

Securing Plastic Money Using an RFID Based Protocol Stack

Abstract. Since 2006, there have been three major systems that have been implemented in an attempt to reduce the threat of credit card fraud - Chip and PIN (United Kingdom), Chip Authentication Program - CAP (European Union), and RFID enabled credit cards (United States of America). In spite of a big effort by the EMV¹, there has been little evidence to demonstrate the success of these schemes in stopping fraudsters, scammers, and identity thieves. This may be attributed to combinations of poor usability, lack of trusted interfaces, the absence of smart-card cryptography that takes full advantage of the available computation resources, and inadequate authentication protocols. In this paper, we explain the shortcomings and vulnerabilities of each of these systems, and then explain requirements of a secure and usable cashless payment system. We also describe a new RFID based protocol stack - SECAPS (Secure Cashless Payment System), which obviates many of the attacks on the current schemes by using the newly available computation resources on modern RFID Tags.

1 Introduction

Credit and debit cards have long been accepted as a convenient alternative to carrying wads of cash in a wallet. However, while it has been accepted by the public, credit card fraud has been a rather expensive problem that has plagued societies around the world for more than a decade. Statistics from the United Kingdom alone indicate losses of over £609 million in 2008 due to card fraud [1]. There has been some significant effort over the last few years by the EMV to quell this problem, such as introducing the Chip and PIN in the United Kingdom in 2006 [2], RFID enabled credit cards in the United States in 2006 [3], and the Chip Authentication Program in the European Union in 2007 [4].

1.1 Types of Credit Card Fraud

We first describe the major types of credit card fraud that occur around the world. In the coming section we will describe how the information is obtained through a combination of privacy and social engineering attacks to carry out these crimes.

1. Card not Present

This kind of fraud is committed by misusing a victims credit card information over channels such as the Internet, fax, mail order, or telephone. This kind of fraud is most commonly used by fraudsters since it saves them the trouble of cloning a credit card with the information they have. It is also harder to identify and trace than other types of fraud. Further, it is possible that a victim will be unaware of this fraud until they receive their monthly account statements, giving the fraudsters ample time to make an escape.

2. Counterfeit Cards

Once a fraudster obtains information such as a card number, card expiration date, and card holders name, it is possible to replicate the magnetic strip and clone a credit card. This required information can be obtained by having the victim swipe his card at tampered terminals, reading information through RFID readers, and sometimes they are even available for sale online! [5]

3. Lost and Stolen Cards, Mail non-receipt

This kind of fraud occurs when the victim's lost or stolen credit cards are misused by a fraudster. A growing number of cases of credit cards being intercepted by fraudsters while making their way from the bank to the user have been reported over the last few years.

The above three types of fraud accounted for over 80% of the losses through credit card fraud in 2008.

Losses through credit card fraud have increased by over 42% since 2006. This indicates that there is obviously a need to re-evaluate the payment systems currently deployed around the world.

1.2 Related Work

There has been little work in the area of securing credit cards. This may be attributed to the fact that a large number of documents and technical reports that describe the functioning of these payment systems and the protocols behind them have not been made available to the public and to investigating researchers.

Reverse engineering was carried out first by Heydt-Benjamin *et al.* in [6] for the RFID enabled credit cards which were introduced in the United States in an effort to understand the underlying protocols. Later in 2006, a detailed study on the security of the EMV Chip and PIN system and the secure messaging system used by its back end API was performed by Adida *et al.* in [7] and [8]. An attempt at understanding the Chip Authentication Program (CAP) protocols through reverse engineering was made by Drimer, Murdoch, and Anderson in [9].

1.3 Organization

In section 2, we give a brief description of the attacks that have been carried out on the three credit card protocols - Chip and PIN, RFID, and CAP. In section 3, we describe the requirements of a new generation of secure credit cards. In section 4, we present a protocol stack (SECAPS) which secures credit card transactions in public environments and even in Card not Present transactions. Finally in section 5, we make our conclusions.

2 Motivation: Weaknesses of Popular Payment Card Schemes

2.1 Chip and PIN Credit Cards

The Chip and PIN was made mandatory in the United Kingdom in February 2006. This was done by banks to completely remove their liability in the case of Point-of-Sale (POS) frauds. Compatible credit cards are actually smart cards (i.e. they contain an embedded chip capable of complex computations) which follow the ISO/IEC 7816 [10] specifications. Its operation procedure is described below.

1. The card is inserted at a POS terminal by the customer.
2. The customer inputs a PIN into the POS terminal.
3. A cryptographic matching algorithm is executed to verify the correctness of the entered PIN.
4. The transaction information- destination account, transaction time, and transaction amount, along with the customers card information such as card number, expiry date, and card holders name are passed along to the back end processing system where the charges are made to the card holders bank account.

