Improve ATM Withdrawal Security and Usability with your Smartphone

Thomas Maqua, René Neff and Matthias Wbbeling
Informatik IV, Universität Bonn, Germany
{clasen, maqua, neff}@cs.uni-bonn.de

Abstract

This paper presents a new conceptual idea to use an automated teller machine (ATM) by using modern technologies and their combined power to enhance usability and security: Google Glass, QR codes and NFC. An extensive concept that combines these technologies is presented alongside its implementation: This includes two android applications - one to be used on a smartphone and one to be used on Google Glass, an online-banking website and a mocked ATM. Within the evaluation, improvements in usability and security in comparison to common alternatives are presented and discussed.

1 Introduction

Today's modern smart phones include more and more features, new wearable technologies are presented, e-commerce is booming: new opportunities arise. Nevertheless, introduction of new technologies still takes its time to change established forms of use. [10] Especially innovation to longstanding established technologies need further motivation to allow for change to happen. Improving security can be such a motivator. Within this paper the an improvement to well-known automated teller machines is presented, it aims to provide a more secure and user-friendly system.

The presented system tries to rise the burden of so called skimming attacks. Skimming is a type of fraud and a special case of a Man-in-the-middle attack. This term summarizes the method of illegal electronically exploration of banking data on automated teller machines (ATMs), one kind of skimming is the capture of data from the magnetic stripe on the back of an ATM card. Another way of skimming is to attach cameras to the ATM to fraudulently capture the PIN numbers. The captured data can be put onto a manipulated or false card and money can be withdraw from the accounts all over the world. Skimming is not only limited to ATMs, instead any kind of cash register or card reader could be victim of a skimming attack. Even if the attackers were not able to gain access to someones PIN it is possible to use them, as for example online retailers do not ask for any PIN or other security codes. [25]

In 2010 the number of skimming attacks reached a record high rate in Germany. In the following years the amount of attacks decreased significantly but nevertheless the stolen money summed up to more than 35 million EURO in 2011 only in Germany. 150.000 cards had to be frozen due to skimming. [4]

To skim or scam different areas on an ATM criminals can extend it with different devices. It is hard for the user to reduce the risk of skimming as ATM look different from bank to bank, why it is difficult to suspect anything unusual with the card reader or the area around. Even if the users hand shields the PIN area during the entering and therefore reduce the risk of shoulder surfing - an overlaid skimmer plate over the existing keypad as a method of PIN capturing and track the digits anyway. Fake card readers are another typical technique to capture information about a customers account data and to read out the information stored on the credit card. [25]
Within our lab in total four ideas were developed. A full description of each idea and it’s implementation to improve security in the context of banking is provided within our lab report. At first the idea of signed QR codes which allows for a cryptographic signature within a QR code. Next is the idea of encrypted QR codes which can be used to conquer shoulder surfing and other attacks. With the NFC TAN Device an alternative to existing TAN generators and distribution is presented. Finally the Mobile Device ATM, as presented in this paper, provides a new user-friendly but more secure idea on ATM cash withdrawal.

This paper will present the necessary background information, the conceptual design and it’s implementation. The products of this implementations are: an online banking website, a mocked ATM (website), an Android smart phone application and a Google Glass application. Combined they provide a viable proof of concept.

2 Related Work

To increase the level of security during the money withdraw process, as well as to optimize the processing performance, several new approaches have been presented and discussed in recent years. In addition to the named aimed-for improvements the including of latest technology was a purpose.

In 2012 the NCR Corporation introduced a way to withdraw money from an ATM in about ten seconds. In this approach the customer is not forced his card out of the wallet as the authentication process is moved to his smartphone and works without it. The customer only needs to be in possession of an Android or iOS smartphone with a camera built in and to launch the application NCR developed. This application administrates all bank accounts of the customer, if he authenticated himself with the correct PIN. The banking card is no longer necessary to log in. In the launched application the user picks the account he wants to withdraw the money from and chooses the amount he wishes to get. Therefore he touches the corresponding dollar figures which represent the different amount possibilities. If those decisions are made the customer is encouraged to use the smartphones built-in camera within the application using the scan button. Displayed on the ATM he wants to withdraw the money from is a QR code. This QR code is used to confirm the transaction and to dispense the money if it is scanned within the application. Consequently, the log in on the application of the customers smartphone could be done outside the bank or in the queue in front of the ATM, whereas the named ten seconds transaction time could be improved. [5]

