The Implementation of a full EMV Smartcard for a Point-of-Sale Transaction

Oludele Ogundele, Pavol Zavarsky, Ron Ruhl, Dale Lindskog
Department of Information Systems Security Management
Concordia University College of Alberta
7128 Ada Boulevard, Edmonton, Alberta, Canada
delesquared@gmail.com, {pavol.zavarsky,ron.ruhl,dale.lindskog}@concordia.ab.ca

Abstract— **This paper examines the changes in the payment card environment as they relate to EMV (named after Europay, MasterCard and Visa). This research shows that if the combined dynamic data authentication (CDA) card variant of the EMV card is deployed in a full EMV environment, given the relevant known vulnerabilities and attacks against the EMV technology, the consequences of unauthorized disclosure of the cardholder data is of significantly reduced value to a criminal.**

Keywords- EMV, Magnetic-stripe, Chip and PIN, Payment card, Point of sale terminal.

I. INTRODUCTION

Presently, there is an estimated 1.2 billion active EMV smart cards used for credit and debit payment worldwide (of different variants, processing options and capabilities) at over 18.7 million EMV acceptance terminals [1]. EMV smart cards, commonly referred to as "Chip and PIN" by bank customers, were designed and introduced by the banking system as a solution to the high rate of financial fraud occurring by the use of magnetic-stripe cards and also to improve the security of payment cards used in face-to-face environments. The PIN is used to prevent the abuse of stolen or lost cards while the chip is used to protect against card counterfeiting. The goal of the EMV chip technology is to secure both credit and debit card transactions by authenticating the card and the individual presenting the card at the point of sale (POS) terminal and sometimes the transaction itself. However, despite the wide deployment of EMV technology in various parts of the world, there still exist successful attacks and known vulnerabilities that undermine its success in combating fraudulent activities that occur in face-to-face environment once there is an unathourised disclosure of the cardholder data during a transaction.

The present EMV deployment in the payment card is referred to as the "hybrid environment", in which the payment card has both the magnetic stripe and the chip embedded into it. The main purpose of the chip is to securely store the cardholder data and protect the data stored against unauthorized modification and also to reduce the number of fraudulent transactions carried out with the use of counterfeit, lost and stolen cards [2], but despite these improved security measures, the payment cards are still not immune to some known attacks. This is due to the presence of the magnetic stripe on these cards and the type of the EMV variant of the smartcards used by the banking systems. At the moment, countries in Europe, North America (Canada), Latin America and Asia are all at various stages of the EMV chip migration from magnetic stripe cards except the United States, where the pressure is mounting for the adoption of the EMV smartcards for face-to-face banking transactions.

1.1 Our Contribution

In this paper, we will show by the comparative analysis of the security features of the different variants of the EMV payment cards, the need for the banking systems to move from the present hybrid environment to a full EMV environment (where only the chip is embeded in the payment cards and the magnetic stripe removed) and the need for the implementation of the CDA variant of the EMV chip technology in order to overcome known successful attacks and vulnerabilities against the EMV technology (SDA and DDA).

The structure of this paper is as follows: In Section 2, we provide an overview of the EMV transaction process and the different variants of the EMV payment cards. Various forms of attack against the different variants of EMV payment cards and the consequences of unauthorized disclosure of cardholder data in hybrid and full EMV environments will be described in the last section.

II. AN OVERVIEW OF THE EMV TRANSACTION PROCESS AND THE DIFFERENT VARIANTS OF EMV PAYMENT CARDS

In this section we will give an overview of the EMV transaction process. We will also compare the security features of the different variants of the EMV payment cards.

2.1 The EMV Transaction process

During an EMV payment card transaction, the chip on the smart card must make contact with the chip reader in an acceptance terminal (e.g. POS terminal). This terminal connection could be either contact or contactless. For the contact method, there must be a physical contact between the chip and the card reader for transaction to occur while for the contactless, the chip is required to be within sufficient proximity to the card reader for information exchange between the chip and the acceptance terminal [2]. The EMV specification details the technical requirements for chip embedded payment cards and for the various point-of-sale infrastructures.

In the EMV chip technology, if the transaction is approved, the Intergrated Circuit Card (ICC) generates a Transaction Certificate (TC) which is passed to the terminal and used to claim payment during the clearing process. On the other hand if the transaction is declined, the ICC generates an Application Authentication Cryptogram (AAC) message. If the transaction needs to be approved online, the ICC generates either an AAC or Authourization Request Cryptogram (ARQC), which is sent to the issuer (e.g. the cardholder's bank). The issuer replies with the Authourization Response Cryptogram (ARPC) message signifying wheather the transaction should be approved (the ICC generates the TC message) or declined (then the ICC generates the ACC) [3]. The terminal manages the level of risk by requiring certain

transactions to be authorized online instead of being authorized locally in order to safeguard against fraudulent use. This includes checking the transaction amount against the floor limits, detecting when the defined limit of consecutive offline transactions have been reached and in addition, offline-capable terminal will also randomly select certain transactions to be performed online. The online processing (depending on the variant of the EMV chip card used) allows the card issuer to analyze the transaction details and to decide whether to go ahead with the transaction or reject it. This provides the issuer with the opportunity to check the account status and apply criteria based on the acceptable limits of risk predefined by the issuer, the acquirer and the payment scheme (e.g. Visa, Mastercard). If a valid response is received from the issuer, the card analyzes the result of the online processing, authenticates the data received from the card issuer and request the terminal to complete the transaction by either declining or accepting the transaction based on the result of the online processing. Therefore, the card cannot request the acceptance of the transaction if the issuer declined the payment [8]. At the completion of the transaction, the card is removed from the chip-reading device.