Vulnerabilities of the Chip and PIN Card The Chip and PIN Reader shares many properties with the Readers of Automated Teller Machines, including their vulnerabilities. In fact, the Chip and PIN Reader may be seen as a portable Automated Teller Machine which is used to process merchant-customer transactions rather than dispense currency. The Chip and PIN payment system has failed to deal with some very basic attacks such as - Observation attacks with hacked and counterfeit terminals and terminal interception attacks. Other more complex attacks that may be carried out are - relay attacks and phishing attacks [7].

- **Observation Attacks:** In this type of attack, the merchant has his own magnetic strip reader inserted in the card reader. When the customer places his card in the reader for authentication, the reader makes one copy of the data on the strip for the merchant and sends the other to the back end processing system (as a legitimate reader would do). The information on the magnetic strip includes the cardholders name, card number, and card expiration date. Since addresses of individuals are easily available and is considered to be public information, the merchant now has enough information about the card and its owner to carry out Card not Present fraud. Using the information obtained from the magnetic strip, the merchant may also clone the customers card. A hidden camera or a key logger may be used with the PIN pad to record the PIN entered by the customer. Using the cloned card and this PIN, the merchant may access the customers account through any ATM. Such attacks are not complicated and have been on the rise recently [11–14]. A recent survey [15] also shows that guessing attacks are a surprisingly effective way to compromise the security of Chip and PIN cards whose magnetic strip data is known.
- **Terminal Interception Attacks:** Here, a small hardware device known as an “interceptor” is placed in the Chip and PIN reader. It intercepts the magnetic data stream that is sent to the back end for further processing. The PIN is recorded using a key-logger which is part of the interceptor. This is easily possible since the magnetic strip information is sent to the transaction server in plain text with no encryption. The Reader may or may not process the transaction. The first suspected instance of such an attack was recorded at various ATM’s in Las Vegas in July 2009 [16].
- **Relay Attacks:** In a relay attack, the Chip and PIN Reader is a device which transmits messages from the card to another device in real-time. The receiving device then sends these exact same signals to a collaborator which now appears to any other Reader to be the card of the victim. These attacks are more difficult and expensive to perform, but are still a major threat to the security of all credit cards. Recently there have been several proposed methods which detect such attacks in Chip and PIN cards [17, 18]. Unfortunately, none of these have been implemented by the EMV.
- **Phishing Attacks:** A phishing attack is a social engineering attack that involves attacking the weakest link in any security protocol - the human user. This type of attack was first suggested by Adida *et al.* in [7]. In such an attack, the victim is sent seemingly “official” instructions by a scammer (who has a fake bank account) to replace their existing card, these will include filling out a form that asks for their name, address, current card number, expiration date, etc. The attackers then send the victim a compromised card which has a chip that transmits information regarding the victims PIN to the attackers by having this information encoded in some field such as the transaction certificate or transaction number. The attackers now can obtain an account statement of the victim which will contain the users PIN encoded in the transaction number. This will be easy to do since the attackers are aware of the victims online banking credentials (since the card was originally theirs). The attackers can now create a card exactly identical to the original card of the victim. With their knowledge of the victims PIN, this card can be misused at any ATM or POS. These attacks are more difficult to execute than others, making it less attractive to fraudsters. It is unlikely that we will see such attacks executed in the future since there are easier ways to compromise card security.



Fig. 1. A PIN and Chip based relay attack

2.2 RFID Enabled Credit Cards

RFID enabled credit cards were introduced in the United States by American Express (ExpressPay), Mastercard (paypass), and Visa (payWave) in 2006. Since then there has been a lot of opposition to the technology by the public and the press. While some of the concerns regarding the implementation of RFID have been well founded, a large amount of it has been based on near-facts and half truths. RFID enabled cards (also referred to as Tags in this paper) contain a processor with the ability to perform some simple calculations. This processor is connected to an antenna that allows it to communicate wirelessly with any device that supplies power to it in the form of RF signals (at 13.56 MHz). These cards do not contain any embedded sources of power, they are called passive RFID Tags since they obtain all their power only from other devices called RFID Readers (also referred to as Terminals in this paper). They follow the ISO/IEC 14443 [19] specifications for communication. We describe briefly the operation procedure of these cards below.