NCR states that it not necessary to purchase new hardware for the ATM, because only the ATM software needs to be updated to display the mentioned QR codes. The approach is therefore easy to implement and to use on a big scale. Another main argument raised by NCR is the improvement of safety. Because the bank card is not used with the ATM to withdraw the money nor the ATMs keypad is used to type in the PIN, skimming attacks could be prevented. [5]

Only limited access about the security policy of the application is given. Obviously the authentication security level decreased, due to the reduction of the two factor authentication process in the common withdrawal approach using bank card and PIN on the ATM to the only one factor authentication on the users smartphone. The application can be launched on any Android or iOS device and is not tied to a specific smartphone. Therefore the ownership factor is eliminated. An attacker who is in possession of the PIN (and presumably he needs information about the account number additionally) would own all necessary information to have complete access to the money. To gather knowledge about the PIN phishing attacks, over the shoulder surfing or a successful prior skimming attack could be used. It might be assumed
that the common approach and the one presented by the NCR would co-exist at least for a while. As a consequence skimming attacks to customers using the usual withdrawal method would make it easier for the criminals to get access to the money. The would not have to replicate the bank cards any more but use the NCR application to get any money they want only with the gathered information.

Of course the smartphone itself is another vulnerability factor, which, however, cannot be estimated because too less information about the security of the app is given. One possible problem is the easier shoulder surfing on smartphones compared to ATM keypads why PINs could easier be spied. The QR codes itself are another vulnerability to the system. If the displayed code on the ATM is manipulated either by manipulated software or a fake notice to use the printed QR code instead due to technical problems the money would be withdraw on any ATM the attackers have chosen. The user is not able to check if the displayed QR code refers in any way to the ATM he is standing in front of. It is evident that the NCR approach is not as secure and safe as it states and therefore does not improve the common approach satisfactorily. [5]

Basically the same approach was presented by FIS and Wintrust Financial Corporation in late 2013 on the ATM, Debit and Prepaid Forum in Las Vegas [6]. The firm Diebold chose a different but similar approach to address the 'millennium generation' [30]. The first stage of authentication is similar with almost same security issues. One difference is that the user has to scan the QR before he is able to use the application on his phone, what eliminates the improved interaction time as no steps of the process can be done without standing in front of the ATM. The mentioned vulnerability of faked or manipulated QR codes is reduced as the Diebolds withdrawal method adds another step to verify that the user is physical next to the ATM: To confirm the withdrawal a onetime PIN is sent to the users smartphone application which he has to enter into the ATM. [30]

Even with this improvements the overall method is vulnerable to attacks and fraud, as a user only must know the PIN to declare his ownership of the account. Only the QR code vulnerability is better than in NCRs approach.

An older but different idea was installed on several Spanish banks in spring 2011. It is based on contactless ATMs and tries to prevent skimming and fraud by enabling the user to only tap cards or NFC phones to withdraw money instead of inserting the cards into the ATM. The problem with this approach is that skimming is still possible as the NFC readers on the machines should as easily be manipulated or changed than the card readers before. The article states nothing about the security policy of the NFC phone technology or other security measures, but calls the technique 'fully secure'. [22]

3 Background

This section aims to provide an extensive background analysis on various topics which are important for this paper’s approach.

3.1 Nearfield Communication

Near Field Communication (NFC) is a standardized communication technologies that enables radio communication between two devices within a small proximity to each other. Since 2011 NFC is becoming more and more popular within media and public perception [12], by now nearly 300 different smart phones support NFC [14]. Introduced by a summary on the development of NFC, this following part will give a technical overview on the concept of NFC and the NFC
Data Exchange Format. Throughout emerging and already in use use cases are presented and finally some possible security problems are discussed.