In summary the payment transaction process of the EMV protocol can be divided into three phases namely card authentication, cardholder verification and the transaction authorization. The card authentication phase assures the terminal of the legitimacy and authenticity of the card. The cardholder verification phase confirms to the terminal that the PIN entered by the cardholder matches the one assigned to the card. The transaction authorization phase reassures the terminal that the card issuer (i.e. the bank) approves the transaction.

2.2 The Different Variants of the EMV Payment Cards

The payment card issuer (i.e. the cardholder's bank) selects the EMV card variant to use and also have the autonomy to choose the subset of the EMV protocols that will be implemented on the card. This includes making decisions concerning the digital signature methods, message authentication code (MAC) algorithm to be used for the card processing, card authentication and risk management options; ensuring that the selected options comply with both the payment card network rules and the EMV framework [4]. The EMV specification defines three main types of card authentication methods otherwise known as variants of the EMV chip card namely, the static data authentication (SDA) card, dynamic data authentication (DDA) card and the combined dynamic data authentication (CDA) card [5]. The EMV chip technology supports cardholder authentication through the use of Cardholder Verification Method (CVM) and ICC authentication by the use of the SDA, DDA or CDA cards [3]. Although the EMV specification supports signature authorization, the PIN verification method remains the dormant method used for the CVM.

Table 1 shows the different variants of EMV smart cards emphasizing the basic characteristics that differentiate each card from the other. The SDA card is a cheaper alternative for the banking industry considering the cost of production because SDA uses only a symmetric key on the card, which is shared only with the issuing bank. The drawback with the use of the SDA card is that it is susceptible to the various successful attacks as evident in [4], [6], [7], [8], [9]. The terminal sends a summary of the transaction data once the PIN entered by the cardholder is verified by the card and the MAC is generated using the shared symmetric key with the issuer. The SDA card can only prove a payment card is genuine only if the terminal is online (i.e. connected to a bank via a network). The SDA submits a static cryptographic certificate to the terminal, incorporating the primary account number (PAN) and a digital signature when performing card authentication [10]. The SDA cards are found to be even less secure when used with an offline terminal as compared to when the magnetic stripe cards are used, since it is the duty of the payment card to verify the PIN and reply with a "yes" or "no" to the POS terminal. The SDA cards can be easily cloned and can be used by criminals to perform fraudulent offline transactions. The issuer has a public-private key pair used for asymmetric cryptography if the transaction is required to be performed offline by the payment card. The POS terminal only knows the public key of the issuer because of the high risk involved in putting both keys on the terminal which could result into fraud if discovered.

The DDA cards are more complex than the SDA cards in that the chip in the card can do asymmetric cryptography and have a public-private key pair assigned to the card itself. DDA involves a challenge-response mechanism to prove the authenticity of the card by using a private asymmetric key (S_{IC}) to sign a challenge chosen by the terminal. It can verify that both the card and the information on the card have not been altered and also prove that the card is not a copy of the original card issued.

TABLE 1: THE CHARACTERISTICS OF THE DIFFERENT VARIANTS OF EMV SMARTCARDS

Static Data Authentication (SDA) Card	Dynamic Data Authentication (DDA) Card	Combined Dynamic Data Authentication (CDA) Card
Can prove a card is genuine online only.	Can prove a card is genuine either online or offline.	Can prove a card is genuine either online or offline.
Has only certificate of static data signed by issuer and certficate of issuer signed by CA.	Has same certificate of static data as SDA, "card level" Public Key signed by issuer, and certificate of issuer signed by CA.	Has same certificate of static data as SDA, "card level" Public Key signed by issuer, and certificate of issuer signed by CA.
Signature remains the same everytime the card is authenticated.	Signature is only valid for one card authentication.	Signature is only valid for one card authentication.
Digital certificate is signed by issuer's private keys.	Digital certificate has public key of card and the private key of the card stored in a secured area of chip.	Digital certificate has public key of card and the private key of the card stored in a secured area of chip.
Support only sym-metric cryptography, Terminal must have CA Certificate.	Supports asymmetric cryptography. Terminal must have CA Certificate.	Supports asymmetric cryptography. Terminal must have CA Certificate.
Card cannot protect itself against cloning and certificates for messages.	Able to generate digital certificate combining the time of transaction and card, cardholder and merchant details for each transaction.	More advance than the DDA card, able to generate digital certificate combining the time of transaction and card, cardholder and merchant details for each transaction. It also combines a dynamic signature and an application cryptogram and the data signed by the card include a random number provided by the terminal and the ARQC. [11]

The DDA card mechanism is used to prevent card cloning and to authenticate the genuineness of the card but not the subsequent transactions performed by the card [12]. The DDA and the CDA cards are capable of performing dynamic data authentication while the SDA card can only perform static data authentication. The unique feature of dynamic data authentication is that it prevents the transaction data from being fraudulently reused even if it was stolen from the processor's or merchant's database thereby making the data of little or no use to the criminal.