1. The customer holds his card within a distance of 10-15 centimeters from the POS RFID Reader.
2. The Tag in the card is activated by the RF signals sent by the Reader.
3. The transaction is authorized without a PIN for transactions under \$25. Otherwise, the customer needs to enter a PIN at the POS terminal.
4. Once the PIN is entered, a cryptographic matching algorithm verifies the correctness of the entered PIN.
5. The card sends via an RF signal, the information that would normally be obtained from the magnetic strip of the card - *i.e.* card number, expiry date, and card holders name. This information is sent in plain text for some banks, other banks use pseudonyms, transaction counters, or cryptography to conceal some of this very sensitive information.
6. The RFID Reader transfers this information to the back end processing system along with other transaction related information such as destination account, transaction time, and transaction amount. The charges are made and the amount is transferred to the merchant from the card holders account.

Note: The cryptographic protocols used in RFID enabled cards are proprietary have not been made available to the public and are therefore not explained in full detail here. Information has only been obtained through reverse engineering in [6].

Vulnerabilities of RFID Enabled Credit Cards The attacks performed on RFID enabled credit cards are similar to the attacks performed on other RFID enabled devices which are used for identification such as ePassports, eIDs, etc. These include - Skimming attacks, eavesdropping attacks, user tracking, replay attacks, and relay attacks. An attack that is specific to RFID enabled credit cards is the cross contamination attack. The RF signal from the Tag (the card) to Reader is sent in the same form as magnetic strip data in plain text for most RFID enabled credit cards, making them vulnerable to the mentioned attacks. RFID cards which use strong cryptographic techniques for the data transmission between the Tag and Reader are not vulnerable to all the mentioned attacks - except the relay attack (this is implemented only by 2 of the 3 major RFID enabled credit card distributors).

- **Skimming Attacks:** Since there is no concept of mutual authentication in RFID enabled credit cards, it is possible for anyone with an HF RFID Reader to “talk” to the RFID Tag on the credit card. This means it is possible for any Reader to get magnetic strip data (name, card number, and card expiration date) from a credit card Tag. This information can be used to create a duplicate swipe-only card. It is possible to prevent such attacks by using Faraday cages which prevent cards from talking to Readers when they are enclosed within them.
- **Eavesdropping Attacks:** Eavesdropping attacks are carried out by having a Reader record the data stream between the Tag on the card and another (legitimate) Reader. The attacker now has magnetic strip data from the card enabling him to create a duplicate swipe-only card. This attack cannot be stopped by using a Faraday cage since the card is taking part in a legitimate conversation while the attacker records its data stream.



Fig. 2. An eavesdropping attack on an RFID enabled Credit Card

- **User and Transaction Tracking:** Since RFID Tags are activated by any Reader in range, and Tags emit fixed identifiers on activation, they can be used to track the movements of an individual without their consent. Since many RFID enabled credit cards also maintain a transaction counter, it is also possible to follow an individual and use a Reader to figure out how many transactions they completed since the last reading.

- **Replay Attacks:** Some RFID enabled credit cards do not make use of time stamps or transaction numbers, this means there is no way for the processing system to verify the validity of a transaction. In these cases, it is possible for an attacker to capture a data stream from a legitimate transaction between the Tag and Reader, and then replay it as many times as they wish to. The replayed transactions are always processed successfully since there is no way for them to be detected.



Fig. 3. A relay attack on an RFID enabled Credit Card

- **Relay Attacks:** Relay attacks for RFID enabled cards are easier to carry out than the relay attacks on Chip and PIN cards because of their wireless communication capabilities. These attacks can be executed successfully even on cards that have strong cryptographic protocols. An adversary uses a Reader to communicate with a victims RFID enabled credit card, and relays the data stream to the his associate who possesses a credit card emulator which communicates with a nearby POS terminal for a transaction. The emulator then relays the POS' data stream back to the victims card through the associates malicious Reader. The victims card then believes this is a legitimate transaction and carries on the conversation. The transaction is authorized by the POS and is charged to the victim. These attacks have been carried out successfully in the past [6, 20, 21].
- **Counterfeit and Hacked Terminal Attacks:** These attacks require legitimate RFID Readers at POS terminals to be replaced with counterfeit or hacked Readers. These hacked Readers record all RFID communication received by all interacting cards and also log key strokes of the PIN pad along with a time stamp. The crooks at the end of the day can look up the data stored in the terminal and note down the victims name, card number, and card expiration date. Since a PIN is required for all transactions over \$25, they can also look up the keystroke log to note down the PIN if it was entered by the victim. Using the magnetic strip data and the PIN, it would be possible to obtain a swipe only card and use it at an ATM to clear out the victims account. These attacks would be very easy to carry out with co-operation from the merchant.
- **Cross-Contamination Attacks:** A cross-contamination attack combines any of the above mentioned attacks with a public information search to locate the victims address. Once the information from the above attacks is combined with the victims address, it can be used to commit Card not Present fraud. Since an individuals billing address is usually their residential address, which is public information, it is very easy to carry out these types of attacks.