Several standardization bodies already passed norms and standards to allow a broad use of NFC: ISO 18092 [13], ECMA 340 [21], ETSI TS 102 190 [9]. Those standards facilitate the radio-frequency identification (RFID) technology, which only describes the broad field of using electromagnetic fields for identification, as printed barcodes do too. Until 2002 RFID lacks a global standard, this changed when the NFC standard was developed by NXP and Sony. [19]

Today ongoing development is done by the NFC Forum a non-profit industry association which assembles over 100 different companies and institutions [17]. This led to a wide spread and still growing acceptance and availability of NFC [29].

NFC’s bidirectional communication allows a maximum communication speed of 424 kbps over a distance of less than four centimeters [16]. Communication can be arranged between self-powered devices (active mode) as well as non self-powered devices (passive mode), such as smartcards [9]. Communication is done at a radio frequency of 13.56 MHz which is part of the international industrial, scientific and medical radio band (ISM), which can be used free of charge [15]. Additional communication speeds are 105 and 212 kbps, all of them are supporting different operating modes: Card Emulation Mode, Peer-to-Peer Mode and Reader / Writer Mode.

In passive mode a so called tag, which can be a sticker carrying a small NFC chip, can be powered by the initiating active NFC device. Necessary power is provided by an RF field generated by the active NFC device which therefore has to be in close proximity to the passive tag. Communication in such a scenario would be handled in the Reader / Writer Mode, a typical use case would be reading contact information from a business card. [16]

Similar to this use case NFC devices can emulate smart cards used for banking or access control, in such a scenario the Card Emulation Mode would be used. This mode allows for contact free interaction with already established infrastructure and still preserving expected security requirements. Card Emulation Mode is also done by interaction between an active and a passive NFC device. [16]

In contrast to those use cases a third way of interaction has been standardized: Peer-to-Peer Mode. It is used when interaction between two active devices, for example two smart phones, is desired. [16]

Message exchange is done in the NFC Data Exchange Format (NDEF) which specifies exactly how data has to be arranged when stored or send by a NFC device or communication with one [2] [18]. A single NFC message, for example the content of a NFC tag, can be divided into single records, called NDEF records, which then hold a header and payload part. [27]

With his paper [25] Collin Mulliner presented some possible attacks on NFC tag reading phones and pointed out some, by now fixed, bug within the handling of read tag data. Overall his research conducted NFC tags may hold malicious data and NFC could be used to create a NFC worm. More on a technical point of view Ernst Haselsteiner and Klemens Breitfu propose that eavesdropping a NFC communication can be done with in one meter range even if one of the two communicating devices is a passive one, for to active components this extends to ten meters [26]. To counter such attacks a secure channel for communication is proposed, as NFC communication is not encrypted or checked for integrity of exchanged data. By using modern NFC (debit) cards this is taken into consideration, as they use especially designed techniques to comply with the EMV specifications while they exchange data [8].
3.2 Quick Response (QR) Codes

A QR Code is a two dimensional, or matrix, barcode designed by Denso for Toyota to mark Products in their stock [20]; it consists of a square white field where smaller black squares placed in a grid. The structure of a QR Code is explained in Section 3.2.1. The coding includes a error correction and is variable in how many percent of faulty reading can be corrected. Also it can include up to 2953 byte of data, depending on its size, which is described in Section 3.2.2. All information about the definition of QR Codes can be found in ISO/IEC 18004:2006 at [3].

3.2.1 Structure

In three corners of the QR Code a Position pattern is placed, to help the reader to identify the correct orientation of the QR Code. For bigger sizes of QR Codes more position pattern will show up in the big data part in the lower right to help recognizing the right orientation. Further there are two lines of alternately black and white dots between the tree position patterns to define positions inside the matrix grid. In the blue parts of the shown QR Code the error correcting level and the data mask pattern is described. The red squares specify the version of the shown QR Code which is explained in Section 3.2.2. At last there are the yellow data parts where the data and the error correction is stored. If the numeric encoding is choosen, it is possible to store up to 7089 characters and each 3 are encoded in 10 bit. In alphanumeric encoding 2 characters are encoded into 11 bit, while it is possible to store 8 bit encoded data with up to 2953 Byte of data. It is even possible to encode kanji, Chinese characters which are used in the Japanese language, with up to 1817 characters in 13 bit per character. A part of this memory space can be used to provide error correction. [3]