The CDA cards are made as an advancement of the DDA cards, in that the message signed by the ICC includes an additional Application Cryptogram (AC), which is used to protect the transaction messages generated by the ICC using the AC Session Keys (derived from the ICC AC Master Key, shared only between the ICC and the card issuer). The ICC AC Master Key is unique per card and it is derived by a combination of the cardholder's PAN, the PAN sequence number and the issuer Master Key. The sequence number of the PAN is used to identify a specific card among several cards owned by the same bank customer with the same PAN [3]. This advance feature of the CDA card gives it the ability to overcome sophisticated types of attack at the POS terminal which is very difficult for both the SDA and DDA cards to surmount and thereby making the CDA card the most secure and the best alternative for use in EMV chip cards.

III. A SURVEY ON ATTACKS AGAINST PAYMENT CARD AND CONSEQUENCES OF UNAUTHORIZED DISCLOSURE OF CARDHOLDER DATA

In the magnetic-stripe environment, data which are embossed on the surface of the payment cards are stored in the magnetic strip. These include the cardholder's name, expiry date, card number and other necessary data such as the card verification value (CVV) of the card [10]. The magnetic stripe becomes easy to clone, once the data content of the magnetic stripe has been copied by a fraudster. Even though there are possible arguments on negligence on the part of the customer but in practical terms the bank ends up paying for the costs if a fraud eventually occur [9]. However, the magnetic stripe environment is however not of much concern as the financial institutions in most regions of the developed world have transitioned or are gradually transitioning beyond this stage except in the United States. This section will examine various attacks carried out in the hybrid environment and the possibilities of the success of the same kind of attacks in a full EMV environment in relation to the consequences of unauthorized disclosure of cardholder data.

A. Various Attacks against Payment Card in a Hybrid Environment

The various attacks against the payment card in the hybrid environment can be viewed as either cross-border or domestic frauds. The cross-border frauds are fraudulent transactions that are performed by criminals with the use of the payment card outside the country in which it was issued to the cardholder while domestic frauds are fraudulent transactions performed by criminals with the use of the payment card within the country in which it was issued.

The attacks applicable in cross-border frauds include attacks that rely on the fall back mechanisms of the magnetic stripe, which have been successful due to uneven migration to the EMV payment cards in various regions of the world. Fraudsters intensify efforts on compromising the magnetic stripe data so as to fraudulently use the POS terminal and ATMs in regions yet to migrate to the use of the EMV payment cards.

In the hybrid environment, cross-border fraud can be performed when the chip on the payment card presented at the POS terminal has been damaged or is unreadable. The terminal falls back to the magnetic stripe operation of the payment card, and this makes it possible for a fraudulent merchant to skim the customer's payment card successfully, using the obtained cardholder data to make a counterfeit card provided that the PIN has also been recorded as well. This means that a fraudster who is not able to clone the chip will simply copy the magnetic stripe to make a counterfeit card, which operates just like a legitimate card from the perspective of the terminal and transaction is permitted to proceed [9], [10]. The use of cloned payment card on the POS terminal has a high rate of success when the terminal is offline and the EMV chip variant is a SDA card. The fraudster does not even have to know the PIN, once a counterfeit card is made; he simply programs the payment card to say "yes" in response to PIN verification request irrespective of the PIN entered at the terminal [9], [10]. This type of attack is called the replay attack (also referred to as "yes cards"). It is important to note here, that if the cloned card is an SDA chip card from a magnetic stripe, the replay attack will not work online, since the PIN used will certainly not be correct (unless social engineering was used) otherwise the PIN Verification Value (PVV) will not be correct.

Similar attack method can also be done by using a stolen card that has both the chip (SDA) and magnetic stripe embedded in it in an environment where the POS terminal permits only the use of the magnetic stripe, which then enable the card to bypass the chip mechanism and fall back to the magnetic stripe operation. What enhances this attack is that the fallback mechanism in this situation tend to be less robust than the magnetic stripe mechanisms are in the environment where the magnetic stripe is the main technology used [9]. Still on the issue of backwards compatibility, for payment cards in certain countries like the UK and most European countries, that have both the chip and magnetic stripe - when used at ATMs that don't have chip readers or if the chip is damaged/unreadable, run the risk of falling back to the magnetic stripe operation. In such cases, fraudsters who have been able to clone the cards by learning the contents of the magnetic stripe and the cardholder's PIN can withdraw money from such ATMs by causing it to fall back to the older system or by using such cloned cards in a country like the United States that is yet to adopt EMV chip technology [6]. It should be noted that the fraudster may not necessarily have to read the magnetic stripe to reproduce the card, since the chip contains a copy of the magnetic stripe and this data is always sent to the card issuer during a legitimate transaction, the fraudster can therefore read the chip or intercept the communication between the terminal and the issuer to copy the magnetic stripe.

The attacks used for domestic frauds in the hybrid environment include damaged/unreadable card chips, replay attack, eavesdropping of the POS terminal, POS terminals with poor anti-tempering mechanisms, the man-in-the-middle attack, untrusted user interface, the back end API (Application Program Interface) attacks at the bank data center and finally the attack implemented during card-not-present transaction, which can also be used for cross-border fraud.