2.3 Chip Authentication Program Enabled Cards

The Chip Authentication Program (CAP) was introduced by Mastercard primarily to curb the amount of Card not Present fraud. It introduces the concept of multi-level user authentication. The CAP protocol has three authentication modes: Identify, Challenge-Response, and Signature mode. Of these, most of today's implementations make use only of Identify and Challenge-Response mode of operation. CAP cards are smart cards that have the ability to compute basic cryptographic operations. Each card has a secret key stored in its restricted access memory. CAP cards also have a 16 bit transaction counter that is incremented on every reading. The operation procedure for Card not Present transactions using these cards is described below.

1. The customer inserts his CAP card into the CAP Reader and enters his PIN.
2. The CAP Reader runs a cryptographic matching algorithm which verifies the correctness of the entered PIN.
3. He is then asked by the POS terminal (different from the CAP Reader, usually his PC), to select a specific authentication mode on the CAP Reader.
4. The CAP Reader now generates a one-time password that is to be entered at the POS terminal. This password is generated differently according to the operation mode selected.
In the Identify mode, the CAP Reader generates this password based on the transaction number.
In the Challenge-Response mode, the POS Terminal asks the user to enter challenge into the CAP Reader, the CAP then supplies the password which is a combination of the transaction number and the challenge.
In the Signature mode, the POS Terminal asks the user to enter their account numbers into the CAP Reader, the CAP Reader then generates a password based on this number and the transaction number.
5. Since the POS terminal is online, the entered password is sent to the card issuer who has all the data that was used to generate the password. Once verification is complete, the transaction is authorized.

Note: The cryptographic protocols used in CAP cards are proprietary have not been made available to the public and are therefore not explained in full detail here. Information has only been obtained through reverse engineering in [9].

Vulnerabilities of CAP Cards The CAP card is by far the most secure payment card and has been difficult to attack by traditional means. However, because of their inconvenience, CAP card Readers are easy to compromise (users refuse to carry them around, and often borrow Readers from unknown people when a transaction is to be made).

- **Hacked CAP Attack:** For such attacks a CAP Reader is compromised by an adversary in such a way that it has the ability to record and store magnetic strip information from the card and also has an inbuilt keystroke logger. The keystroke logger records the PIN entered by the user during the user authentication process, while the CAP Reader records the Card Number, Card Expiry Date, and Name of Account Holder. This information can later be sent to the adversary via Bluetooth, IR, or RFID. The adversary can use this data to create a magnetic strip based swipe card. While it will not be useful to participate in Card not Present frauds, it can be used to access the victims primary account via an ATM, or for transactions at foreign destinations where CAP is not prevalent.
- **Malware Based Relay Attacks:** Relay attacks can still be carried out in the Identify and Signature operation modes with the help of a compromised terminal, and a compromised online CAP Reader. Here, the adversaries modify the CAP Reader to give it online connectivity. The

terminal used for the victim's transaction is compromised such that it refuses to make any online transactions. The victim attempting a legitimate transaction in the Identify or Signature mode, enters his card into the compromised CAP Reader which uses its wireless connectivity to relay card information to the adversary. The terminal now asks the user to enter one of the three operation modes and give necessary user input. The user complies, and the CAP Reader sends the entered mode, the user input, and the generated one-time password to the adversary. The terminal at this point will not complete the transaction. The adversary has all the needed information to perform a Card not Present transaction in the same mode as the victim.

