virginia 22030
Email: syadagir@gmu.edu

***Abstract*—Near Field Communication(NFC) has expanded rapidly and is set to be the new big thing in communication. Contact-less payment processing being its biggest beneficiary, there are many other uses of NFC technology like passport identification, security access etc.The growth in NFC technology does not come without any security vulnerabilities. Malicious attackers can carry out many attacks like Eavesdropping,Man in the middle attack and Relay attacks on this type of communication.Among all these types of attacks relay attack is currently not entirely detectable and cannot be predicted in most scenarios.Relay attacks can be carried out on encrypted data and is much more dangerous when compared to other attack.In this paper we give an overview of security issues in NFC communication, describe the relay attack in detail, present our timing based solution to the problem and its implementation, and give results of our evaluation of timing based two-factor authentication protocol for prevention of NFC relay attacks.**

I. Introduction

Near field communication is a type of wireless radio communication designed for a very short range, usually up to 10 cm.It was originally based on RFID (Radio-frequency identification) and operates at 13.56 MHz communicating very small amounts of data, at rates between 106 and 424 kbps. NFC was created to allow the exchange of different information types, such as telephone numbers, pictures, Digital authorizations or perhaps mp3 files between 2 NFC enabled devices like mobile phones, or even between an NFC enabled mobile phone and a compatible RFID chip card NFC is intended to be used as an access key to contents and for services like cashless payment, ticketing and access control. When NFC is used for communication of secure and confidential data, security is always going to be an issue that has to be dealt with priority.Secure transactions, such as financial, the need for security is even higher. Credit card in its natural form is prone to fraud and is rapidly moving to a contact-less medium for conducting business which adds more complexity. The security of NFC is a significant concern and as the use of this technology grows, that concern grows with it.One of the inherent security structures built into NFC technology is the proximity for communication.Because tags are only able to communicate over about 10 cm eavesdropping on the communication can prove to be challenging. It has previously been shown that distance can actually be increased to at least 50 cm. While any attack involving the transfer of secure information is a concern, the relay attack is one of the most critical because there is not a lot of hardware that is required to carry out this attack and the user would not have an idea when the attack takes place. The main security issue is that relay attack circumvents encryption.Contactless systems, as a result of the limited operational range, operate on the implicit assumption that successful communication with a token proves that the token is actually in close proximity of the contactless reader. Therefore, once authentication has been achieved at the application layer, the reader will approve a transaction or perhaps render a service as it believes that the legitimate token is actually in its presence. A relay attack exploits this assumption by placing a proxy-token within the communication range of the reader, which communicates with a proxy reader located in close proximity to the legitimate token. The proxy token is always able to answer with a legitimate response to any reader command since it simply forwards the command to the proxy-reader, which subsequently sends it to the reputable token and returns the valid response from the legitimate token to the proxy token. For the duration of the relay attack the proxy token exhibits the same behaviour as a legitimate token from the reader's perspective. This attack effectively circumvents application layer security mechanisms., an attacker is able to use the authentication protocol by simply relaying a challenge to the real token, that will provide him with the correct response, which can then be relayed back to the reader via the proxy token. It doesn't matter what application layer protocols or perhaps security algorithms are actually used, as the attacker just relays all the application layer data, thereby making sure that both the legitimate reader and also the legitimate token always receive the data they expect.In the last decade, nonetheless, it's been found that these devices are actually susceptible to different attacks. To give an example, the commonly used Digital Signature Transponder (DST) RFID transponder from Texas Instruments has been attacked from a research group from Johns Hopkins Rsa and University Laboratories in 2005. The transponder provided encryption capabilities and was used in millions of automobiles to protect against millions and theft of payment transaction systems which allows to pay contactlessly in supermarkets and restaurants deployed the system in more than 400 restaurants. In order to do the attack 16 FPGAs were used and performed a brute force attack to

reveal the secret key. Some other examples are actually the attack on Mifare Classic and the KeeLoq system that had been used in many remote keyless entry systems including automobile immobilizers and garage doors.

