RFID and the Middleman

Ross Anderson

Cambridge University
http://www.ross-anderson.com

Abstract. Existing bank-card payment systems, such as EMV, have two serious vulnerabilities: the user does not have a trustworthy interface, and the protocols are vulnerable in a number of ways to man-in-the-middle attacks. Moving to RFID payments may, on the one hand, let bank customers use their mobile phones to make payments, which will go a fair way towards fixing the interface problem; on the other hand, protocol vulnerabilities may become worse. By 2011 the NFC vendors hope there will be 500,000,000 NFC-enabled mobile phones in the world. If these devices can act as cards or terminals, can be programmed by their users, and can communicate with each other, then they will provide a platform for deploying all manner of protocol attacks. Designing the security protocols to mitigate such attacks may be difficult. First, it will include most of the hot topics of IT policy over the last ten years (from key escrow through DRM to platform trust and accessory control) as subproblems. Second, the incentives may lead the many players to try to dump the liability on each other, leading to overall system security that is equivalent to the weakest link rather than to sum-of-efforts and is thus suboptimal.

1 Introduction

Card payment systems have come under repeated attack from forgers and other fraudsters. The mechanisms used in magnetic-strip cards, and attacks on the ATMs that rely on them, are documented in [2], while the back-end backing systems are described in [3]. Banks in Europe are now moving to smartcards following the EMV standard, branded in the UK and Ireland as 'Chip and PIN'. In the most commonly-deployed variant of this standard, bank cards contain a smartcard chip that will verify a customer PIN against a locally-stored value, and also one or more cryptogram generation keys. These are used to compute a message authentication code (MAC) on transaction data; the MAC is verified by the card-issuing bank. As many countries, including the USA, have not adopted EMV, the cards also have a magnetic strip for fall-back operation. The introduction of EMV has altered the fraud landscape, with a reduction in cardholer-present forged-card losses, coupled with increases in cardholder-not-present fraud and mag-stripe-fallback fraud.

Now that European cardholders are used to entering cards and PINs for all bank card transactions, rather than just ATM transactions, it has become much easier for attackers using false terminals to harvest card and PIN data.

Previously, card forgery attacks typically involved skimmers attached to the front of ATMs; now we see a spate of offences involving wiretapping of the links from EMV terminals to branch server equipment. With frauds reported in France and Italy, both card and PIN data were collected using surreptitiously-installed wiretap devices, which send the data to waiting criminals by wireless [11]. In a series of recent UK cases, card data have also been harvested using an in-store wiretap, but the PINs collected by observation. Many offences have been reported in garages, staffed by Sri Lankan Tamils, leading to ATM withdrawals in India, Malaysia or Thailand; it's reported that the Tamil Tigers use these forgeries for fundraising – presumably intimidating their UK expatriates into collaborating [8, 12].

In addition to such fairly straightforward frauds, it turns out that the EMV protocols and their implementations are vulnerable to various middleperson attacks [1]. In addition to capturing card details and PINs for use in magnetic-strip terminals, villains can do various kinds of man-in-the-middle attack, and the transactions provided by cryptographic hardware security modules to support EMV have flaws that enable bank insiders to extract PINs. The growing number of bank customers who complain that stolen chip-and-PIN cards were used without the PIN possibly having been compromised suggests that at least some banks have corrupt staff who sell PINs to criminals; this was a known modus operandi with magnetic-strip cards in the 1990s [2]. Meanwhile, Murdoch and Drimer have shown that real-time man-in-the-middle attacks are feasible by implementing them [9]; and it's also been noted that a bank customer could use a middleperson device to fix the protocol by providing a trustworthy user interface – she could observe the actual data traffic between the card and the terminal, rather than the possibly false data displayed by the terminal [5].

2 Implications for RFID

The bank card industry in the USA and Japan is introducing RFID credit cards, described by Heydt-Benjamin et al. in [7]. Although the specifications are confidential, protocol dumps suggest that the mechanisms are very similar to EMV, although with the crypto (the MAC) missing. The authors of that paper showed that cloning and relay attacks work in principle by scanning a credit card in its sealed delivery envelope and making a purchase with it.

Meanwhile, the mobile-phone industry is introducing Near-Field Communications (NFC), whereby a phone acquires the capability to act as either an RFID device or an RFID terminal under program control. The NFC protocols themselves provide an abstraction layer between the four different RF interfaces already deployed, thus providing a clean interface for the software developer. The NFC Forum envisages about half a billion NFC-enabled handsets in use by 2011 [10].

This technology holds out the prospect of solving the problem of a trustworthy user interface. The plan is that instead of being a relatively dumb device, your credit card will be an application on your mobile phone. You bring your

mobile into close proximity with the merchant terminal, an application displays the sale amount, you authorise this, and the transaction goes through. It may also bring further security benefits, such as a single point of revocation [6].

However, without protocols giving robust protection against man-in-the-middle attacks, the trust that might be placed in this interface will be largely illusory. Worse, NFC is likely to make middleperson attacks much easier. At present, such an attack requires the construction of custom hardware; in future, an attack could be carried out by software installed on commodity mobile phones. One phone could act as a rogue merchant terminal to the cardholder, and communicate with another that acts as a card to a merchant elsewhere.