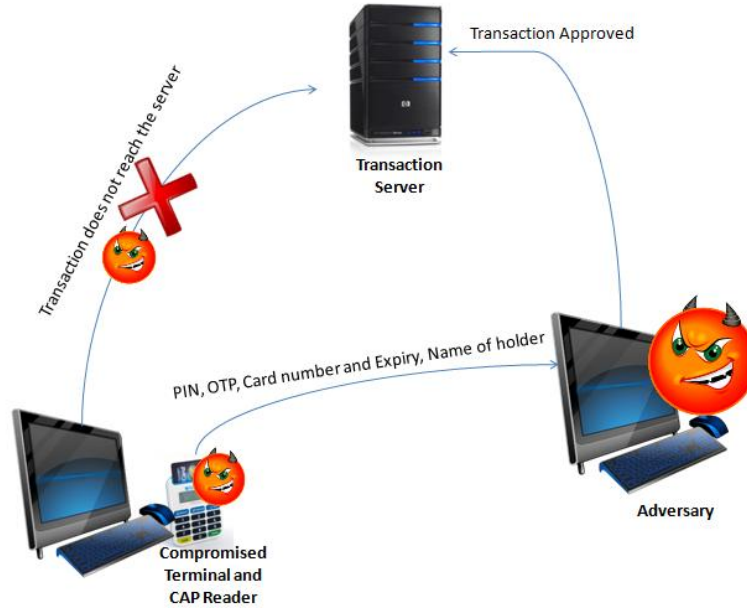


Fig. 4. Illustration of a typical malware based relay attack

3 Requirements of a Secure Cashless Payment System

3.1 Strong User Authentication

One of the weakest links in any security system is the point of user authentication. There have been many studies in the past that have shown that using passwords and PINs for user authentication is not ideal [15, 22–25]. Leaving the choice of a PIN to the user makes it vulnerable to social engineering and guessing attacks, giving the user a PIN that they have no control over leads to poor customer satisfaction, and eventually leads to the user making the system more insecure by having them store PINs and passwords in their wallets, drawers, cars, desktops, etc. PIN and password based user authentication also leaves users susceptible to keystroke loggers (hardware or software based) which are easy to plant [24, 26], and shoulder-surfing attacks. Many solutions have addressed this problem, but most of them fail to provide practical, yet secure user-authentication. There is need for an authentication system that either requires dynamic user input, or hard to forge static user input. PINs and passwords do not fall in either of these categories.

3.2 Mutual Authentication

Most of the attacks described for the Chip and PIN cards, RFID enabled cards, and CAP enabled cards are feasible because of an adversary's ability to compromise a card Reader. The implementation

of a protocol to authenticate the Reader in addition to the existing Card authentication protocol would obviate these attacks. Unfortunately, for mutual authentication to be feasible, there is need for a comprehensive change in the hardware manufacture and distribution system of current schemes, since the implementation of a Public Key Infrastructure is required. There is also a need for a central body that issues and revokes certificates. This according to us is feasible since even the cost of such a major shake-up would be less than the amount of time and revenue lost in credit-card fraud (through legal fees, settlements, man-hours, and other losses) in one year for a large number of users. The existence of a body such as the EMV makes it easier to set up certificate distribution authorities and a public key directory (as the International Civil Aviation Organization did for ePassports [27]). The only problem with building such a system is the time to roll-out and the need to redistribute hardware equipment to customers and merchants.

3.3 Strong Cryptography

There is a need for better cryptography in credit card systems. For a while now there have been implementations of commonly used cryptographic protocols such as the ones described in [28, 29] which work well even on the limited computation capabilities of contactless smart cards. It is important that all sensitive information on the cards is encrypted, or at least inaccessible to unauthorized Readers to ensure that cards cannot be cloned. The data should also be signed by the card issuer to ensure that no data on the card has been modified. A lot of attacks described in previous sections are feasible because of the systems inability to detect and defend against cloned cards. Defense against such attacks can be achieved by using a challenge-response based card authentication protocol in a Public Key environment.

3.4 Trusted Devices and User Interfaces

There is a need for a trusted interface for user input during the authentication process. If it is possible to ensure that the point of data entry or user input is tamper-resistant or tamper-evident, then users can easily detect compromised Readers and authenticators from good ones. The presence of a device containing such an interface coupled with user awareness will obviate many attacks described in the previous section. The CAP enabled card payment system does exactly this by providing each user with a CAP card Reader. However, it is still a failure because of its poor portability and usability. An alternate technique to achieve this is through a Terminal authentication protocol. Another requirement is that cards are readable or accessible to readers only when the user expresses a desire or need to take part in a transaction.

3.5 Improved Usability

Security and usability are essential in any payment system and one usually comes at the expense of the other. For devices such as payment cards which are mostly used by the average Joe, it is important that users feel comfortable (and not challenged) while using them. CAP enabled cards while achieving good security, fail to make users comfortable because of their varied operating instructions and poor portability. In fact there are several blogs and even facebook groups dedicated to abolishing CAP enabled cards! [30–32]

4 SECAPS: A Secure Cashless Payment System

Having discussed the attacks on current cashless payment systems and the requirements of a secure yet usable cashless payment system, we are now in a position to describe a protocol stack which satisfies the above stated requirements and obviates all the previously described attacks. Our system makes use

of an RFID based Public Key Infrastructure which implements mutual authentication. In this section we will describe the Public Key Infrastructure, User Authentication, POS Terminal Authentication, Tag Authentication, and the general operation procedure of SECAPS enabled cards. Before we do this, we will describe the resources available to us on typical high end RFID Tags.