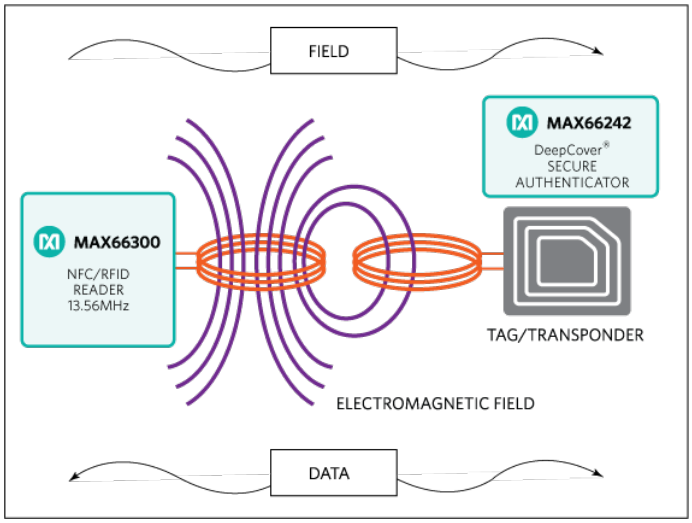

Fig. 1: NFC Technology

II. Existing Work

Several of the more promising work done in relay attack prevention use distance bounding protocols, which set an upper bound to the distances allowed. This is actually attained by requiring a cryptographic challenge response mechanism to measure the proximity to the authentic device. This approach is able to present challenges, such as time for processing such responses, elimination of time extensions for responses, and dealing with noisy channels. The more error resistance is actually incorporated into them the far more open to attacks these bounding protocols can become. A few promising works incorporates a full duplex secret key agreement scheme into the distance bounding. This would hopefully help with computation and power consumption demands as a result of the cost of cryptographic operations. Even when proper encryption strategies are actually used the protocols still remain vulnerable to relay attacks. Without authentication of the reader, an attacker could query the tag in the beginning and replay the info at a later time, which is sometimes called terrorist fraud. Eavesdropping, or perhaps listening in on such conversations, has been identified as a problem. Unless encryption of information is needed in some form this remains an issue. The ECMA even state in one of their documents that users must plan implementations carefully around possible vulnerabilities. As encryption isn't built into NFC, security is actually left up to the application developer, or perhaps the end user of the technology, which isn't a comforting prospect. Even encryption schemes may be ineffective without prior sharing of device authentication. With the relay attack in case we could keep the information from the tag and send it to the proxy to emulate the tag then detection in this manner becomes impossible. There would be no additional timing delays in such a setup. This is also not accurately described as a relay attack however this is much more of a true man-in-the-middle or perhaps just store and replay attack. This particular kind of security vulnerability has already been seen in attacks such as ticket cloning. Even a replay attack for passport control was discussed and is actually vulnerable, but again, this is a replay, not a relay attack. To add to security a user behavioral profile can be created and updated on devices that are mobile. These profi les monitor a users normal activity. Once created they can compare the present activity with the profile and perform a security risk assessment based on that task. This may create additional authentication needs, such as requiring a pin to be entered to activate functions considered unusual for that user. This sort of detection and pro ling, called active authentication, is still in the first stages of development but shows promise. Additional prevention techniques also include gesture recognition and position data. These strategies make use of the device location to verify close proximity to the actual reader for security. Malicious software installed onto devices could also take advantage of the NFC technology, and software-based attacks could be activated, without the victim's knowledge. Google Wallet provides users with a convenient strategy to store credit card data and uses NFC by only carrying the phones of theirs, but software-based attacks are able to create vulnerabilities. Additionally, these contactless payment methods provide no protection against relay attacks. There have been more extreme techniques suggested, such as enclosing readers inside of Faraday cages. This limits all signals from traveling outside of a predesigned physical area. Faraday cages could also be used to enclose tags when not in use, but if used on a phone would limit connectivity or perhaps create a burden on the user. There is actually work suggesting attaching such a device to phones to prevent eavesdropping. Devices like these seem impractical and would most likely not be popular in the market unless they might be included in the phone at the manufacturer. Asking consumers to purchase and attach additional hardware is actually an unreasonable solution and simply addresses eavesdropping, not relay attacks, which would remain a security risk. There have been RFID blocking wallets available for some time, but they are not wide spread. generic, practical, and Recently relay attacks were implemented, just using 2 NFC enabled mobile phones and software apps. It's been found that many EMV-compliant systems still appear to be vulnerable. Previous work has also proven that an extension of the classic relay attack is actually possible. Such an extension can mean an expansion of the distance between the reader device as well as the genuine card. The additional distance varies between forty cm to Fifty cm and also the additional cost is under hundred dolars. Far more precisely, a potential attacker could discreetly access a foreign card from fifty cm far away. This's a fivefold increase in distance as compared to the distance of a genuine ISO 14443 contactless smart card transaction. Additionally, EMV transactions have a typical structure. Hence, if a transaction is actually recorded and the redundant and static data, and