Attacks used for domestic frauds in the hybrid environment include the eavesdropping of the POS terminal. Once the account and PIN data can be eavesdropped from a transaction made by an EMV card at a POS terminal, it becomes easy to make a magnetic stripe containing that data for a fraudulent use in an environment where EMV chip technology is not supported. Alternatively, the fraudster can intercept the communication between the POS terminal and the chip to copy the magnetic stripe, and can also obtain the PIN entered by the customer as it is sent to the chip during the cardholder verification stage (i.e. if the PIN is sent to the chip in plaintext) [10]. There are various approaches by which the eavesdropping can be done but one of the preferred approach in terms of cost, development time and convenience is to create a skimming device (i.e. POS terminal skimmer) that conveniently sits on the smartcard slot of the POS terminal and used to capture the card and PIN data without the suspicion of the cardholder. This type of attack is usually employed to capture and store account details of majorly SDA cards in large quantities and this device can be removed quickly from the terminal should there be a problem or suspicion [13].

The POS terminals cannot be fully relied upon because they are viewed as being under the control of the potentially untrusted merchants and their proper functioning cannot be fully guaranteed and for this reason they are tamper resistant to prevent malicious merchants or staff from extracting cardholder data. Nevertheless these terminals are still prone to attack due to poor anti-tampering mechanism and research has shown that it may not even be necessary to open the device to do so; manipulating the power supply or transmitting a radio signal may be sufficient [13]. Most of these attacks are possible due to design errors in the tamper resistance protection measures discovered in some of the POS terminals and thereby making it possible to circumvent the protection in place [6], [14]. Drimer et al in [10] demonstrated that some of the POS terminals have inadequate tamper resistance, and therefore a possibility of adding a tapping device to record the cardholder's PIN and necessary details to allow a cloned magnetic stripe card to be produced [4].

The man-in-the-middle attack which was discovered to be successful in the hybrid environment, allowed the use of any payment card (e.g. stolen/lost) without the knowledge of the card's actual PIN as demonstrated by Murdoch et al in [4]. The "man-in-the-middle attack" (a.k.a. the wedge attack) is found to be effective against both offline and online transactions. This type of attack is possible because of the central flaw discovered in the EMV protocol used in the payment cards in the UK, in which the PIN verification step is never explicitly authenticated [4]. Since the POS terminal is permitted to communicate directly with the legitimate card during card authentication, with the use of a stolen card, a man-in-the-middle attack can be used to intercept and modify the communication between the stolen card and the terminal so that during the cardholder verification, the man-in-the-middle suppresses the messages as they are being sent to the stolen card. Therefore, irrespective of the PIN entered by the fraudster the man-in-the-middle tells the terminal the PIN entered is correct and the transaction is authorized to proceed. This attack is also successful when the transaction is done online, which is due to the oversight of the design of the transaction authorization stage. It was discovered in the EMV specification that the ARQC (Authorization Request Cryptogram) and TC (Transaction Certificate) messages include the result of the cardholder verification but the result only indicate whether the verification was attempted but failed, but does not distinguish if the verification succeeded or if it was not attempted at all. Therefore, there is the high possibility that the man-in-the-middle can suppress the cardholder verification and then relay the ARQC and TC between the legitimate card and terminal. So when the issuer receives these cryptograms, the authorization would succeed because the cryptograms are seen as being sent from a legitimate card and the bank would accept the transaction [10], [20].

Another concern when considering the hybrid environment is the user interface trust issue. This in fact, questions the assurance of the cardholder concerning the amount displayed on the POS terminal during a transaction. What if the POS terminal has been compromised? When the cardholder puts the card in a POS terminal, what is the guarantee that the payment card is genuinely interacting with the terminal? This serves as an advantage for a fraudster to implement a relay attack. This is a situation whereby the POS terminal has been compromised by a fraudulent merchant and the unsuspecting cardholder is made to authorize payment for another transaction done elsewhere, while thinking that the payment being made is based on the current transaction done at his present location [9], [7], [14]. The relay attack is done successfully because the compromised terminal appears legitimate to the untrained cardholder, who gets fooled easily without suspicion. The relay attack is outlined in [9] and demonstrated by Drimer and Murdock in [7].

The back end API (Application Program Interface) attacks at the bank data center where the HSM (Hardware Security Modules) is located with the aim of providing necessary backend support to EMV chip cards such as processing authorization requests and responses, and also sending secure messages. This attack is serious because it reveals that the EMV protocol has not been able to mitigate the risk of abuse by bank programmers at operation centers and insider attacks which can circumvent the system. This was possible with the inevitable implementation issues with protocols coming out of the woodwork and with such an attack, the key update message between the bank and card can be eavesdropped, and then a cloned card produced or PINs could be discovered at will [14].

Also, the dotcom boom which has resulted into huge activities of card-not-present transactions during online shopping, attacks which make use of "phishing attack", aimed at capturing online banking details of customers, perpetuated by sending emails impersonating banks, asking customers to click a link under false pretense that their accounts have been compromised, and when they do, a malicious copy of their bank's website asking for their account details or authentication data is displayed. This form of attack on the EMV chip cards result into online fraud and have been reduced with the use of Chip Authentication Program (CAP) introduced by the banks [15], [16]. EMV was adapted to the CAP protocol and used to generate codes for two-factor authentication [17]. Another common attack for online transaction involves the use of malwares in which authentication details are stolen by a software key-logger installed on the customer's computer [15]. Other social engineering tricks – by telephone or post (this is by brazenly calling up customers and asking for their PIN or CVV or using postal redirection scams) [9], [14].

