UNIVERSITY OF CAMBRIDGE

Computer Laboratory
Security Group

Chip & PIN (EMV) relay attacks

by Saar Drimer and Steven J. Murdoch

Executive summary

This article discusses the *relay attack* on Chip & PIN which could be exploited by criminals to perform fraudulent transactions, using a card with a fake chip. We describe the situations in which this fraud could be perpetrated and suggest ways to mitigate the risk. Chip & PIN currently does not defend against this attack, despite assertions from the banking community that customers must be liable for frauds in which the PIN was used. We thus propose measure to detect, and prevent such attacks in the future.

Background

EMV (named after its founders Europay, Mastercard and Visa) is the standard on how smartcards used for payment communicate with the terminal in shops. In the UK, the system based around EMV is known as *Chip & PIN*.

Chip & PIN is intended to reduce fraud by requiring the genuine card and matching PIN be presented for a successful transaction. The process starts by the terminal sending the card a random number, known as a *challenge*. The customer then enters their PIN into the terminal and it is sent to the card. The card computes a cryptographic *response*, which incorporates the challenge, whether the PIN was entered correctly, and a secret known only to the card and the bank which issued it (the terminal does not know this secret). The purpose of including the challenge is so that the terminal can detect whether an old response is being replayed. This response is sent back to the terminal which then *goes on-line* and sends the challenge and response to the bank, who will verify them.

Let us consider some potential scenarios of fraud which Chip & PIN is intended to protect against:

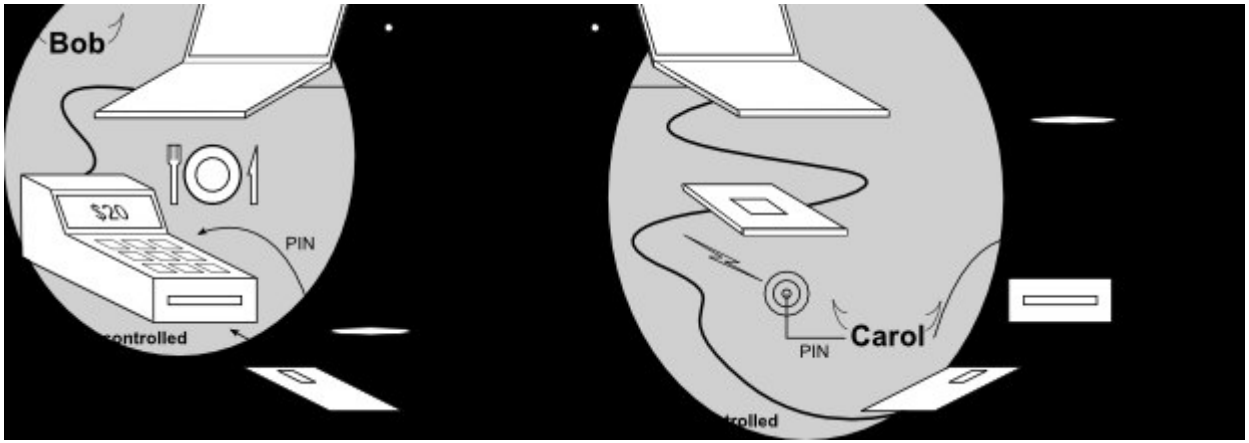
Stolen card. Without the correct PIN being entered, the card will not produce the correct response, and so cannot be used in an *on-line* Chip & PIN transaction.

