White Paper

Combating Fraud In The Retailer Payment Environment

January 2010

Prepared by

Maria Arminio

AVENUE B Consulting Inc.

With contributing editor
Paul Reimer

WHITE PAPER ON COMBATING FRAUD IN THE RETAILER PAYMENT ENVIRONMENT

TRANSACTION PROCESSING FRAUD IS AN EVER INCREASING PROBLEM.

Over the last several years data breaches and compromises of large credit card processors and retailers have had devastating consequences to the retail payments industry. In the United States, breaches have compromised the "sensitive" data of millions of card accounts, potentially leaving consumers exposed to identity theft. The hacking technology and the professional criminals that perpetrate the hacks are evolving faster than the standards are being improved or subscribed to by retailers and processors. The situation has become so bad that card issuers of all sizes routinely block and reissue large percentages of their credit and debit card portfolios on a daily basis to the point where it has now become business as usual to do so.

As a result, many stakeholders in the industry including retailers, acquirers, processors and issuers are questioning whether the Payment Card Industry (PCI) Data Security Standard (DSS) adequately addresses the security challenges facing the industry. These standards may help prevent cardholder data breaches, but certainly have not proven to be a panacea for fraud prevention. Because of this, industry players are balking at the exorbitant costs associated with maintaining compliance with the PCI-DSS annual audit requirements.

Even those retailers that have not been victim to security breaches are confronted with the economic realities of the growing cost of fraud. According to a 2009 survey conducted by LexusNexus, merchants are paying over $100 billion in fraud losses due to unauthorized transaction and fees and interest associated with chargebacks. Adding the cost of lost and stolen goods, the U.S. industry-wide fraud losses jump to a total of $191 billion.[1] Moreover, retailers are facing increasing pressure to cut costs and maximize return on investment in the current economic environment. This makes the decision to implement security measures to combat fraud a challenging one, particularly when major changes to the payment infrastructure are required.

Large retailers and major acquiring processors are clamoring for methods and approaches beyond those provided by PCI-DSS compliance to minimize and reduce their exposure to data breaches and compromises, while simultaneously reducing the costs of maintaining compliance with PCI-DSS standards. This white paper addresses the emerging security measures that retailers could employ to effectively thwart fraudulent transactions and provide consumers with safety and soundness in the payment processing environment.

BACKGROUND: EVOLUTION OF SECURITY FOR THE RETAILER

Since electronic transaction processing achieved widespread market acceptance in early 1980's, retailers have needed to employ fraud prevention and detection security measures. Mitigation strategies included performing online authorizations, coupled with the use of neural network and rules-based fraud detection tools, to track potential fraudulent transaction activity. Concurrently retailers fortified their enterprise payments infrastructure with advanced encryption capabilities for PIN debit including hardware-based PIN debit encryption, host security modules with key management and zone encryption, PIN debit dynamic key exchange using Master Key-Session Key and DUKPT, and use of Triple DES.

[1] 2009 LexisNexis True Cost of Fraud Study.

This trend for employing additional security measures accelerated with the advent of the Internet in the early to mid 1990's when major retailers and card processors by necessity began adding web facing interfaces to their enterprise payment processing technology platform. Secure Sockets Layer (SSL) and Virtual Private Network (VPN) connectivity have been widely employed for POS and ATM applications whose interfaces were enabled through the Internet thereby providing added transaction security for the transaction data. Additional security measures such as Verified by Visa and MasterCard Secure Code were implemented by bank issuers. However, consumer acceptance has been tepid. Still these measures have proven partially effective at checking the growth of major compromises to credit and debit card data.

To counteract the growing threat of compromises and successful hacks to private and proprietary computer networks that process credit and debit card transactions, the card processing networks and major card brands have established the Payment Card Industry (PCI) Data Security Standard (DSS). PCI-DSS focuses on card data and information security relating to account number storage and protection practices. The standards are designed to bolster the security capabilities of transaction processing networks. Nevertheless, achieving "PCI-DSS compliance" does not assure the retailer or processor adequate protection from a major compromise. Almost without exception the companies that have had major security breaches had received a clean bill of health from a PCI Qualified Security Assessor firm months before the compromise took place.

Despite all of the security measures that have been widely implemented in the payments industry, the occurrences of data breaches keep growing in size and frequency. To this end, security breaches have resulted in increased demand for tougher standards, and have ushered in a new wave of security measures.