The '3-D Secure' protocol was also introduced by Visa (Verified by Visa) and MasterCard (MasterCard SecureCode) to help reduce online fraud during card-not-present transactions. It is essentially a single-sign on system. Although it has flaws in its implementation, it was adopted easily by merchants because of the contractual terms of

liability which offered reduction of liability in case of disputed transaction for merchants [16].

3.1.1 Consequences of Unauthorized Disclosure of Cardholder Data in a Hybrid Environment

The consequences of unauthorized disclosure of the cardholder data in a hybrid environment, differs based on the variant of the EMV payment card used. It should however be noted that in the hybrid environment, the common payment card in use are the SDA and the DDA cards [1], [4], [20].

When the SDA card is used, the consequences may be the successful cloning of payment cards by taking advantage of the fallback mechanism to copy the magnetic stripe, the use of skimming devices to eavesdrop on the POS terminal to capture and store account details and PIN of the card made possible by the presence of the poor anti-tampering mechanisms [6], [13]. In reality, the consequences of unauthorized disclosure of cardholder data relate to the risk involved if the PIN and card details are intercepted when they are sent unencrypted between the card and the PIN entry device (PED) during a transaction, a fake magnetic stripe card can be produced by fraudsters and due to the backward compatibility of the EMV smartcards, the possibility of falling back to the older system if the chip is damaged or unreadable [6], [13]. Since the PEDs are expected to protect both the cardholder's PIN and card details, the presence of vulnerabilities in the PEDs will lead to disclosure of cardholder data which indicates a high risk of fraud of which the cardholder will still be held accountable in most cases because he would have entered his PIN during the transaction. It was also discovered that the PEDs appear to protect both the merchants and banks secrets while leaving the cardholder's card details and PIN inadequately protected [6], [13].

When the replay attack is used in the case of a SDA card, the attack is easily successful with the unsuspecting cardholder entering the PIN willingly to authorize the transaction while if the DDA is used, the replay attack will not be successful because of the challenge-response mechanism built into the chip [12]. The consequences of unauthorized disclosure of cardholder data with use of the replay attack on SDA cards will result to successful fraudulent transaction done by criminals with the use of cloned SDA cards.

However, the man-in-the-middle-attack as demonstrated in [4] will still be successful against both the SDA and the DDA cards because the terminal cannot verify the transactions performed by these types of EMV chip variants when an offline PIN verification method is used. The success of this attack could be attributed to the cardholder verification phase, in which the offline PIN verification is not authenticated. This hereby enable a criminal to use man-in-the-middle attack to trick the terminal in such a way that any PIN entered is accepted as correct while the card suppose that the CVM was done by the cardholder signature option. The consequences of unauthorized disclosure of cardholder data with use of the man-in-the-middle attack will result to the ability of criminals to modify the transaction elements to perform successful fraudulent transactions without entering the valid PIN number.

Also, in the case of resolving disputes between the cardholder and the bank over a transaction, the reliability of the EMV smartcard technology and the evidence generated has become increasingly important. In a situation whereby, the man-in- the-middle attack is used by a criminal to compromise a genuine payment card of a customer. When the defrauded customer eventually notifies the bank of the unauthorized transaction, the disputed transaction is often reversed under the old system(i.e. the magnetic stripe) but in the EMV chip technology, the bank may alternatively believe that the customer has acted negligently by not protecting the payment card or PIN adequately or both the card and PIN. If such a decision is challenged, the bank frequently sees the EMV chip technology as infallible, stating categorically that the customer's payment card was read and the PIN was used for the transaction thereby concluding that the customer must be grossly negligent or lying. In other words, in resolving disputes, the consequences of unauthorized disclosure of cardholder data makes it difficult for the customer to obtain refunds once transaction made was authorized by PIN except when proved otherwise by fraud experts. Anderson et al [9] described how cardholders might have difficulty in obtaining refunds when a fraudulent transaction was proven to have been authorized by PIN. Furthermore, in the case of a disputed transaction, if the transaction was authorized by a signature the merchant will be held liable for fraud while if the same transaction had been authorized by PIN, the cardholder is liable for fraud [4]. Arguably, the bank is expected to show that their position is defensible and that the transaction was not performed by a third-party fraudster exploiting a security vulnerability of the system [4], [10]. Murdoch et al [4] described and demonstrated the EMV flaw that allowed fraudsters to use a genuine card to make fraudulent transaction without knowing the card's PIN and to be undetected despite the fact that the merchant has an online connection to the banking network and thereby making it evident that the fact the terminals print "verified by PIN" does not mean that the valid PIN was entered for that particular transaction [4].

For attack done with the use of replay attack, the customer is not expected to lose money but since it was done offline the merchant will be held accountable for this kind of fraud because he will be accused of negligence once the fraudulent transaction is reversed, for not verifying transaction with the issuer of the card before proceeding with the transaction [10].

Ultimately, the issue of liability shift (from the payment card issuer to the cardholder and merchant) was never considered as significant under the magnetic stripe technology because the cardholder can always complain when he suspects a fraudulent activity on his bank statement or when charged wrongly for a purchase, but under the hybrid environment, the conditions are totally different, and the liability has shifted from the payment card issuer (bank) to the cardholder and the merchant [4], [6] therefore making the consequences of unauthorized disclosure of cardholder data very serious as relate to the cardholder and merchant when exploited by criminals.

