A Practical Relay Attack on ISO 14443 Proximity Cards

Gerhard Hancke

University of Cambridge, Computer Laboratory JJ Thomson Avenue, Cambridge CB3 0FD, UK
ghancke@ieee.org

Abstract. Contactless smart cards are used in access control and payment systems. This paper illustrates an attack which effectively allows an attacker to 'borrow' the victim's card for a short period without requiring physical access to the victim's card. As a result the legitimate owner will remain unaware of the attack. We show that our hardware successfully executed a relay attack against an ISO 14443A contactless smart card, up to a distance of 50 m. Simply relaying information between the card and reader over a longer distance does not require the same technical resources from the attacker as hardware tampering or cryptanalysis. This attack is therefore a feasible method for circumventing current security protocols with little effort. Since application-level measures fail to protect against relay attacks, we discuss possible solutions involving characteristics of the physical communication medium.

Keywords: RFID, contactless smart card, relay attack, distance bounding

1 Introduction

Contactless smart cards have found use in applications such as access control, public transport schemes and other cashless payment systems [1][2]. Due to their nature of use, these cards often have added security capabilities. In this area the Philips Mifare system is the most common, although many manufacturers supply Mifare compatible hardware. Contactless cards use near field communication to communicate, and receive power, over short distances. These devices do not need physical contact with a reader, which simplifies operation and increases transaction speeds. This lack of human interaction has led to fears that this technology could be abused by big business. It is alleged that information about the user could be acquired without consent and that the user's right to privacy would be violated. Consumer groups have therefore campaigned against

[1] Project report, posted January 2005.

the "big brother" potential of similar RFID technology [3]. As more applications start to use contactless technology for transactions of value, the threat from a malicious "little brother" becomes more of an issue. If a contactless card could be read while in a pocket, purse or wallet, a thief might be able to engage in the act of digital pickpocketing while standing next to or merely walking past his victim. This paper shows that, because of the way contactless communication operates over a distance and without user intervention, it is easily possible to execute a successful relay attack in order to impersonate, and therefore gain benefits from, another user.

2 Contactless Smart Cards

For transponders to work they require power, even though the levels are very small. Passive devices operate without an internal battery source, deriving the power to operate from the field generated by the reader. Consequently passive devices offer an unlimited operational lifetime but have shorter read ranges and require a higher-powered reader. Contactless smart cards operate at 13.56 MHz and are further divided into proximity (ISO 14443 [4]) and vicinity (ISO 15693 [5]) devices with nominal operating ranges of up to 10 cm and 1 m, respectively. ISO 14443 specifies A and B operation modes that use different communication and card selection procedures.

This paper concentrates on the ISO 14443A standard which is used with most contactless cards and is compatible with the lower layers of popular commercial products such as Philips Mifare. The standard specifies the operating frequency, modulation and coding schemes (ISO 14443-2), anti-collision routines (ISO 14443-3) and communication protocols (ISO 14443-4). It uses 100% Amplitude Shift Keying (ASK) modulation with Modified Miller coding (106 kbit/s) in reader to card communication. The width of the modulation pulses is only 2–3 μs to ensure continuous power supply to the card. For data transfer from the card to the reader, load modulation with a 847 kHz subcarrier is used. The subcarrier is modulated by using an On/Off Keying (OOK) code with Manchester-coded data (106 kbit/s). The anti-collision protocol is based on the principle of a simple state machine. A card placed in the reader's field powers up and is *IDLE*. Typically the reader periodically sends a *REQA* command, which places all cards into the *READY* state. The card then returns an *ATQA* command so that the reader knows that it has at least one card within range. If multiple cards are present, a binary search tree algorithm is used

to select a specific card. The reader transmits a *SELECT* command, a Number of Valid Bits (*NVB*) parameter, and a bit mask. The number of bits in the mask is indicated by *NVB* and is compared by each card to its ID. If it matches, the card responds with the remainder of its ID. This process is repeated until one card ID is selected ($NVB \geq 64$), after which this card responds with a *SAK* command and becomes *ACTIVE*. The reader and card will then negotiate the protocol parameters used for data communication. The reader will transmit a Request Answer to Select (*RATS*) command that the card answers with an *ATS* command containing its settings. If the card supports Protocol Parameter Selection (*PPS*) the reader can request setting changes through *PPS* requests. The application layer communication protocol is based on the "T = 1" (ISO 7816-3) protocol [6].