LATEST WAVE OF SOLUTIONS

In the last several years new solutions have been implemented by retailers and processors that extend beyond those required by PCI-DSS that enhance the level of transaction and data security, and protection for magnetic stripe card data. Some of the new solutions which have started to gain the most traction in the marketplace include:

1. **End-to-End Encryption**

   Gaining new credence as one of the industry best practices for protecting payment card data, end-to-end encryption protects customer data from the first point of contact at the POS to the third party processor, and potentially through the payment network to the final destination at the card issuer.

   With end-to-end encryption, the card account number and magnetic stripe data are captured and encrypted at the first point of entry (i.e., magnetic-stripe reader head or smart-card reader contacts), in a tamper evident security module or in an independent software crypto module. Triple DES (or AES) is used as the cryptographic standard for securing the confidentiality and integrity of sensitive data and PIN security, coupled with dynamic key management (or DUKPT). Additionally, some transaction data is exposed, supporting "partial clear text" card data for POS level functions, e.g., 4-6 digit BIN routing, last 4 digits of PAN for receipt printing. The encrypted payment card data thwarts external "skimming" or "data tapping" attacks.

   While numerous iterations of end-to-end encryption implementations exist and have been employed in the payments industry over the last decade, the historical implementations of end-to-end encryption have been limited to the retailer and the retailer's processor. To address this

shortcoming, the ANSI Standards Committee X9 work group is currently developing an end-to-end encryption standard that extends from the POS or ATM device all the way out to the card issuers.

Even if end-to-end encryption were supported by merchants, processors and issuers, this focused security effort alone could not stop fraud entirely. Management of keys and the cryptographic process can be difficult. And, while encrypted data travels securely through the payment process using end-to-end encryption, security breaches could still occur if the transaction data is compromised *prior to* transmission or after the transmission has been completed. For example, in the case of skimming where counterfeit cards are created, end-to-end encryption would securely transmit the data on the card, but the data being read and transmitted would still be fraudulent.

2. Tokenization

Tokenization has been gaining popularity with both large retailers and small merchants alike. With tokenization an encrypted or random value (a "token") replaces the card number (PAN) or the magnetic stripe track data in an electronic transaction. Tokens are most commonly used in lieu of storing the card number in a transaction database, and they are also used "in transit" where the card data token is contained in a transaction message that is sent between two end-points. The token then becomes the reference number representing the card number, so all tokens can be referenced back to the original card number. Tokenization greatly reduces the possibility of the theft of actual credit card numbers because the account numbers are stored only in the dedicated tokenization database, and not in the other payments processing platforms used by retailers.

Tokenization is most commonly deployed using Format Preserving Encryption (FPE). FPE preserves the length and formatting characteristics of the token in alignment with the data element associated with storage of the card data, thereby overlaying it with the encrypted token data. Some digits of token data are commonly left unencrypted (i.e., in the clear) in order to facilitate BIN routing (e.g., first six digits) or for research purposes (last 4 digits). FPE typically permits a Luhn Check (mod 10 checksum) to be utilized in the tokenized card number.

Retailers have embraced tokenization because its use significantly reduces the scope and hence the cost of PCI-DSS compliance. Implementation of a tokenization solution by a large retailer is typically a major project initiative requiring a dedicated project team, and the procurement of a commercially available tokenization engine. With proper planning in the implementation phase of an enterprise level tokenization solution, retailers can reduce the scope of PCI- DSS compliance audits by 50-75 percent.

Since data is not stored or sent in its actual form, tokenization provides merchants with an added layer of security for transaction processing. However, this does not address all of the data used by the merchant in transaction processing, and therefore must be used in combination with other fraud solutions to stop fraud completely.

3. Enhanced Authentication Techniques

While transaction security has been strengthened with the use of encryption, security breaches due to stolen data are still on the rise. These types of breaches cannot be completely protected by encryption. As such, retailers are beginning to use authentication to prevent stolen data from

being used, as it is much more powerful than encryption alone in protecting cardholder data from a sniffing or skimming breach.

In 2005, the FFIEC (Federal Financial Institutions Examination Council) issued guidelines on security, endorsing authentication and identity management solutions for financial transactions. These guidelines mandate the use of two-factor authentication, a process in which the user provides at least two independent means of identification among these options: something you have (such as a card), something you know (such as a PIN), and something you are (such as a fingerprint). With each additional factor that is used, authentication becomes more reliable.

Authentication solutions can be either static or dynamic, although the latter is significantly more secure. With static authentication the same credential data is used for validation, whereas dynamic authentication uses different credential data for each authorization, and the credential used is typically specific to the transaction being performed. Dynamic authentication solutions provide added protection against counterfeit cards and skimming.