B. Possible Attacks against Payment Card in a Full EMV Environment

In a full EMV environment, the magnetic stripe will be removed or phased out from the payment cards, and only the chip will be used for business transactions at the POS terminal. Technically, in a full EMV environment, frauds that are perpetuated by taking advantage of the presence of the magnetic stripe in the hybrid environment such as the use of fallback mechanism and card cloning will be drastically reduced [20]. In essence, EMV implementation that utilize the different card verification values present on the chip (iCVV – Card Verification Value for ICC) from those available in magnetic stripe image (CVV - Card Verification Value) will

prevent the cloning of payment cards from compromised EMV magnetic stripe-image data and also limit the impact of lost/stolen/mail not received types of fraud [6], [20].

EMV does not protect the confidentiality of or the inappropriate access to sensitive authentication data and/or cardholder data during any point of the transaction at the POS terminal [20]. This is because the PAN is needed to be processed in clear text to complete critical steps in the EMV transaction. The data elements of the cardholder include the PAN, cardholder name (which is not required to be transmitted) and expiration date. These are all visible on the surface of the payment card, except the service code which is only present in the Track 2 Equivalent Data on the chip. The PAN is sent in the clear during EMV transaction in order to identify the cardholder and facilitate the transaction routing, enable key derivation and perform data authentication at the POS terminal. The expiration date is sent in clear text during an online authorization and is included in the Track 2 Equivalent data. The data elements of the sensitive authentication data for a full EMV transaction are the ICVV and the PIN [20]. The EMV specification allows for offline verification of the cardholder through the use of the PIN, the online PIN verification for the cardholder verification is also supported depending on what is required in the CVM process and also the presence of the iCVV stored on the EMV chip card will prevent the production of counterfeit cards [20].]. This means that in a fully EMV environment, the SDA cards will no longer be vulnerable to the well-known replay attack because the PIN, along with the other card information cannot be used to create a counterfeit magnetic stripe version of the card [4], [20], [18] but still the SDA card is not a good option to be used in a full EMV environment because it can only use the PIN to verify the cardholder only when it is online (i.e. connected to the bank network in real-time) and it will still be vulnerable to attacks, such as the man-in-the-middle attack and cannot also resolve the user interface trust issue experienced because this type of card is only verified by the static digital certificate signed by the issuer that it presents to the terminal. It cannot also do RSA asymmetric cryptography and thus cannot sign any transaction with a private key.

For the EMV chip technology, one of the PIN algorithms used is the VISA PIN Verification Value (PVV). It is the cryptographic signature of the PIN and other cardholder data. Using VISA PVV algorithm with 3-DES encryption is considered very resistance (very difficult to brute force) and is regarded as secure, therefore making the chip impossible to counterfeit. The encryption process it involves serves as an antifraud mechanism. L. Padilla made a demonstration of the attempt to brute force VISA PVV algorithm with DES encryption in [19].

The DDA cards can do RSA asymmetric cryptography and also involves a challenge-response mechanism to prove the authenticity of the card by using a private asymmetric key, S_{IC} to sign a challenge chosen by the terminal. It can prevent the replay attack in an offline transaction but not the man-in-the-middle attack and cannot resolve the user interface trust issue. This is because the DDA card cannot prove the authenticity of a transaction done by a card, in such a way that it is able to verify that the transaction was done by that particular card since it was not designed to have that ability.

The man-in-the-middle attack can be eliminated in the full EMV environment with the use of the device called “Smart Card Detective” as demonstrated by Choudary [8]. This same device can also help to resolve the user interface trust issue due to the flaws discovered in the PED used for payment by the customer to the merchant. This device is able to intercept the communication between the card and the terminal and therefore can provide the cardholder with the ability to observe the amount requested by the terminal and the option to continue or reject the transaction based on the cardholder’s discretion. This experiment is demonstrated in [8], and further used for test scenarios in [17]. For Choudary’s experiment, a specially designed chip card was used but without the magnetic-stripe, and it was built on the same ISO 7816 standard as required for any EMV chip card, so as to make it compatible and also provide necessary functionality.

Alternatively, the man-in-the-middle attack will not be successful against the CDA card because during the cardholder verification stage, the CDA combines the authentication of the card with actual transaction done with the card which will result in the issuer authorizing the transaction therefore any alteration of the transaction elements implemented by man-in-the-middle attack will invalidate the signature on the message generated by the card and the transaction will be denied [12].

Therefore, the CDA card can also be used to eliminate the man-in-the-middle attack, relay attack and also resolve the user interface trust issue in a full EMV environment. This is so, because once there is any alteration in the elements of the transaction, the issuer will not authorize the transaction because a different message will be generated by the card which will not correspond with the message received by the issuer [12].

The relay attack demonstrated by Drimer and Murdock in [7] will be successful against SDA or DDA chip card but not against the CDA chip card because the CDA chip card computes a MAC, encrypting it with K_{MAC} (i.e. the key of the MAC) over the details of the transaction and interestingly K_{MAC} can only be computed by the chip in the card and the issuer. Therefore, any cardholder data or sensitive authentication data extracted by a fraudster during an EMV transaction such as PAN will be of reduced significant value, due to the dynamic nature of the CDA card because the value (amount) of the transaction coupled with the PAN and the transaction and all these are encrypted with the MAC so the cardholder cannot buy beyond the amount specified in the transaction number and thereby preventing the success of the relay attack. This explanation is also applicable to the man-in-the-middle attack and the user interface trust issue. With all these major successful attacks experienced in the hybrid environment being solved with the use of the CDA card variant of the EMV cards, the CDA card is the most preferred EMV smartcard [12], that all banks will eventually want to migrate to, coupled with reasons such as the card can verify the legitimacy of the card being used by the cardholder; the assurance that the cardholder knows the PIN and that the transactions have message authentication code (MAC) so that they can’t be relayed and misrepresented because the only place the secret key is on the secured area of the chip and with the card issuer (i.e. the bank). Only the bank can recreate the key at any time because it has the master key and it created the symmetric key by the formula it applies to the customer details and their master key. The EMV protocol itself is indeed robust and can be adapted to different technology in the face of changing business demand as explained in [17].