4.1 Available Computation Resources in Modern RFID Tags

Recently there has been a noticeable increase in the available computational power and resources on RFID Tags. RFID Tags such as Infineons' (EMV certified) SLE66/78CLxxx PE family [33] have been built specifically for applications such as contactless payment and they provide us with the following resources which enable us to think about solving our security issues with public key cryptography.

- Memory constraints: User ROM - 288KB, EEPROM - 144KB, RAM - 8KB.
- Symmetric cryptography: 3DES and AES encryption with up to 256 bit keys.
- Asymmetric cryptography (encryption and signature verification): RSA operations with keys up to 4096 bits, ECC operations with keys up to 521 bits.
- Random number generation: True Random Number Generation (AIS 31, FIPS 140 [34]).

4.2 The Public Key Infrastructure

A Public Key Infrastructure is required to aid the process of public key distribution and authentication. A Certificate Authority (CA) issues signed certificates to every institution in the chain. The PKI is usually hierarchical in nature in the case of a large number of institutions. The key elements in our PKI are the Division I Certificate Authorities(D1CA), Document Verifiers (Banks), and POS Terminals (RFID Readers).

The highest level body in each region is appointed by a universal body such as the EMV and it acts as the D1CA (Note that there will be a very small number of these). The D1CA generates and stores a public-private key pair (KPU1CA, KPR1CA). The private key of the D1CA (KPR1CA) is used to sign each Document Verifier (DV) certificate (from its own and from all other D1CA's). There are usually many Document Verifiers in each region. Each of these document verifiers generates and stores a public-private key pair (KPU1DV, KPR1DV). The private key (KPR1DV) of the DV is used to sign each POS Terminal certificate in its region and also the Security Data Element (SDE) and public key of every RFID enabled Credit Card that it issues. The SDE is nothing but the computed hash on all the information stored on the card (including its public key).

Note: User owned POS Terminals (such as PC's and cell-phones) that may be used for Card not Present transactions can be fitted with DV certified RFID Readers for very nominal costs (estimated to be under \$60). The size of these Readers does not reduce the usability and portability of these devices [35].

Dealing with Reader Revocation A check for the validity and revocation status of received certificates is required before proceeding to the authentication step. This is one problem that has plagued other RFID PKIs for a long time. We would like to state that describing a solution to the problem of reader revocation and is beyond the scope of this paper but it has been successfully and efficiently dealt with in our previous work [36].

4.3 POS Terminal Authentication

We require a POS Terminal to first authenticate itself to the Tag on the RFID enabled Credit Card to ensure that Denial of Service attacks by malicious Readers are not feasible. The Reader is authenticated to the Tag using the following protocol before access to data stored on the card is granted.

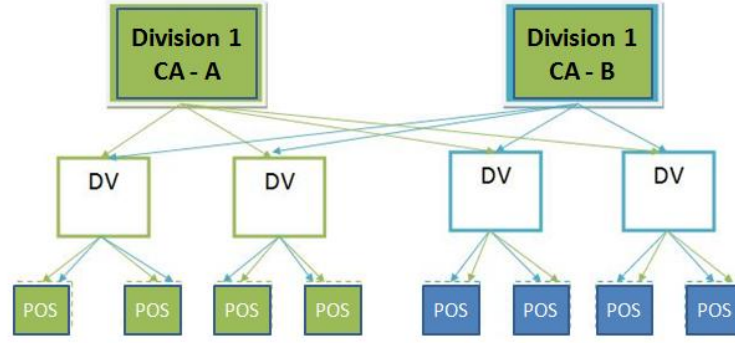


Fig. 5. Public Key Infrastructure for RFID enabled Credit Cards

Here, $CERT(DV)$ refers to the certificate issued by the D1CA to the issuing DV and $CERT(POS)$ refers to the certificate issued by the DV to the POS Terminal. These are sent to the Tag for verification (the Tags have prior knowledge of the public key of the D1CAs) and to extract the public key of the POS (RPuK). Now, the POS Terminal and Tag generate a ephemeral Diffie-Hellman key pair (public keys - R' , R'' ; private keys - x , y) using the typical protocol [37] (with parameters g , q). The Terminal then computes a fingerprint (hash) of R' (its ephemeral public key). This step ties the Terminal Authentication protocol with the following Tag Authentication protocol. The Tag then generates a 128 bit challenge (r) and sends it to the Terminal. The Terminal signs ($r || h(R')$) using its private key RPrK. The Tag can verify the correctness of the received signature using its knowledge of r , RPuK and $hash(R')$.