3.2.2 Capacity

The data capacity increases by size of the QR Code, as it increases its version number. It starts with version 1 at 21*21 pixel and rises for each version number increase of 1 for 4 pixel. So version 2 is 25*25 pixel and version 40 177*177 pixel which is the maximum size of a QR Code. This maximum size is able to hold the in Section 3.2.1 described number of bytes. So it is possible to save 2953 byte of binary encoded data in a 177*177 pixel QR Code which have to be presented quite big to read it properly, because of limited resolution in display and limited resolution of pictures taken of the QR Code. [3]

3.3 Google Glass

Google Glass is a optical head mounted display, with a resolution of 640x360 pixels, which is build in a pair of glasses with bigger frame as usual. Inside the Frame is a 570mAh battery and a full Android device, which runs up to 8 hours. In contrast to a Android Smartphone it has only W-Lan and Bluetooth and needs a Android device to run applications. An application can read all sensors of the Google Glass and show anything on the Display. Beside a accelerometer and a gyrometer to gather data about head and body movement it has a microphone for voice commands. Furthermore it has a light sensor and a proximity sensor. The 5 MP camera can take videos with 720p and save everything on its 16 GB Flash. Besides controlling the Google Glass with the Phone App or its Touchpad on the right frame it is possible to control it with head movements and eye tracking. Applications for Google Glass are developed in Java with the API which is provided in androids wear editions. Google provides his own development editor ”Android Studio”. [7]
3.4 Authentication Techniques

The FIPS Publication 199 describes three different main goals of security, also called objectives of security [1]. They are relevant for dealing with information over all as well as with information systems. Those three goals are commonly known as the CIA triad of information security, not because of any relation to the central intelligence agency of the United States but due to the initials of the goals: Confidentiality, Integrity and last but not least Availability. [23]

Regarding authentication three basic distinctions are made between three main protection methods. Those methods are bases on knowledge, ownership and biometrical identifiers. Techniques in the class of knowledge are for example passwords or PINs as well as cryptographic keys. The ownership of something can authenticate someone as well, like the use of a SIM-card or smart cards in general.

In case of authentication based on knowledge the user needs to know certain information to authenticate himself towards the system. One possible mechanism is the challenge response procedure (CR). The idea is to use an authenticity check to login, whereby the user first identifies himself via his user name, MAC address or something similar. The instance responds by sending a challenge - e.g. a random number - and the user calculates is response based on it. The instance again checks the calculated response and therefore if the user was in possession of the correct secret. This method is based on a pre-shared secret (the key to calculate the response). Variations of this technique are one time passwords also known as OTPs (RFC 2289) or time based OTPs, the so called TOTPs (RFC 6238).

An example for ownership based authentication is the use of the Elektronischer Personalausweis (nPA) in Germany. With this smartcard passport the user is able to authenticate himself on terminals or online with the contact free RFID chip. The card can be used as digital signature, this possibility is named eID function. The user has to be in possession of the nPA to authenticate himself.

Multifactor authentication combines elements of the previously describes features. One commonly known example for two factor authentication is the log in process on mobile phones: the user owns the SIM card (category ownership) and knows the PIN (category knowledge). Of course the use of an ATM to get money is another example for two factor authentication, as the user needs to be in possession of his bank card and the knowledge about the PIN. [24]

4 Concept Idea

As already motivated the current system for automatic teller machines as well as so far presented improvements lack of security in different ways. Especially fake keypads or card readers are typical examples of attacks, aiming to read the customers PINs or bank card information. The user is forced to trust the security preparations of the ATM alone or, as in the named other approaches, the security levels of his phone.

To avoid this, the NFC technology presented in chapter 3.1 can be used to separate the two factors of authentication. The first development stage of this work was to separate only the authentication process physical from the ATM as given infrastructure and move it to the users mobile phone, similar to NCRs or Diebolds methods. As a result the source of the problem would have been displaced by more or less the same vulnerability. Both information values would have stayed in one place although the place changed. Based on this considerations one must assume that the security level could not have been considerably enhanced.