3 The Relay Attack

Contactless cards are used as a means of electronic identity in access control and payment systems. Many authentication schemes between the reader and the card is based on the principle of mutual authentication, in accordance with ISO 9798 [7], where both participants confirm that the other party has possession of a shared secret key. After authentication all communications are encrypted. It is assume that even if the authentication is breached, the attacker would still not have access to the secret key and therefore any data received as a result would still be protected. An attacker would have to tamper with the hardware or perform cryptanalysis of the chosen encryption algorithm to recover the key. This might not be feasible for the attacker as this would require substantial technological resources and time with the original card.

A number of factors combine to make a relay attack on a contactless device possible. A contactless card operates over a distance and is activated automatically when close to a reader. An attacker can, therefore, access the card discreetly, without knowledge of its owner, and relay information through a communication link between the card and a remote reader. The reader will assume that the card, and by implication the user, is in close vicinity and provide access to the attacker. Using this attack on the authentication schemes mentioned above the attacker would be able to convince both reader and card that they share a common secret key. The attacker would not be able to view in plaintext any subsequent communications but it is not needed as long as he can continue relaying the respective messages. The attack can be given an active twist by re-

laying the initial authentication sequence after which subsequent data is modified and relayed. As some cards use stream ciphers with only linear error detection codes for integrity checking, a malicious user might easily implement a known plain text attack, which is feasible if the value stored on the card is known.

4 System Implementation

In theory a relay attack is quite straight forward. Our goal was to demonstrate the effectiveness of the proposed attack by developing the hardware to relay data between a card and a reader. The system has two main units connected by a communication link. The Mole interfaces with the user's card and appears as a valid reader. The Proxy appears as a valid card to the reader, and passes instructions to the Mole, who responds with the required information. The simplest implementation will be to perform the required analog modulation/demodulation at each end and then to relay the digitally coded data packets immediately using short range RF communication. An alternative would be to buffer a whole packet before transmitting. The first method is cheaper, easier to implement as it does not require a DSP or microcontroller to encode/decode and store the Modified Miller and Manchester coded data. The second method buffers data packets causing longer delay although this allows for easy data modification and the use of other available communication methods, such as a laptop with Wi-Fi, which should give it a range advantage. The ability to store data also helps during the anti-collision stage as the Mole can read the card's ID and forward it to the Proxy, which can then react to the readers subsequent *REQA* and *SELECT* commands within the specified response time. We decided to design a system that would be cheap to build with simple discrete components. A block diagram outlining the major function in each unit is shown in Figure 1.

4.1 Design Considerations

System Delay Relaying data introduces delay into the system. It is therefore important to consider the timing constraints in the original system. ISO 14443-3 specifies timing requirements to maintain bit synchronization during the anti-collision process. Response time is specified as $(n \times 128 + 84)/f_c$ if the last data bit sent by the reader was '1' and $(n \times 128 + 20)/f_c$ if the last data bit sent was '0'. Response times are calculated using $n = 9$ for *REQA* and *SELECT* commands, and $n \geq 9$ for all