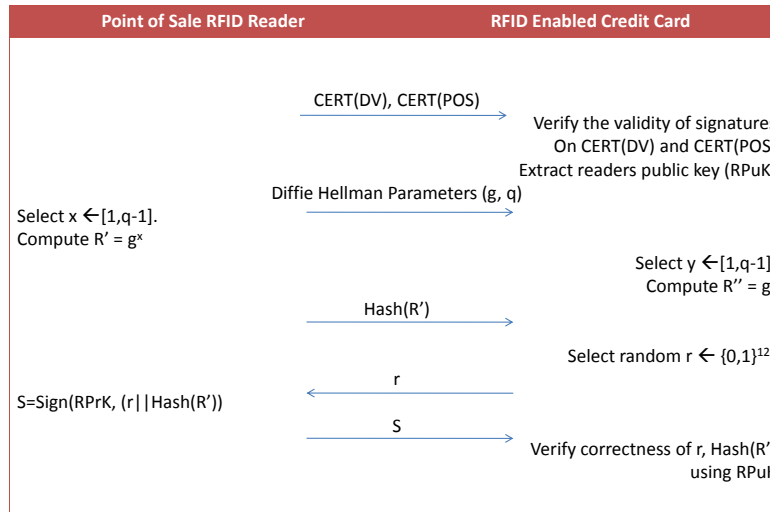


Fig. 6. POS RFID Reader Authentication Protocol

What actually happened? At the end of this protocol execution, if executed successfully, the Reader has proved knowledge of the private key that corresponds to the public key in its certificate. The Tag also now has the hash of the Readers ephemeral public key.

Informal Security Analysis The security of the Terminal Authentication protocol is straightforward since we do not concern ourselves with the certificate revocation checking process in this paper (this is dealt with in prior work [36]). Only the POS Terminal which has the private key that corresponds to the public key extracted from the D1CA and DV signed certificate will succeed in authenticating itself to the Tag. The security of the protocol can thus be reduced to the security of the underlying signature scheme.

4.4 Tag Authentication and Establishing a Secure Communication Channel

The Tag Authentication protocol is executed after the POS Terminal Authentication protocol is executed successfully. To ensure that the same POS Terminal that was authenticated is used in the Tag Authentication protocol, we tie the two protocols together by requiring the use of the previously generated Diffie-Hellman keys in the Tag Authentication protocol to enable secure messaging.

In this protocol the Tag initiates communication by sending the Terminal its public key (signed by the issuing DV. For verification, the DVs public key may be obtained from the EMV owned Public Key Directory). In addition to this, the Tag also sends the Terminal a copy of its ephemeral public key (R'') that was generated in the previous protocol. After verifying the correctness of the DVs signature, the terminal uses the Tags public key to encrypt its ephemeral Diffie-Hellman public key and sends this to the Terminal. The Terminal decrypts the key and verifies its correctness by comparing it to the hashed version it obtained in the previous round. After successful verification, both parties possess a shared secret (g^{xy}). This shared secret is used as a seed key to generate a new encryption and MAC key for future communication (in which regular transaction data will flow). The new symmetric keys for secure messaging may be obtained by simply hashing the seed key repeatedly.

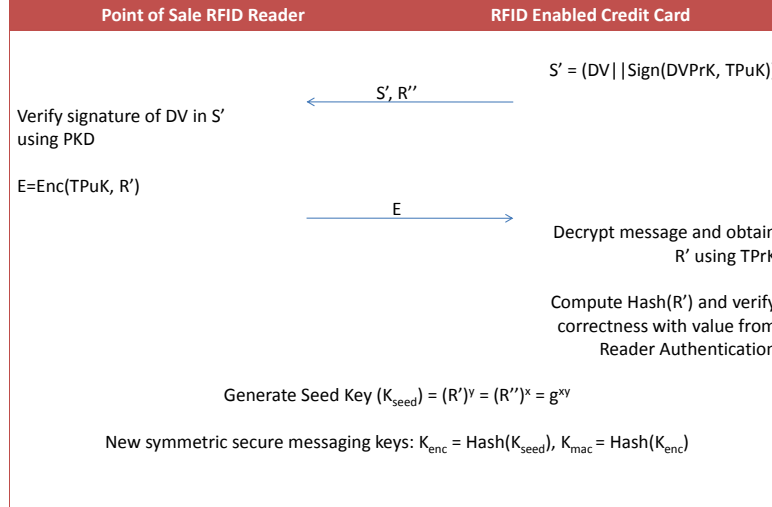


Fig. 7. Tag Authentication

What actually happened? At the end of this protocol execution, if executed successfully, the Tag has proved that it has knowledge of the private key that corresponds to the DV issued certificate. The Tag and Reader have also established a secure messaging channel based on the shared secret obtained from the ephemeral Diffie Hellman protocol.