Under the new approach the two factor authentication process is separated into two different parts. The possession based authentication part, more precisely the users ownership of credit
or bank card, is switched to the mobile phone. The well-known example persona Alice would use her card to authenticate herself. But instead of giving the card away (and under the control of the ATM) she uses the NFC technique as authentication method. This is an improvement to the related work because not the customers PIN has to be known by him but he has to be in possession of the bank card. As the NFC technique works even through the wallet it is not necessary to take it out (and therefore to lose it). The authentication can be done in beforehand, for example in the privacy of the own home or car. Afterwards an application can be used to choose the amount of money the user wants to withdraw.

Only when it takes the next step Alice needs to be in front of the ATM. She scans the displayed QR code of the ATM, to locate and verify it. Subsequently after the withdrawal information is send from the phone to the banking server and then to the ATM. But instead to withdraw the money immediately the ATM asks Alice for her PIN to verify herself.

Consequently the second factor of the two-factor authentication is still on the ATM and separated from the ownership factor. Figure 1 shows the method as a sequence diagram. The added level and separated authentication factors are highlighted.

Common skimming attacks to the ATM would only give the attackers the information about the PIN but not about the corresponding account. The same holds true for the phone application, as only the card information is stored and processed there. A successful attack should therefore aim at both device a disproportionately higher level of effort. Additionally the transaction time can be improved as the first authentication factor as well as the selection of the money amount can be done in beforehand.

### 4.1 Use Case

Alice as sample user owns an NFC enabled debit card as well as a smartphone. Sometimes her friend Bob lends her his Google Glasses. In this scenario Alice wishes to withdraw $100 from
the local ATM near the train station, which is usually strongly attended by other customers.

The use case is illustrated by an activity diagram in figure 2. Alice starts the authentication by holding her NFC enabled debit card near to her smartphone. The NFC triggers the start of the mobile App on her phone and she is automatically logged in. She is now able to check the credit balance on the phone screen. If the amount is high enough to allow the withdrawal of $100 she presses the ATM button on the screen. On the next screen she is asked to choose the amount of money she wants to withdraw. After selecting $100 the app waits for information about the local ATM she wants to withdraw the money from. Therefore she could scan the QR code displayed on the selected ATM near the train station either using the smartphone camera or with Bob’s Google Glasses, if she has got them with her.

Figure 2: Activity diagram illustrating the cash withdrawal process. Actions concerning smartphone and/or Google Glass are highlighted in dark blue, ATM interactions in light blue.

Both scanning possibilities trigger the App to Server communication process to exchange the needed information about customer, wanted amount and identification of the ATM. Depending on the chosen device the ATM identification information is shown on smartphone display or Google Glass. If the correct ATM is selected Alice enters her PIN on the local ATM and thereby confirms the cash withdrawal. After taking her money the transaction ends.

Since it is possible for Alice to do most of the needed authentication and amount choosing on her phone, she is not forced to those steps in front of the teller machine. Instead she is able to prepare the withdrawal in beforehand for example during waiting in the queue. This increases usability as well as it decreases time consumption.

5 Proof of Concept

Our concept is realized by a system consisting of a smartphone application and an ATM which is mocked by a website displayed on a touchscreen computer display. A background infrastructure
for managing requests and providing said website is setup too.

We propose the use of cryptographic signature within QR-Codes to enable Alice to authenticate her chosen ATM. This technique would include two components: One to generate such a QR-Code, and one to decode it on a mobile device.

A signed QR-Codes has to include a timestamp and a unique identifier $i_1$ as a secure way to enable an observer to identify the device on whose display the QR-Code is shown. These QR-Codes would be valid for a defined amount of time $t$, and have to be regenerated after this time. An observer is able to verify whether the cryptographic signature is correct and whether the timestamp is not older than $t$, which proves to him that the device is correctly identified by $i_1$, if it has not been attacked within $t$. This holds only if an attacker is not able to compromise both this device and another device with the identifier $i_2$, causing the device $i_1$ to display the QR-Codes of $i_2$. While it is possible to circumvent this protection, it is still more secure than most other forms of identification, which can be for instance painted over or forged in a similar way. Thus, it is a form of authentication of the device to the observer.

The implemented banking server is capable of generating a QR-Code with an embedded signature. It is displayed via the mocked ATM website and refreshed every 30 seconds, a status bar indicates the refresh cycle.