that is the same for every transaction, are actually omitted in the relayed communication, a relay attack transaction can possibly be conducted faster than an actual real transaction. This results in an optimized, time-saving relay attack. All of these Authentication methods improve the security of contactless transactions. Nevertheless, prior research has also observed that the payment terminal itself can be made to fall back to old Cardholder Verification methods (CVM) methods, such as downgrading a full EMV credit card to do a EMV Mag-Stripe transaction. If such an attack vector is actually possible, all the other security measures are actually rendered useless. Another crucial issue concerning EMV, is actually the EMV Personal Identification Number (PIN) verification "wedge" vulnerability. This vulnerability allows an attacker to use stolen cards without knowing the correct PIN. To do so, the attacker copies a card and modifies that counterfeit card in such a way, that the counterfeit card is going to accept some PIN entered, for both offline and online transactions

III. Relay attacks on NFC Communication

The intruder launches a relay episode by putting a malicious master device (reader) close to the tag owners genuine tag and a malicious slave device (emulated tag) close to the genuine reader. The master device as well as the slave device are connected in the form of a wireless connection to pass all activity forth and back between the authentic tag as well as authentic reader. The relay may be attached over any medium including the Bluetooth or Internet, and all visitors in between the genuine tag and genuine reader is transferred straight to the opposite end. Right here we are able to observe the malicious tag(slave device) comes into the field of the authentic reader which requests communication. This particular request is relayed to the malicious master device that forwards it to the genuine tag disguised as a authentic reader. Tag responds and that information is transferred again throughout the link on the authentic reader. The authentic tag as well as authentic reader both believe they're speaking directly to the genuine unit in close proximity. This particular confidence of proximity is bypassed by the relay, and the inherent security feature breaks. Since tags are made to be activated with no user interaction they're especially susceptible to this particular attack. A malicious party may just provide one end of the relay of theirs up to a genuine credit card while it's in a pocket or perhaps a bag and the other slave device near a payment terminal to use that card without owner's knowledge.Security is an essential aspect of the success of NFC technology. The high interoperability of the popular collection of standards must be integrated with appropriate mechanisms to protect data.

Implementation of security mechanisms to a tag requires analysis of costs versus benefits. There are various solutions that imply different economic and computational costs, therefore it is crucial to understand exactly what information has to be protected and which are the main threats.

Newer tags have security functionality built into the chip but are not a part of the NFC tag specification; the principal objectives to pursue for data protection are:

Authenticity Integrity Confidentiality Principal menaces are represented by an attackers ability to intercept and manipulate the data without detection. In both cases, the above principles are violated.