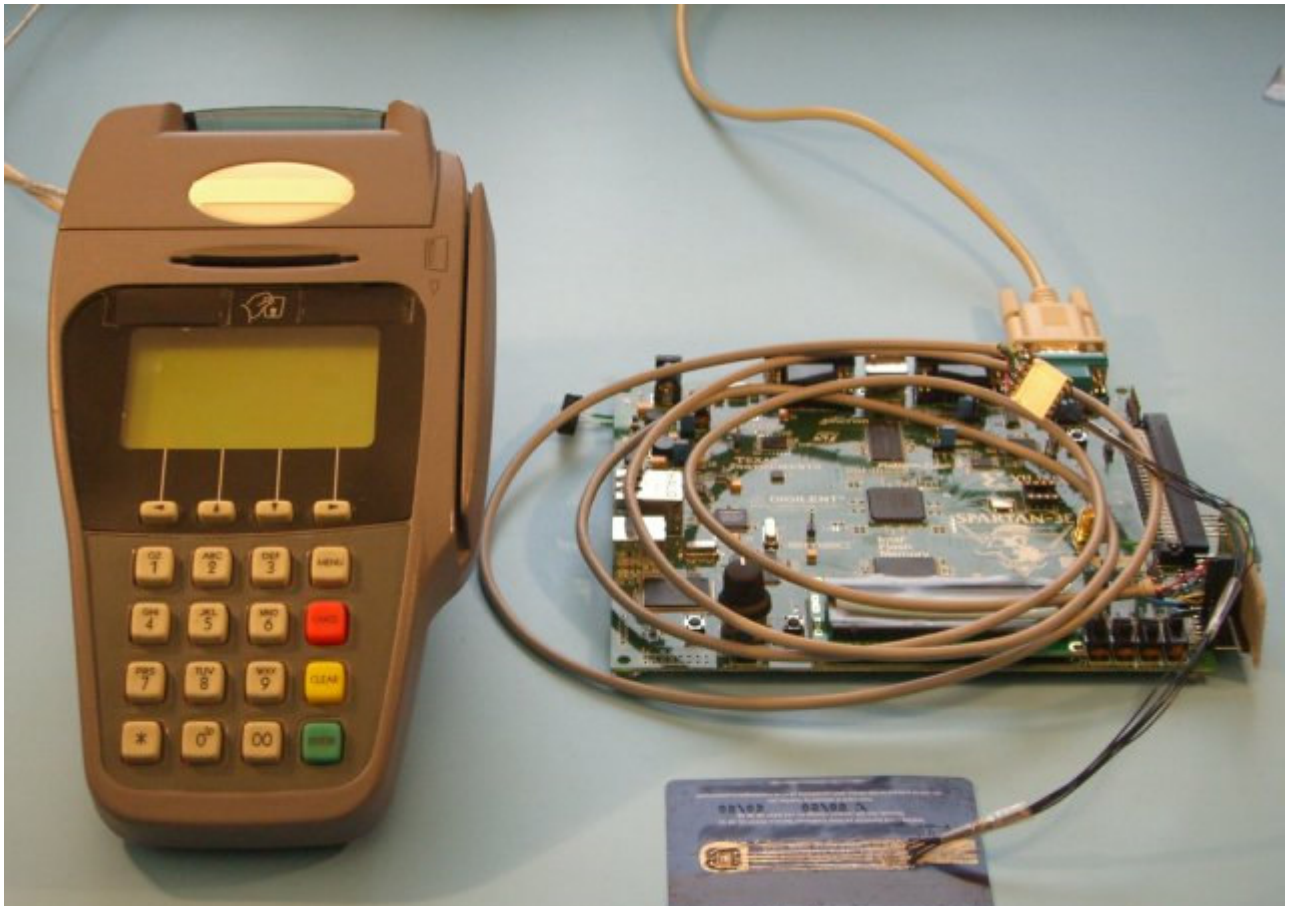
Observed PIN. Without the card, a fraudster who knows the PIN will find it difficult to produce a fake card which will compute the correct response.

Stolen card and its PIN observed. Now the fraudster can use the card and PIN to produce a valid response and use it as though he is the rightful owner. The account holder, however, will eventually notice the fraudulent transactions and promptly contact his bank, causing the card to be cancelled.

Observed response. If the fraudster knows the PIN (or persuades the customer to enter it) and gets temporary access to the card, the card will produce correct responses. These response, however, cannot be used later as the challenges from legitimate terminals are meant to be unpredictable.

Despite these protections, vulnerabilities remain. For example, when customers pay with a Chip and PIN card, they have no choice but to trust the terminal when it displays the amount of the transaction. The terminal, however, could be replaced with a malicious one, without showing any outward traces. When the customer pays for a low-value product and enters the PIN into the terminal, the challenge from a different shop selling a far more expensive product could be *relayed* to the card. The PIN and response from the card could likewise be relayed back to the other shop, which will accept the transaction.


Bob
$20
PIN
Carol
PIN



We will illustrate the relay attack with a description of a real life scenario. For that we will introduce the characters:

- **Alice** is the innocent customer who is about to be defrauded from her savings.
- **Bob** is a crook. He is now employed as a restaurant waiter.
- **Carol** is Bob's accomplice who is loitering at a jewellers shop waiting for Bob's signal to participate in the attack. She is carrying all the equipment needed for the attack in her backpack.
- **Dave** is an honest merchant who operates a jewelry shop and is not associated with Alice, Bob, or Carol.

Alice, our innocent customer, is about to pay $20 for a meal in a restaurant. Unbeknownst to her, Bob, the crook waiter, presents her with a terminal which looks and behaves like a real Chip & PIN terminal, but is secretly relaying a transaction with Alice's card to Bob's partner in crime – Carol. This is done by using a hidden laptop hidden behind the counter.

Carol is in Dave's jewellers shop about to buy a $2000 diamond. Just as Alice inserts her card into the restaurant's terminal, Carol is notified via a radio link or SMS message to insert her specially modified card into the jewellers shop's reader.