3.2.1 Consequences of Unauthorized Disclosure of Cardholder Data in a Full EMV Environment

The consequences of unauthorized disclosure of the cardholder data in a full EMV environment is reduced to a significant extend because it is difficult to make a copy of the

original chip on the card [19] and also in a full EMV environment, a chip-generated dynamic data element is done uniquely by the use of the CDA card for each transaction made in a face-to-face transaction, coupled with a robust authentication process for card-not-present (CNP) transaction, this makes it difficult to create counterfeit cards even if the cardholder data (e.g. PAN) or authentication data (e.g. PIN) is sent in clear-text. EMV transaction is therefore capable of reducing fraud value of cardholder data and preventing the compromise of sensitive authentication data [20]. This situation helps to prevent the fraud liability from being dumped on the cardholders or merchants by the banking industry as opposed to the practice in the hybrid environment. Even in a situation in which the cardholder data is compromised the consequences of the unauthorized disclosure of the data in a full EMV environment with the deployment of CDA card is significantly reduced in that the breach of the cardholder data confidentiality cannot be used for fraudulent transaction by criminals.

The implementation of the CDA variant of the payment card also help to provide valid evidence of transaction in case of a transaction dispute between the cardholder and the merchant which may also involve the issuer or acquirer. It also helps to protect the payment card system from abuse by any of the parties involved in any transaction.

Given the above findings about the consequences of unauthorized disclosure of cardholder data in a full EMV environment compared with that of a hybrid environment, we can come to the conclusion that the consequences of unauthorized disclosure of cardholder data in a full EMV environment in a face-to-face transaction is significantly reduced because of the ability to protect MAC message with a symmetric key and protect authentication with an asymmetric key.

But it should be noted that as the technology improve so also will criminals move with the trend with time. However, this research is not intended to focus or put into consideration the privacy legislation in certain regions of the world, in which the privacy law may prohibit the disclosure of the full PAN of the cardholder, in a bid to protect the private/personal information of the individual presenting the card. In such an environment, this research will likely be viewed from the perspective of privacy rather than for security reasons.

C. Cost Implementation of the CDA variant of the EMV chip cards

Initially, at the early deployment of the EMV chip cards, the SDA variant was used. The rationale is that the SDA cards were the cheapest to produce but a critical examination of the present hybrid environment makes it evident that the EMV technology has to support the security requirement of the "weaker" technology (i.e. the magnetic stripe), therefore the hybrid terminals and system must provide adequate protection for the cardholder data and PIN while being processed through the payment system [20]. The fact that financial institutions (banks) are able to issue DDA variant of the EMV payment cards in the hybrid environment makes it evident that the cost to implement the CDA cards may not be too expensive after all since they both involve the use of the private and public keys for the use of asymmetric cryptography and the primary difference between the two variances is that the CDA card is able to authenticate the card as well as the transactions performed by the card.

In contrast, a full EMV environment will benefit from the lower cost of implementing a chip-only infrastructure since most of the needed infrastructure are already in place via the hybrid environment and coupled with better payment card security by the deployment of the CDA chip card which will not involve the integration of complex software and hardware components as seen in the present hybrid environment, thereby making the cost of system integration and units cheaper for merchants, processors and banks.

However, it should also be noted that the cost of EMV chip card production has reduced over the last couple of years which helps to resolve the issue of the cost of card production. Also the card production cost will further decrease if banking system in the United States eventually adopts the EMV chip technology due to the presence of large customer base and the purchasing power of the big market players [20].

IV. CONCLUSION AND FUTURE WORK

We have been able to show based on the survey on the known attacks against the payment card and consequences of unauthorized disclosure of cardholder data that the CDA cards will be the most desirable option if EMV becomes the sole means of payment in a face-to-face transaction, coupled with a globally robust authentication process for card-not-present (CNP) transactions and the need to keep the PAN and other sensitive authentication data confidential would be significantly reduced. The future work of this research can be done in the area of implementation of a full EMV transaction using mobile phones for payment transaction due to the advancement and the increase use of smart phones worldwide for payment transaction such as direct mobile billing due to the significant progress made in the field of telecommunication.

ACKNOWLEDGMENT

The first author will like to thank the research team of Concordia University College of Alberta for their positive criticism, support and guidance in the completion of this research. He will also like to thank his parents, siblings and Katherine W. Mitchell for their support and love throughout his studies and also his profound thanks to God Almighty for the gift of life.

REFERENCES

[1] Douglas King, "Chip-and-PIN:Success and Challenges in Reducing Fraud" Retail Payments Risk Forum, January, 2012. [Online]. Available: www.frbatlanta.org/documents/rprf/rprf_pubs/120111_wp.pdf [Jan. 30, 2012].