To date only a limited number of retailers and processors use any form of authentication. Ultimately the best solutions will use multi-factor authentication and dynamic authentication, providing the most protection from unauthorized individuals compromising the payment transaction. Here are some of the competing authentication technologies available in the market today.

- **Security tokens** (e.g., one-time password tokens; USB tokens, display cards, or software-based tokens) generate a one-time password in a token device (like a mobile phone) and use an algorithm that only the authenticator knows. Security tokens that use hardware encryption devices (such as card readers) leverage a familiar form factor, and offer the most robust encryption, but adoption and fulfillment (i.e., getting handheld devices in the hands of consumers) remain a challenge. Software tokens are easier to work with and interface to, but they are less secure because they are prone to malware such as key loggers.
- **Knowledge-based authentication** is typically performed using a password and challenge responses, and site key. In recent years, this authentication method has become more prevalent in online banking programs, but there are some shortcomings. First, knowledge-based authentication often is implemented as single factor, e.g., something you know. Adoption can be difficult as some consumers have problems remembering the answers to the challenge questions. And, with so many online accounts using challenge questions for authentication, the answers to these questions are now becoming overused thereby diluting their inherent secrecy. Also it has been demonstrated that consumers may be redirected to a fraudulent site that may not contain the picture image or site key. Not realizing they have been spoofed, unsuspecting consumers enter user names and passwords anyway, defeating the security.
- **EMV/Chip cards** have only gained traction in the U.S. in closed-loop environments. Chip cards using PINs provide a high level of security by combining secure cryptograms with dynamic transaction data, each time creating a unique and therefore highly secure authorization value. Keys need to be systemically generated and managed in a chip card program. Recently there have been reported incidences of hacked chip cards, which suggests that increasing level of cryptographic security may be needed for the next generation of chip cards.

We are beginning to see the deployment of contactless cards in public transit and merchant locations with low dollar average ticket size. Contactless cards use a radio frequency identification (RFID) chip and some use dynamic CVV (Card Verification Value) cycling. However, the small amount of memory and power available on an RFID chip limits the size of the encryption algorithm that can be used and hence, the level of security supported. Moreover, encryption does not occur until the data is tapped into the contactless reader so additional efforts are needed to further secure the environment from sniffing or breaches that occur at or before the card reader.

Adoption of both chip cards and contactless cards has been slow in the U.S. as the card issuers have spent millions to promote the payment infrastructure based on magnetic stripe. While there has been considerable discussion by the major card issuers about moving to chip cards in the U.S., the cost to the industry to change the payments infrastructure in terms of card (re)issuance and back-end changes to implement a chip card program is estimated at $25 billion[2], and that does not include retrofitting the terminals to read the chip.

- **Magnetic Stripe Unique Profiling** offers a highly reliable method of card authentication. This dynamic card authentication technology is based on the unique physical properties of the magnetic stripe that appear naturally on each magnetic stripe card as a byproduct of the manufacturing process. It provides validation that the card itself is genuine and that its encoded data has not been altered.

  This solution can be implemented at low price point compared to other authentication solutions in the market. Since existing magnetic stripe cards contain this unique authentication technology in their inherent state, there is no need to reissue cards to consumers. However the cards must be registered. The card reader technology tied to this solution is now sufficiently advanced to encrypt the magnetic stripe card data at the reader head, providing added security. Retailers can readily upgrade their POS technology as part of the routine device upgrade/replacement cycle.

- Among the dynamic authentication solutions available in the market, magnetic stripe unique profiling best leverages the existing payment infrastructure and minimizes cost expenditures to the retailer. The solution does require a working agreement between merchants, acquirer processors, and cards issuers before the benefits are realized, and this has lagged in the market place. However, the automatic card registration process may help to circumvent this stumbling block in the future.

- **Out-of-band authentication** uses a secondary channel and different medium to communicate to the user. Out-of-band techniques (delivered via email or SMS text message to the mobile phone) have emerged to track near real-time monitoring of card misuse. This method of authentication has been quite popular in online and mobile banking programs, but is still in the nascent stages of development and can be cumbersome for the consumer and more time consuming at POS.

- **IP Geolocation** leverages mobile phone technology by comparing the user's current location (identified by satellite) to that previously registered by the user. Two factor authentication is supported, e.g., the consumer's cell phone and physical location.

[2] Reported at Visa Security Summit, 2009.