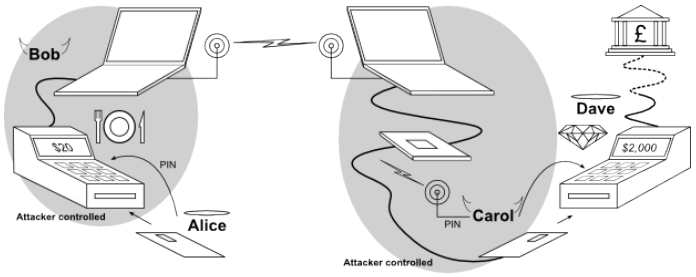

Fig. 2: Relay attack depiction in real world scenario

A. Countering relay attacks

An effective countermeasure solution against the relay attack addressed here will be an implementation that will work within the operating constraints of the existing ECMA and ISO standards. By giving a little change to the software program at the reader or initiator we are able to in essence provide a solution with a software program or perhaps firmware update. This wouldn't create changes to any of the existing tags, which have already been manufactured and distributed. Rather changes will remain limited to the readers, and could produce an an easy-to-implement and affordable solution to our issue. We don't propose a change to any of the standards, to the communication techniques or anything to compromise alternate tag type compatibility, but rather a time based two-factor authentication mechanism to detect and counter possible relay attempts. This remedy has been implemented and tested with Mifare Classic 1K card. We have used a typical NFC tag to emulate our security protocol that asks for a pin if it takes longer than the threshold time used for indicating a possible relay attack.For more security this implementation could allow for any alert mechanism along with pin requirement . If different tag types were to be used additional research into timing for those tags should be evaluated. We want this particular solution to be robust against practical relay attacks and consistent in its warnings and not ask for pin for every genuine transaction as well. To implement this in devices that are Android for instance a quick code update could create the time based two-factor authentication protocol needed to protect against relay attacks and could be pushed out to the devices as any normal software update would be.

IV. Contactless Payment System

From new form factors like watches and phones, to new acceptance locations - like vending machines - contactless payment technology is actually fueling innovation in the payments industry. Issuers and payment brands have introduced a

variety of form factors including fobs, mini cards, stickers, mobile phones and wearables (jewellery, wristwatches and wristbands) - all enabled with contactless payment. As a result, consumers now view contactless as an established payment method for low value general retail transactions, ticketing and transport, and toll or perhaps parking transactions. And contactless is actually happening on the mobile in a huge way. Already available in the US, Canada and Australia, Apple Pay - Apple's mobile payment and digital wallet service has recently launched or will launch in China, Hong Kong, Spain and Singapore. Apple isn't alone. Google, Samsung and Android have implemented contactless payment using Host-Card emulation on devices. Indeed, Samsung is actually taking things even more with plans to extend service beyond payment to transit passes, coupons and membership cards - an approach also possible on today's multi- application contactless payment cards.NFC technology is emerging as a helpful accessory for consumer transactions. NFC is not a payment technology; it's a set of standards that enables proximity based communication between consumer electronic devices like mobile phones, tablets, wearable devices or perhaps personal computers. An NFC enabled mobile device is able to communicate with a POS system that currently accepts contactless payment cards. Contactless payment transactions can be made using NFC enabled devices that are actually provisioned with a mobile payment application and are actually processed the same as contact and contactless EMV chip card transactions.

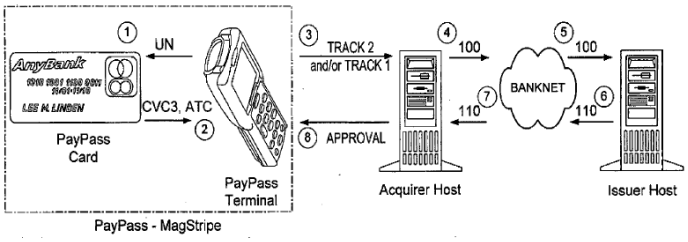

Fig. 3: Credit Card Payment System

A. Devices in Contacless Payments

Almost any device capable of making payments using radio frequency identification (RFID) technology is actually using contactless payment technology. The device doesn't have to be a smartphone though this is by far the most commonly used device for contactless payments. An antenna and NFC chip embedded into the device lets the customer wave their smartphone over a card reader to create a purchase. Our work here is related specifically to NFC enabled credit cards and preventing relay attacks on them. Apple-pay is secure because of it's two-factor authentication as it asks for passcode or fingerprint before sending information through NFC and uses secure chip to store its confidential data to protect from data theft in this case card details.

Security for contactless payments in credit card does not have the feature of two factor authentication for practical reasons. A user wants to use NFC to pay and entering a pin every time is not efficient. Fraud protection laws all apply, and secure channels and encryption are actually used for sending credit card info and PIN numbers. For several purchases or high priced purchases within a very short period of time, the user is actually asked to manually enter the PIN number to ensure theft hasn't occurred. Typically contactless payments are faster because the PIN number or even a signature is not necessary. It also, nonetheless, can result in the customer to spend much more since paying is very quick and easy.