[2] A Guide to EMV, Version 1.0, May 2011. [Online]. Available: http://www.emvco.com/best_practices.aspx?id=217 [Oct. 10, 2011].

[3] S. Balfe and K.G. Paterson, "e-EMV: Emulating EMV for Internet payments using Trusted Computing Technology", Information Security Group, Royal Holloway, University of London, UK. 2008. [Online]. Available: http://www.isg.rhul.ac.uk/~kp/EEMV.pdf [Oct. 20, 2011].

[4] S. J. Murdoch, S. Drimer, R. Anderson, and M. Bond, "Chip and Pin is Broken "IEEE Symposium on Security and Privacy, p 433-446, 2010, 2010 IEEE Symposium on Security

and Privacy, SP 2010. pp. 433 – 444. [Online]. Available: www.cl.cam.ac.uk/~sjm217/papers/oakland10chipbroken.pdf [June 12, 2011].

[5] EMV – Integrated Circuit Specifications for Payment Systems, Book 2: Security and Key Management, version 4.2 ed., LLC, June 2008. [Sep. 9, 2011]

[6] Drimer, S., Murdoch, S.J., Anderson, R. "Thinking inside the box: system-level failures of tamper proofing". In IEEE Symposium on Security and Privacy (Oakland).(May 2008) pp. 281 - 295. [Online]. Available: www.cl.cam.ac.uk/techreports/UCAM-CL-TR- 711.pdf [Sep. 9, 2011].

[7] Saar Drimer and Steven J. Murdoch. Keep your enemies close: Distance bounding against smartcard relay attacks. In USENIX Security Symposium, August 2007. [Online]. Available: www.saardrimer.com/sd410/pres/usenix07_relay.pdf [Sep. 12, 2011].

[8] Omar S. Choudary., "The Smart Card Detective: a hand-held EMV interceptor", June, 2010. [Online]. Available: www.cl.cam.ac.uk/~osc22/docs/mphil_acs_osc22.pdf [Nov. 9, 2011]

[9] Ross Anderson, Mike Bond and Steven J. Murdoch. "Chip and spin", March 2005. [Online]. Available: http://www.chipandspin.co.uk/spin.pdf [Sep. 14, 2011].

[10] Steve J. Murdoch, University of Cambridge Computer Laboratory, 2009 [Online]. Available: www.cl.cam.ac.uk/~sjm217/papers/deaeslr09reliability.pdf [Sep. 15, 2011].

[11] Keith E. Mayes and Konstantinos Markantonaiks, "Smart cards for Banking and Finance" in *Smart cards, tokens, security and applications,* New York: Springer Science+Business Media, LLC. NY, 2008, pp. 120 -125.

[12] Joeri de Riuter and Erik Poll, Formal Analysis of EMV Protocol Suite, Digital Security Group, Radboud University Nijimegen, Netherlands. 2011 [Online] Available: http://www.cs.ru.nl/E.Poll/papers/emv.pdf [Dec. 14, 2011].

[13] Saar Drimer, Steven Murdoch and Ross Anderson, "Failures of Tamper-Proofing in PIN Entry Devices" in IEEE Security and Privacy v 7 no 6 (Nov- Dec 09), pp. 39 - 45. [Online]. Available: www.cl.cam.ac.uk/~sjm217/papers/ieeesp09tamper.pdf [Sep. 16, 2011].

[14] Adida, B., Bond, M., Clulow, J., Lin, A., Murdoch, S. J., Anderson, R. J., and Rivest, R. L. Phish and chips (traditional and new recipes for attacking EMV). In Security Protocols Workshop (Cambridge, England, March 2006), LNCS, Springer (to appear). [Online]. Available: www.cl.cam.ac.uk/~rja14/Papers/Phish-and-Chips.pdf [Sep. 15, 2011].

[15] Drimer, S., Murdoch, S.J., Anderson, R." Optimised to fail: Card readers for online banking". In: Dingledine, R., Golle, P. (eds.) FC 2009. LNCS, vol. 5628, pp. 184–200. Springer, Heidelberg (2009) [Online]. Available: www.cl.cam.ac.uk/~sjm217/papers/fc09optimised.pdf [Sep.16, 2011].

[16] Steven J. Murdoch and Ross Anderson., "Verified by Visa and MasterCard SecureCode: or, How Not to Design Authentication", 2010. [Online]. Available: www.cl.cam.ac.uk/~rja14/Papers/fc10vbvsecurecode.pdf [Sep. 18, 2011].

[17] Ross Anderson, Mike Bond, Omar Choudary, Steven J. Murdoch, Frank Stajan. "Might Financial Cryptography Kill Financial Innovation? - The Curious Case of EMV", 2010 [Online]. Available: www.cl.cam.ac.uk/~osc22/docs/fc11_p2pemv.pdf [Sep. 17, 2011].

[18] Richard J. Sulliva., "Can Smart Cards Reduce Payments Fraud and Identity Theft?" [Online]. Available: www.kansascityfed.org/PUBLICAT/econrev/pdf/3q08sullivan.pdf [Jan. 12, 2012].

[19] L. Padilla, "Breaking Visa PIN" [Online]. Available: www.gae.ucm.es/~padilla/extrawork/visapvv.html [Nov. 19, 2011].

[20] Toni Merschen, Fraud dynamics in the card payments industry: A global review of the realities of EMV deployment, In: Journal of Payments strategy & systems, Bd. 4 (2010), 2, S.156-169. January, 2010.