The implemented Android application is able to decode a scanned QR-Code and to check the embedded signature against the bank’s public certificate. Furthermore, a summary of the desired transaction and a detailed description of Alice request to withdraw money is transferred via a TLS secured connection and her identity is authenticated by her NFC-enabled banking card. A mocked banking card containing a NFC tag holding the user’s authentication token was used.

To acknowledge his desired withdraw request the user has to provide only his PIN via the ATM’s interface. A mocked ATM withdraw is presented to show the transaction’s success if the correct PIN was provided.

The final system can be seen in figure 3.

6 Evaluation

This section regards the evaluation of the developed concept and system. Therefore it discusses the security improvements of the presented project in a first step. In a second section the usability aspects of the present work are discussed briefly.

Afterwards, the section concludes with a summary. In there a conclusion is done, reflecting the work done in the original project as well as the concept shown in this paper.

6.1 Security

The main driving forces of this work was the improvement of security in the ATM withdrawals. This was to be achieved using an NFC authentication process and signed QR codes. The former handles the communication between the communication between the banking card and the Android application, the latter is displayed on the local ATM and enables the user to check the apparent authenticity of the ATM as well as to verify himself to be standing in front of it. As the QR code displayed on the ATM changes after a certain length of time he can be sure the QR code is currently valid. This fact is new compared to the related work presented in section 2.

Another major new development given by the QR code displayed on the ATM which is signed by the bank itself is that the user is able to check the authenticity of the ATM. In the
current applications this was not possible to him; he had to trust the integrity blindly. Thereby another level of security is given to the user and enables him to detect manipulation. As the QR code displayed on the ATM changes from time to time and it is digital and not printed the manipulation of the code more complicated for attackers in addition. In preliminary research it was not possible to find any comparable approach of signed QR codes to authenticate ATMs, making our approach in this application unique for the time being.

The last enhancement allowed by this works approach is the two-factor authentication process not only supported by two different factors but also by two different devices. In the current applications of cash withdrawal the user is forced to trust the ATM. He is charged to hand over the ATM not only his possession factor (the debit card) but also the PIN. If the ATM is manipulated both needed authentication factors are given to the attackers in one step. Our approach resolves those security gap by separating the factors onto different devices.

The information stored on the debit card is only given to his smart phone, more precise to the app. The needed information for the ATM is only communicated directly to the server, without use of the input possibilities of the ATM. However, the PIN is the only information presented directly to the ATM. It is quite more complicated for attackers to obtain both information in one single attack. Therefore the number of skimming attacks could be reduced or made less profitable.

6.2 Usability

As mentioned in section 4.1 the cash withdrawal with our approach accelerates the process in some ways. By preparing the withdrawal on the smart phone in beforehand and without the
need of physical proximity to the ATM, the main part of withdrawing money can be lead away from the bank and ATM. Queues in front of ATM can be avoided as users only have to scan the QR code displayed on the ATM and take their money in the bank branch.

Another increase of usability is enabled due to the fact, that the user is not forced to take his debit card out of the purse. The authentication based on NFC is possible even if the wallet is brought only near by the smart phone. Therefore the risk of loss or oblivion of debit cards is minimized.

6.3 Summary and Conclusion

The main goal of the original work was to use three arising technologies to create a more secure and user-friendly system. The field of banking was chosen as application area and to show the advantages of the technologies. The technologies NFC, encrypted and signed QR codes and Google Glass have been integrated in a first step into four ideas, one of which have been presented in this paper: A new method to make ATM withdrawals safer.

This project relied on existing and established security libraries when implementing the encryption and signing of QR codes. Those libraries and techniques are generally accepted guidelines. Due to this the presented system is reliable and easy to extend.

The most challenging aspects of this work was to deal with android development. The construction and operation of android apps was quite complex and needed more time than calculated before. The integration of external libraries under android was a tremendous challenge leading to different decisions concerning the selection of the development tool. Under Android SDK it was almost impossible to integrate the libraries in the intended way, so Eclipse was used as the development environment instead.

To conclude this work was challenging but nevertheless resulting in a working system, proving the developed concepts. It could be shown that the integration of Google Glass, NFC and encrypted as well as signed QR codes is possible to increase the security level of online banking. Based on this foundation derivative work is possible and preferable.

References