As Alice keys her PIN, it is read out to a earphone worn by Carol. All communication from the jewellers shop terminal will be sent through Carol's card and Bob's terminal to Alice's card, and vice versa. Dave will see that the transaction has succeeded and will hand Carol the diamonds that will be charged to Alice's account. Alice leaves the restaurant thinking she paid $20 for a meal, while her statement will show $2000 for a diamond. As the malicious terminal will never communicate with the bank, Alice will not be changed for the meal.

Alice will have used a terminal which looks perfectly normal and it will have shown $20 on the display. Dave will see that the transaction went through without any problem. The bank will see that Alice's card appeared to have been used with the correct PIN. Carol and Bob, however, have walked away with a $2000 diamond without paying for it.

A demonstration of this attack was featured by BBC Watchdog on 6 February 2007. A video of the segment is available.

Questions and answers

Why does this attack work?

Customers insert their card into a terminal without any way to verify that it has not been tampered with. They depend on the integrity of the terminal to:

- Display the correct value of the transaction
- Keep the PIN secret
- Use the card only for transactions approved by the customer

Merchants accept customer's cards without verifying that they are genuine. In some cases they may even not handle the card and they are encouraged look away while the PIN is being entered. They depend on the security of EMV to detect unauthorised transactions although EMV does not prevent relay attacks such as the one described above.

How can the attack be prevented?

Merchants can try to identify fake cards by taking them from customers, checking the counterfeit detection features (such as the hologram and embossing) then inserting them into the card reader themselves. This will require the relay card be wireless rather than the wired prototype we have developed. A further protection is for the merchant, before handing over the goods, to confirm that the account number on the receipt matches the one on the card. This would require the attacker know, in advance, the account number of the customer whose card they are about to exploit, making it substantially harder to perpetrate.

Banks could deploy measures to detect such relay attacks. In our academic paper, in USENIX Security 2007, we describe a technique which will allow a terminal to measure how far away the genuine card is. This design, a so-called *distance bounding* protocol, is one example of a secure positioning system. It requires that the terminal measure the time it takes to communicate with the card. Since information cannot travel faster than the speed of light, the maximum distance between card and terminal can be calculated. By carefully designing the communication method cards use, this estimate can be made very accurate and ensure that relay attacks over even short distances (around 10m for our prototype) are detected. This will require modifications to both the cards and terminals and will only be feasible for the longer term.

Banks or third parties could release a device so that the customer avoids entering their PIN into a merchant terminal. If the customer entered their PIN into a keypad they control, a malicious terminal could not record it. Similarly, such a device could also show the value of the transaction and so allow the customers to detect if they are being charged for more than expected. Such device would transfer the trust from the merchant's terminal to a personal device that the customers bring themselves into the transaction. The paper, "The Man-in-the-Middle Defence" by Ross Anderson and Mike Bond, describes this concept in more detail.

Banks could deploy terminals which allow customers to detect if they have been tampered with. Currently, there are hundreds of different designs for EMV terminals and customers cannot tell the the difference between a legitimate or fake ones or if a real terminal has been tampered with. Visible tamper-resistant seals may allow a merchant, if properly trained, to detect such tampering. However, it is unlikely that customers would have the time, skills or patience needed to reliably check a terminal before each use.

Aren't terminals supposed to be tamper-resistant?

Yes, but this doesn't protect the customer. Terminals are designed to stop working *as Chip & PIN terminals* once opened; this will prevent them from communicating with the merchant's bank and placing charges onto a card. To the customer, however, they still appear the same, and as fake terminals do not need to communicate with the bank, the current tamper proofing mechanisms are ineffective for preventing the relay, and other, attacks. Our fake terminal was purchased through eBay and looks the same externally although we have replaced almost all the internal electronics with our own. *We even made our's play Tetris!*

There is further discussion of this question on our page: "Tamper resistance of Chip & PIN (EMV) terminals".

Is it a problem that the terminal used is not approved by the banks

No, because the customer cannot tell the difference. Our prototype uses a terminal which is not approved in the UK. However there are so many different types of terminals currently in use that customers cannot be reasonably expected to remember them all. If the fraudster wished to make an identical terminal, even approved, tamper-resistant terminals can be easily modified, simply by removing electronics as we have done. Alternatively, the fraudster could build a new plastic enclosure identical to those of commercially available Chip & PIN terminals. This option would not be as cheap, but the cost could

be easily recouped by just a few fraudulent transactions.

Will corrupt merchants not be detected?