In order for any of these authentication methods to be broadly successful, they must be must be easy to use, efficient and cost effective. Historically, chip and PIN is viewed as a superior authentication solution because of its use of dynamic authentication, but this technology has never gained widespread acceptance in the U.S. Contactless cards have been somewhat effective in combating counterfeiting, but lack the security of EMVchip and PIN schemes. Plus, card issuance and device retrofitting are significant added cost expenditures.

Magnetic stripe unique profiling holds tremendous promise given that this solution provides dynamic authentication, uses a form factor that is familiar to users, and is readily available in the market. Additionally, magnetic stripe unique profiling best leverages the existing payment infrastructure by minimizing retrofit requirements and eliminating the need for card reissuance.

Viable hybrid solutions are also emerging in the market place. One-time passwords that are delivered through an out-of-band channel provide the benefits of both two-channel and two-factor authentication.

All of these authentication methods require some type of registration process and/or issuance process. Even though the card issuer stands to benefit, the cost burden lies with the retailer.

4. Dynamic Transaction Authentication

While much focus has been placed on dynamic authentication of the card in the payments environment, the pundits of the latest advancements in transaction security endorse the dynamic authentication of all elements of the transaction including the user, user's card, the data on the user's card, the terminal or device, the network switches and host computers of the data recipients and the transaction details. This approach ensures that the transaction is secure not only from the first point of entry at the terminal and across the payment infrastructure, but also makes certain that the card itself and the data on the card are not altered.

The need for dynamic transaction authentication has arisen because end-to-end encryption alone cannot protect retailers from breaches due to skimming or sniffing. Dynamic transaction authentication provides retailers with a multi-layered solution for securing each element of the payment transaction. It leverages a combination of strong encryption, secure tokenization, counterfeit detection, tamper recognition, data relevance and integrity, and dynamic digital transaction signatures - which together validate and protect the entire transaction and each of its components.

In today's payment environment, there are only two ways to support dynamic data in the transaction authentication process – either using chip and PIN or magnetic stripe unique profiling.

- In the U.S., the Chip and PIN solutions require the retailers, processors and issuers to make major infrastructure changes, including issuing chip cards, swapping out POS terminals and implementing significant changes to the back-end infrastructure. Additionally, chip cards rely on a key management process which is systemically generated and thus can be cracked or compromised.
- Magnetic stripe unique profiling leverages the existing payments infrastructure. Magnetic stripe cards do not need to be re-issued, rather the cards are registered "on the fly" as they are captured at the POS card reader device, so the process is transparent to the cardholder. The card reader technology has already been deployed in over 150,000 POS terminals in the U.S. market. Additionally, there is no key to protect with magnetic stripe unique profiling, and the security architecture appears naturally in the micro-particles on the card's stripe.

In October 2009, The Smart Card Alliance published a white paper on the use of chip technology to impact fraud.[3] In that document the authors presented a table comparing U.S. contactless payments security features with EMV and existing magnetic stripe infrastructure. In light of the developments with magnetic stripe unique profiling, we thought a revised comparison of the strengths and weaknesses of these authentication capabilities would be beneficial. Table 1 compares the security features provided by U.S. contactless payments with EMV payment cards and the magnetic stripe unique profiling infrastructure and the benefits of each.

Table 1. Comparison of U.S. Contactless Payments Security Features with EMV and Magnetic Stripe Unique Profiling