The very first example of contactless payment came in the form of Speedpass in 1997. Mobil gas stations offered contactless payment devices that clipped onto a key ring. The customer waved the device over a labeled square at the gas pump and paid instantly. Today ExxonMobil still offers this service, and other gas stations are actually incorporating contactless payment technologies into their payment choices.

1) Working of NFC Credit Card: In order to understand the working of our solution it is important to understand the working of a contactless payment system. In the normal functioning of contactless transaction through credit card,The first step is that the reader sends an unpredictable number to the credit card which uses this input along with the Automatic transaction counter(ATC) which is stored in the discretionary data field of the card and uses the inbuilt special NFC chip for credit cards that is made use of to compute a dynamic CVV as well as for encryption purposes.The card then sends the newly generated dynamic CVV along with the ATC in encrypted format to the POS. The POS verifies this CVV with the provided ATC and sends this information to the appropriate bank for approval. The latter part is not important as it is verification by the bank system which does not involve the security the security of NFC credit cards. The use of dynamic CVV adds greater security to NFC cards as it acts like one time password for the credit cards. So even if a person is able to store the details of the relayed credit card and emulate it, it cannot be successful in using the cards for malicious purposes due to this feature.

V. Proposed Solution

In the following section we present an approach to detect relay attacks by monitoring the communication delays caused by relay medium used and the processing of the data received at each malicious device and by responding to timing anomalies. The premise for this work is actually that (a) we are able to see just how long it takes to communicate with a tag and (b) If the threshold selected to mark the maximum time it takes for a legit transaction then using that threshold it can predict the likelihood of whether or not a relay attack is actually occurring. To be able to be feasible, this solution requires an accurate analysis of the consistency of timing between different cards. A timing based anomaly detection is only as accurate as the underlying behavior is actually predictable. Once timing is actually shown to be consistent we proceed to demonstrate that the implementation of a practical relay attack detection is actually possible at low cost .To be able to collect tag read latency and relay latency data we implemented

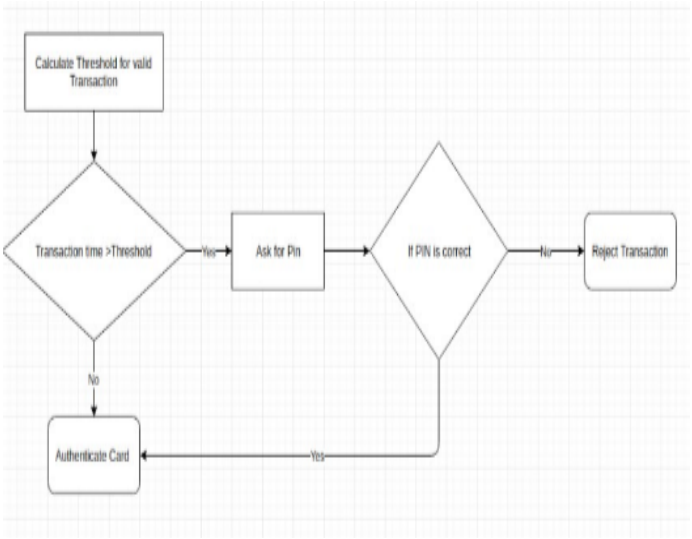

Fig. 4: Proposed Architecture

a relay system. A breakdown of the tag communication steps at a low level can help us with understanding where any timing anomalies occur. We are going to show that the actual time it takes to read a Mifare Classic 1k Card which has certain data bytes and then evaluating delay over the relay system to read the same card. This enables us to distinguish between the actual legit transaction which takes time that is lower than or equal to the selected threshold and a relay attack that takes time that is higher than the selected threshold. Using this demonstrate that it's preventable and detectable. To avoid delays that are caused by anamolies and not by relay attack we add a pin authentication step to let a legit user use the card. In this way the two-factor authentication is required only when the reader predicts that a relay attack is probably happening.In this case if the user is asked for a pin he/she can continue with the transaction as shown in figure 5. If the attacker is asked for the pin, he/she cannot possibly guess the correct pin and the transaction will fail as shown in figure 6.