Perhaps. In our scenario, the only link between the restaurant and the victim is that the victim was *not* charged for the meal. While customers will likely spot the fraudulent transaction when they receive their bank statement, they might not be able to recall the particular missing transaction. By the time the customer notices the fraud, the crooks have already defrauded many others, received the goods, and disappeared. The fraudster doesn't even have to be the owner of the restaurant, just a temporary employee or impersonating a terminal repair technician. The attacker needs only to place a fake terminal and a laptop in the shop.

How long has this style of attack been known?

Relay attacks, also know as the *wormhole* or *Chess grand-master* attacks have been known of since at least 1987. In the context of EMV, we described how relay attacks could be used for fraud in the paper "Chip and Spin". Subsequent to publication, however, there remained some skepticism within industry as to whether these attacks were implementable. We set out to confirm that they were indeed plausible attacks and estimate their costs.

How far away can the fake card be from the genuine card?

It can be anywhere in the world. While testing our prototype we introduced a 2 second delay in all communications and the transaction still succeeded without a problem. Within 2 seconds, a radio signal could travel around the world 10 times. So we expect that the relay attack will succeed regardless of where the genuine and fake cards are, as long as there is reliable Internet or phone coverage.

How does the relay attack compare to card cloning?

Disadvantages: the genuine card must be in the malicious terminal for the full duration of the fraudulent transaction. This means the attacker must coordinate the scam quite carefully to avoid the victim's transaction seeming to take an inordinate amount of time. The fake card in the relay attack, as we have implemented it, has a wired connection to a laptop, so it will not pass inspection by a merchant. However, while a fully wireless version is more difficult, we think it would be within the means of reasonable attacker.

Advantage: Defences against more conventional attacks do not affect the relay attack. Card cloning attacks depend on either the transaction being performed offline, or being able to force the terminal or ATM to fall back to magnetic stripe verification. These weaknesses are gradually being phased out. Another EMV technology, *Dynamic Data Authentication (DDA)* will prevent card-cloning even in offline transactions, but has not yet been deployed in the UK, who still use *Static Data Authentication (SDA)*. None of these defences will be effective in preventing the relay attack.

Aren't transactions encrypted?

Sometimes, but this does not prevent our attack. As demonstrated by our Chip & PIN interceptor, SDA transactions are not encrypted between the terminal and card. Even if they were, as they can be with DDA, relay attacks still work since encryption would only prevent use *reading* or *modifying* the data, whereas we simply relay that data blindly.

How much does it cost to implement this attack?

Not much. We have acquired an EMV terminal off eBay for $50. The built-in electronics were removed and replaced by an FPGA board costing $200 while

we used the original LCD panel and keypad. Two laptops are needed with the sole requirement that they have USB or serial ports and can transmit or receive wireless communication. Creation of the modified card requires some mechanical skill and a processor for communication with the PC. In our implementation we used an FPGA board that costs less than $150. The PC software and FPGA code were written in Python and Verilog, respectively and require moderate technical skill.

Why are you doing this work?

To help reduce fraud. We think it is valuable to understand an important weakness in Chip & PIN so that appropriate defences can be deployed. By building a prototype we gain practical experience of what fraudsters must do to carry out these attacks. This information assists the development of practical, preemptive security measures which the banks could deploy to resist the relay attack. Also, when discussing this problem with members of the industry, we were met with skepticism as they believed the timing constraints would be too tight for the attack to be possible. So we set out to build a practical demonstration that this attack was both feasible and affordable. Finally, should criminals use this vulnerability to defraud customers, we would like the victims to be able to defend themselves against claims that the fraud could not have happened because banks claim Chip & PIN is infallible.

What is the purpose of the relay demonstration equipment?

*Although this equipment might be adapted for use in the commission of fraud, it is **not** designed for such use* nor do we intend for it to be used for fraud. Doing so would be unlawful under the UK Fraud Act 2006. Instead, our prototype is solely made to identify the vulnerabilities in the Chip & PIN system as well as to help design and evaluate defences.

In what other scenarios could the relay attack be applicable?

Cross-border card theft. Currently, if cards are stolen in one country, they are often used in a different one. This complicates criminal investigations and circumvents some fraud protection measures, so is desirable to fraudsters. As magnetic stripe cards are still widely accepted, the criminals can read the details off the stolen card and email them overseas where they will be written onto blank cards for use.

As Chip & PIN becomes more popular, magnetic stripe cards will be accepted in fewer locations. Because the chip cannot be trivially copied (unlikely the magnetic stripe), fraudsters would need to physically send the card overseas to use it. This would take time and allow the legitimate card holder to cancel the card. Instead, fraudsters could use a relay attack to instantly use a Chip & PIN card abroad, without it ever having to physically leave the country.

Please send any comments on this page to Steven J. Murdoch
Last modified 2008-03-21 19:28 by Steven Murdoch