U.S. Contactless Payments Security Feature/Behavior	Comparison with EMV Implementation	Comparison with Magnetic Stripe Unique Profiling Infrastructure	Benefits of Contactless Payments Security vs. Magnetic Stripe Unique Profiling
Cardholder typically maintains possession of a contactless payment card and taps the card on the reader, never relinquishing the card to a sales clerk.	EMV contact chip card is inserted into the reader slot by cardholder or handed to a sales clerk. Cardholder retains possession of contactless EMV chip cards and taps the card on a reader.	Magnetic stripe card is swiped by consumers in a multi-lane retailer or is inserted in a gas pump or ATM. Magnetic stripe unique profile eliminates the potential for skimming because the unique properties of the card cannot be duplicated.	Security measures are comparable—no risk of skimming data. EMV cards and magnetic stripe unique profiling are more secure than contactless cards because the reader head on the terminal has direct contact with the chip or magnetic stripe when read. The contactless chip contains an antenna that allows that chip to communicate with the reader through radio frequency. The communication can be intercepted by hackers depending on level of encryption used.
Card is based on highly secure smart chip technology. Contactless chip card is extremely difficult to counterfeit.	Card is based on highly secure smart chip technology. EMV chip card is extremely difficult to counterfeit.	Magnetic stripe unique profiling data cannot be compromised as the micro-particles on each card are unique.	Chip and contactless cards are extremely difficult to counterfeit, but magnetic stripe unique profiling cannot be counterfeited.
Contactless card produces unique data for every transaction that is a function of a secret key resident on the card and placed there by the card issuer.	EMV chip card transaction produces a unique transaction code that does not allow reuse or replay of the transaction.	Magnetic stripe unique profiling cards carry dynamic data, which is inherent in the natural properties of the magnetic stripe. Cards cannot be counterfeited.	Transaction data cannot be reused/replayed for fraudulent transactions in either case. Magnetic stripe unique profiling is considered more secure because the dynamic properties appear in nature, whereas with chip card the key management process can be subject to hacking.
Contactless card allows online card authentication.	EMV chip card allows authentication of the payment card for both online and offline transactions.	Card authentication is supported by registration of the existing magnetic stripe cards. This process can be done "on the fly", as transactions are performed – totally transparent to the user.	Fits well into the U.S. infrastructure where almost all transactions are authorized online. Magnetic stripe profiling leverages from cards that already exist in the market and with existing terminals that have already been deployed to read the unique profiling features.

[3] Fraud in the U.S. Payments Industry: Fraud Mitigation and Prevention Measures in Use and Chip Card Technology Impact on Fraud, October 2009.

BEST BET FOR SUCCESS: LEVERAGE SOLUTIONS THAT EXIST IN THE MARKET PLACE TODAY!

Security breaches continue to occur at a time when cost cutting is paramount. The retailer community recognizes that there is a problem with security. They are willing to play a role in improving payment security and even make reasonable investments (e.g., retrofit POS devices, etc.) in their systems to stay current. They want assurances that this will resolve the problem and is not done just for the sake of compliance.

We can expect that newer versions of PCI-DSS will be aimed at tightening requirements and increasing security. But, retailers should view PCI-DSS as the bare minimum set of requirements. Retailers need to do more than "check off the boxes" of PCI-DSS compliance, and instead develop strong security solutions.

What solutions should be implemented? **We advocate a flexible mix of end-to-end encryption, tokenization and dynamic transaction authentication as the best solution that could be applied to applications, on an application-by-application basis. Retailers may employ two or more different combinations to achieve the right balance of protection, simplicity, and cost effectiveness for each key application.**

PCI-DSS is steering the industry in the direction of end-to-end encryption for sensitive payment data. Retailers will benefit by implementing end-to-end encryption because the PCI-DSS audit becomes less complex and thus less expensive. But security measures such as end-to-end encryption must be expanded and endorsed by all parties in the life cycle of a payment transaction in order to achieve true success. PCI-DSS falls short of this requirement. The implementation of end-to-end encryption by itself will have limited effectiveness unless it is also adopted by card issuers and the payment processors in a coordinated fashion. The retailer operating as a sole practitioner cannot make end-to-end encryption achieve its true potential. Cross industry standards need to be in place and they need to be open, as proprietary solutions will be summarily dismissed.

Tokenization will become more prevalent for smaller merchants, particularly in those environments where card sensitive data resides in a central location. While effective in protecting data from hackers, implementation of a tokenization solution is a complicated process on the enterprise level, but there is a groundswell of support from very large, Tier I retailers.

On a parallel track, retailers need to stay apprised of developments in cardholder authentication, especially the use of dynamic cardholder authentication, and the impact its adoption would have on the need for end-to-end encryption. Authentication will undoubtedly help prevent fraud due to skimming and sniffing, but the business case needs to exist for enhanced authentication to become widespread in acceptance.

Dynamic transaction authentication combined with end-to-end encryption is the best solution we have available today. In the U.S., the most cost effective way for merchants to support this is using magnetic stripe unique profiling.

We cannot forget consumer acceptance. Security solutions must protect the consumer *before* a compromise occurs. And, they must be easy to use. Those leveraging form factors that are familiar to the consumer will have a better chance of success than new technologies with limited adoption.

Maria Arminio, President and CEO of Avenue B Consulting, is thirty year veteran in the payments consulting business specializing in transaction processing and risk management solutions for merchants and acquirer processors. Paul Reimer, President of Clearkey Consulting, is a technology strategist in the payments industry and a PCI DSS compliance-auditor.