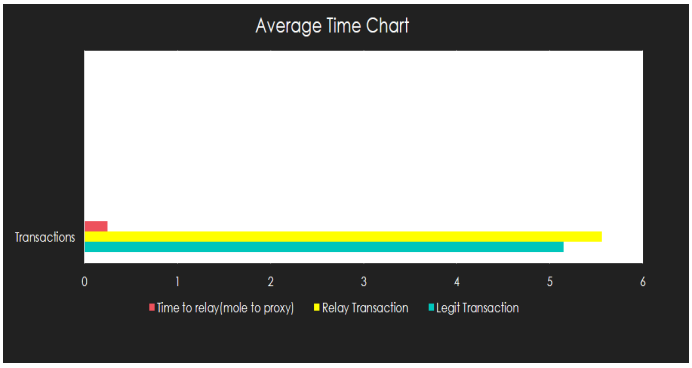

Fig. 5: Average Time Chart

VI. Implementation

1) Emulating relay attacks: To display the working of a relay attack, we have used PN532 NFC reader along with Raspberry Pi connected to a Linux laptop-(ASUS XJ550) running on Ubuntu operating system.We measure the time for every step in relay attack and use SSH to send the Mifare Classic card data over WiFi to the proxy laptop.We also measure the time required for the same card without the relay apparatus.It has been noted that the time difference in actual transaction and the time taken to relay the information from the malicious reader to the malicious slave device is approximately on an average 500 milliseconds.This helps in deciding the threshold to indicate a relay attack has a difference of 500 milliseconds which is quite huge. First we measure the time taken to just read the NFC credit card and then we run a shell script which reads the card and then secure copies the primary database file in this case dump.mfd to the connected linux laptop that plays the role of proxy device in this setup. We also noted that the time to just transfer the dump.mfd file was 18 milliseconds on an average. This represents the time to relay information from mole to proxy device . Time difference greater than 5 milliseconds is significant here as the legitimate transaction mostly showed a +3 or -3 variation. Given that the time difference to relay from mole to proxy being high, the time to relay from proxy to POS is just going to add a few more milliseconds which won't make such a great difference in this architecture. We run the same experiment many times to come to a conclusion that relay attack always took more time than a legit transaction and the time difference was the factor that helped us in detecting a relay attack.Our experiment excludes giving importance to the time difference it took to process the transaction at stack level as it will vary from hardware to hardware. In our case it was around 450 milliseconds on an average.

2) Emulating Time-Based Two Factor Authentication: We have used PN532 Adafruit NFC Shield and Arduino Uno microcontroller version 3 and Arduino IDE and AdaFruit NFC Library.We emulated the relay attack detection and pin authentication to prevent possible relay attack.Time taken to read the MiFare Classic card having the same amount of data for each reading was taken as a threshold for indicati ng relay attacks.According to the proposed solution a relay attack will always take more time than this threshold and hence can be detected.To consider any other network failure that may cause delay in a legit transaction,we've added a pin authentication feature to authenticate the original user in these situations.As the attacker is not aware of the pin he cannot continue with the transaction even after relaying everything.Arduino was used for simulation of the pin authentication feature,to do this we added a small delay that represents the relay attack.This delay causes the reader program to ask for a pin to authenticate the user.

VII. Limitation

This solution is time based which in turn depends on the hardware used in the transaction. Readers have to be faster

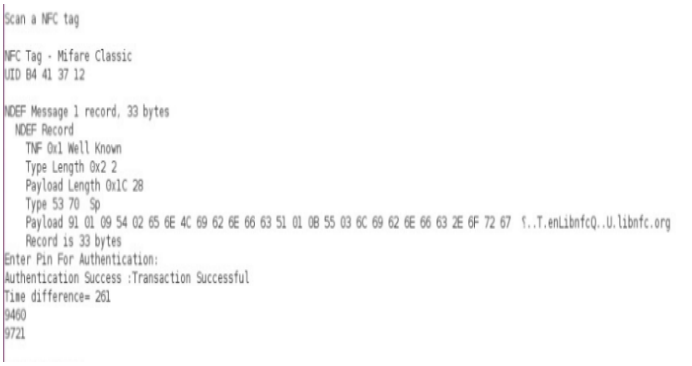

Fig. 6: Pin Authentication for possible relay detection : Success for User

Fig. 7: Pin Authentication for possible relay detection :Failure for Attacker

and process transactions at higher speeds. If a attacker is overclocking its readers to increase their speed and manages to run the relay attack faster than the payment terminal processes a valid transaction, then that relay attack cannot be detected. The only solution here will be to overclock the payment terminals as well which would increase the energy consumption and and decrease the hardware lifetime.