Transaction forwarding does have 'honest' uses. (Once when I stayed at a hotel in Malaŵi I found that the bill appeared on my credit card via a South African merchant in Rands – an obvious attempt to collect a slightly harder currency, which cost me no more.) However, most forwarding is likely to be objectionable. Someone thinking she's buying an hour on a parking meter in Baltimore for $2 might find out, when the credit card bill arrives, that the bank thinks she paid $2000 for casino chips in Macau.

Limiting transactions without cryptography to low values – as some suggest – won't solve all the problems. For example, a crook might collect large numbers of small payments from passers-by; this might be done by malware installed on the phone of an unwitting victim, which would steal a few dollars from everyone he passes. Theft might not be the only objective; RFID will be used for sports ticketing, so known hooligans who are banned from buying tickets will have every incentive to spoof the system.

If cryptography is introduced, it's not obvious how. The banks will want end-to-end protection, but merchants will want access to transaction data; and the police will want access to records for intelligence/evidence, raising all the old issues about escrow. It's unclear that much more can be done than is already being done with EMV – message authentication codes that might just be upgraded to digital signatures eventually.

It's also not obvious who will design any new protocol suites to support NFC payments. The NFC Forum limits itself to interoperability, and the protection issue are not just about bank payments but transport tickets, sports tickets, supermarket coupons and much else. How many protocol suites will there be? What sort of limits are desirable on relaying / cloning bits? Who is going to specify the protocols? (VISA? Microsoft?) Who will regulate them? (The FTC and the EU's DG Comp? The Federal Reserve and the European Central Bank?)

3 Conclusions

The introduction of RFID payments based on programmable devices such as NFC mobile phones may fix one of the problems underlying bank card fraud – the lack of a trustworthy user interface. However, this may well be at the cost of seriously exacerbating the other main problem – the vulnerability of current payment protocols to various man-in-the-middle attacks.

What's more, the responsibility for security is very widely dispersed, and defenders may hope they can rely on each others' mechanisms. Security economics teaches that we get a much more appropriate level of protection where this results from the sum of defenders' efforts than in the cases where it results from the weakest link – from the least awful of the disparate efforts of a number of possibly uncoordinated defenders [13]. Unfortunately, the security of RFID transactions looks set to become a matter of the weakest link. So a large-scale infrastructure may be deployed that's not only systemically vulnerable to man-in-the-middle attacks, but that actually provides the platform required for these attacks to be carried out easily. I'd suggest that application providers think hard about these issues now; it will be much more expensive later.

References

1. B Adida, M Bond, J Clulow, A Lin, S Murdoch, RJ Anderson and R Rivest,"Phish and Chips", at *Security Protocols Workshop*, Mar 2006; paper at `http://www.ross-anderson.com`
2. RJ Anderson, "Why Cryptosystems Fail", in *Communications of the ACM* v 37 no 11 (Nov 94) pp 32–40
3. RJ Anderson, *'Security Engineering – A Guide to Building Dependable Distributed Systems'* Wiley (March 2001), ISBN 0-471-38922-6
4. RJ Anderson, "Why Information Security is Hard – An Economic Perspective", in *Proceedings of the Seventeenth Computer Security Applications Conference*, IEEE Computer Society Press (2001), pp 358–365; available at `http://www.cl.cam.ac.uk/ftp/users/rja14/econ.pdf`
5. RJ Anderson, M Bond, "The Man-in-the-Middle Defence", at *Security Protocols Workshop*, Mar 2006; paper at `http://www.ross-anderson.com`
6. M Baard, "Will new RFID technology help or hinder security?", 27 Apr 2005, at `http://searchsecurity.techtarget.com/originalContent/0,289142,sid14_gci1083417,00.html`
7. TS Heydt-Benjamin, DV Bailey, K Fu, A Juels, T OHare, "Vulnerabilities in First-Generation RFID-enabled Credit Cards", at *Eleventh International Conference on Financial Cryptography and Data Security* Scarborough, Tobago, February 2007
8. Walter Jayawardhana, "Tamil Tigers suspected of scamming millions in Britain", at `http://lankapage.wordpress.com/2007/01/17/`
9. SJ Murdoch, "Chip & PIN relay attacks", Feb 6th 2007, at `http://www.lightbluetouchpaper.org/`
10. *'Near Field Comunication and the NFC Forum: The Keys to Truly Interoperable Communications'*, 2006, at `www.nfc-forum.org`
11. "Clonavano carte con il bluetooth Scoperta nuova truffa telematica", in *la Repubblica* Sep 4th 2006, at `http://www.repubblica.it/2006/09/sezioni/cronaca/truffa-blue/truffa-blue/truffa-blue.html`
12. K Shoesmith, "Garage Scam funded Terror Group", Hull Daily Mail, Jan 16th 2007, p 1, archived online at `http://www.srilanka-botschaft.de/NEWSupdates_neu/Press_Releases/Press_Pol_Government_Statement_070119bE.htm`
13. H Varian, "System Reliability and Free Riding", available at `http://www.sims.berkeley.edu/resources/affiliates/workshops/econsecurity/econws/49.pdf`