Informal Security Analysis S' is precomputed and stored in the Tags memory. It is made available to the Reader after the Terminal Authentication protocol. The Reader extracts the DVs identity and uses the PKD to find the DVs public key. Once this is complete the Reader knows the Tags public key (TPuK) which it uses to encrypt the Readers ephemeral public key. Security in this step comes from the fact that the PKD is a trusted directory set up and maintained by the EMV and the public key cryptosystem is secure.

Forgery of these RFID enabled credit cards is not possible since the private key that corresponds to the DV signed public key is stored in inaccessible secure memory on the Tag. On completion of authentication and exchange of ephemeral public keys, the shared secret g^{xy} may be computed by both parties. Only an adversary that is capable of guessing the shared secret will be capable of eavesdropping on the (sensitive) data exchanged after this point. The shared secret is generated using the ephemeral Diffie-Hellman protocol which is secure even in the presence of active adversaries when used with certified public keys (as is the case here).

4.5 User Authentication

The above protocols obviate eavesdropping and compromised terminal attacks. Now, we turn to proper user authentication to deal with relay attacks and impersonation attacks. Once the Terminal and Tag are authenticated and secure messaging is enabled, it is the turn of the user to authenticate himself to the card. This is a critical part of the protocol stack. It is important that the user authentication procedure is secure, easy to understand, and usable by the average Joe. We need the user input to be either dynamic in nature, or static but hard to forge. We recommend the use of biometric user authentication techniques because of their resistance to forgery. Our protocol uses fingerprint recognition for user authentication because of the widespread availability of fingerprint readers in laptops and cell-phones. This allows the possibility of strong user authentication even in Card not Present environments. It also offers the best balance between security and usability [38]. The user registers their biometric data with the issuing bank (Document Verifier), this biometric data is encrypted and stored on the card (just as a PIN is). During the process of authentication, the user supplies their biometric to the authenticated Terminal (nowadays, these are usually capable of detecting gummy fingers) which sends it to the Tag using the established secure messaging medium. The Tag runs a matching algorithm and verifies the correctness of the fingerprint before authorizing the user. There is no need for the use of trusted third party devices even in public environments since the Terminal is authenticated before the biometric is supplied.

4.6 Operation Procedure

1. The RFID Tag is placed within readable distance (10-15 centimeters) from the Reader.
2. The Reader sends the Tag its Document Verifier issued certificate and a Certificate Revocation List (CRL). The Tag verifies the validity of the issued certificates using the CRL.
3. The Reader authenticates itself to the Tag using the described POS Terminal Authentication protocol.
4. Once the Reader is successfully authenticated by the Tag, the Tag authenticates itself using the described Tag Authentication protocol. Secure messaging is started from this point on.
5. The user now authenticates himself to the card by supplying his fingerprints to the authenticated POS Terminal. The biometric matching algorithm is run on the card rather than on the Terminal. This is feasible for contactless smart cards [39].
6. The card supplies the magnetic strip data to the authenticated Reader. The received information can be verified by comparing it with the signed SDE on the card. This data is then sent by the Reader along with other transaction data to the banks transaction processing server where the transaction is authorized.

The user is required to provide an input only in the user authentication phase after the Terminal and Tag are authenticated, making the procedure easy to understand.

5 Conclusions

The Chip and PIN payment system has failed to deal with some very basic attacks that have been carried out on ATMs such as - Observation attacks with hacked and counterfeit terminals and terminal interception attacks. Other more complex attacks that may be carried out are - relay attacks and phishing attacks. The attacks performed on RFID enabled credit cards include Skimming attacks, eavesdropping attacks, user tracking, replay attacks, and relay attacks. A major problem is that the RF signal from the Tag (the card) to Reader is sent in the same form as magnetic strip data in plain text for most RFID enabled credit cards, making them vulnerable to the mentioned attacks. While the CAP card is by far the most secure payment card and has been difficult to attack by traditional means, their inconvenience to the user makes them easy to compromise. Our analysis shows that current cashless payment systems fail because of a combination of reasons such as poor user authentication, lack of a Reader authentication protocol, poor usability, and easy access to sensitive data by unauthorized Readers.

To deal with this, we proposed a protocol stack - SECAPS which makes use of an RFID based PKI, mutual authentication, and biometric authentication to secure a transaction. The procedure is easy for the user to understand, does not require the use of any trusted third party devices such as CAP Readers, does not work unless the user gives consent in the form of a biometric, and can be used even in Card not Present transactions. The cost of implementation and operation of such a system is estimated to be less than half of the amount lost to credit card fraud in a year.