VIII. Conclusion

NFC is now a rapidly growing technology. It surrounds us today and will continue to in the near future. We've all seen the effects of malicious parties taking benefit of security vulnerabilities in technology fields, and NFC is actually no exception to this. The rapid expansion of the technology and lack of solutions to the security vulnerabilities are a concern. Relay attacks are a present and real very threat to this technology, both diffcult to detect and also to stop. We show in this work that by timing the communication of NFC tags a section at a time during data transmissions we are able to detect relay attacks successfully. The accurate timing at the proper level is crucial in this particular endeavor. After detection, many different actions may be taken. One particular action might be a system that first issues warnings, then completely stops communication. By taking a look at the time needed for a data read on a tag we are able to preserve the existing technologies by sticking with the ISO specifications and enhance security at the same time. This won't result in significant changes to the protocols or perhaps the existing devices on the market today. It would just need a software update at the reader side, i.e. POS. In probably the worst case this solution would ask for a pin from an end user for most of the transactions as it probably detects the chance of a relay attack, and also can allow all current NFC devices to work under the established protocols and simply alert them of potential malicious activity. Applying a solution such as this to it within the current ISO specifications without causing expense or changes significantly is important and this particular solution is that criteria. This technology benefits many industries as well as allows for applications in fields which could be very advantageous. We expect that the solutions presented here will add both a level of security as well as value to NFC technology.The effectiveness and ease of the attack means that ticketing, payment (mobile wallets and credit card) and access control application need to be hardened against relay attacks. Currently, virtually no deployed products implement relay resistant mechanisms, with the exception of NXP's new Mifare Plus smart card and that has up to now only seen limited deployment and it's unknown just how many systems that do use Mifare Plus actually take advantage of this particular security service. There are a number of countermeasures which are considered effective against relay attacks, and mobile platforms have much possibilities when compared to conventional smart cards

Acknowledgment

We would like to thank our Professor Parth Pathak for guiding us and encouraging us to learn about NFC and motivated us to find a way to improve the security of NFC.

References

[1] Prevention of Relay Attack Using NFC *Deepa S Pillai1 , S.Sathyalakshmi*, PG Scholar, Department of Computer Science & Engineering Hindustan University, Padur, Chennai, India

[2] Practical Relay Attack on Contactless Transactions by Using NFC Mobile Phones
Lishoy Francis, Gerhard Hancke, Keith Mayes, Konstantinos Markantonakis, Information Security Group, Smart Card Centre Royal Holloway University of London Egham Hill, TW20 0EX, Surrey, United Kingdom

[3] Practical Experiences on NFC Relay Attacks with Android: Virtual Pickpocketing Revisited
anonymous

[4] Security in Near Field Communication (NFC) Strengths and Weaknesses. In: Proceedings of the Workshop on RFID Security and Privacy. (2006) 111
Haselsteiner, E., Breitfu, K

[5] Nfc attacks analysis and survey.In Innova- tive Mobile and Internet Services in Ubiquitous Computing (IMIS)
C. Chen, I. Lin, and C. Yang.

[6] A Relay Prevention Technique For Near Field Communication
ERIC S. WILCOX

[7] Relay attacks of NFC smart cards-Xiqing Chu

[8] Conditional privacy preserving security protocol for NFC applications
Eun.H, Lee.H, Son.J, Kim.S, and Oh.H

[9] Access Without Permission: A Practical RFID Relay Attack
Roman Silberschneider, Thomas Korak, and Michael Hutter

[10] Relaying EMV Contactless Transactions using Off-The-Shelf Android Devices
J. van den Breekel
[11] Relay Cost Bounding for Contactless EMV Payments
T. Chothia, F. D. Garcia, J. de Ruiter, J. van den Breekel, M. Thompson
[12] Picking Virtual Pockets using Relay Attacks on Contactless Smartcard
Z. Kfir, A. Wool