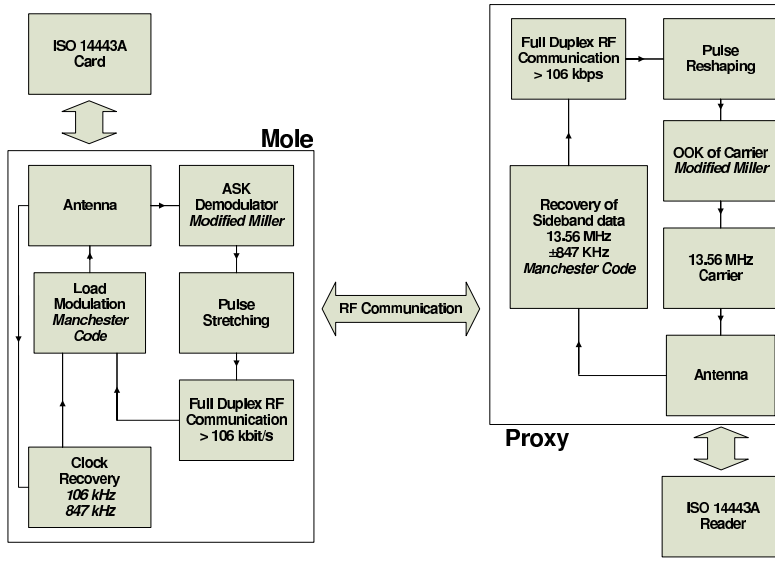


Fig. 1. Functional analysis of the relay system

other commands. The minimum timeout defined for *REQA* commands is $7000 \times f_c \approx 500$ μs. ISO 14443-4 specifies the Frame Waiting Time (FWT) as $(256 \times 16/f_c) \times 2^{FWI}$, where FWI is a value from 0 ($FWT = 300$ μs) to 14 ($FWT = 5$ s) with a default of 4 ($FWT = 4.8$ ms). The FWI value is defined by the card in its *ATS* response. The Frame Waiting Time defines an upper bound, and a modern communication channel should be able to forward a few data frames in 5 s. Therefore the anti-collision timing constraints is more of a concern, especially since the responses are ordered to a bit grid, resulting in the card only answering when the information is due, after 91 μs and 86 μs respectively. An idle reader transmits *REQA* commands periodically and as a result the Proxy held to the reader requests information from the mole with same regularity until the Mole is within range of the targeted card.

Mole The nominal range of ISO 14443 compatible cards is about 10 cm. Under normal operating conditions this means that an attacker's Mole should be within 10 cm of the victim's card. This is possible in a crowded station or a checkout queue but not preferable. The Mole's operating range is determined by distance over which it can power the card, and its ability to receive the card's answer. The range is therefore dependant on the transmitted power, antenna diameter and the Q factor of the coupled antennas [8]. An attacker may not feel bound by the FCC or ETSI limits

imposed on contactless devices and therefore the transmitted power can be increased. The attacker can also use a larger antenna, for example a large loop antenna could be concealed within a briefcase, which can be placed close to a victim. The Q factor could then be tuned for optimal power and data transfer. An attacker should therefore have no problem in increasing the operating range of contactless cards. The operation should be covert and simple as the attacker needs to maintain a few seconds of contact time with the victim's card. A single attack does takes less time but for reliability we should allow time for additional attacks in case the relayed data is interrupted. The only part of the hardware that needs to be covert is the antenna as it needs to be close to the victim without being noticed. The creativeness of the implementation is left to the attacker but an antenna can be built into a briefcase, clothes or even a fake racquet.

Proxy The Proxy's operating range is not a problem as the device can be held in close proximity to the reader. In fact the person executing the attack needs to appear as normal as possible and will probably be holding a wallet or bag against the reader. Once again antenna design is left to the attacker.

4.2 Experimental Hardware

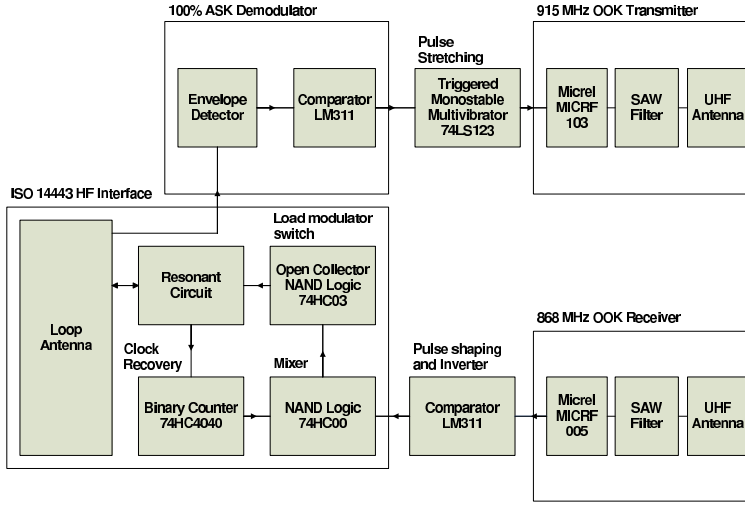


Fig. 2. Hardware diagram of the Proxy

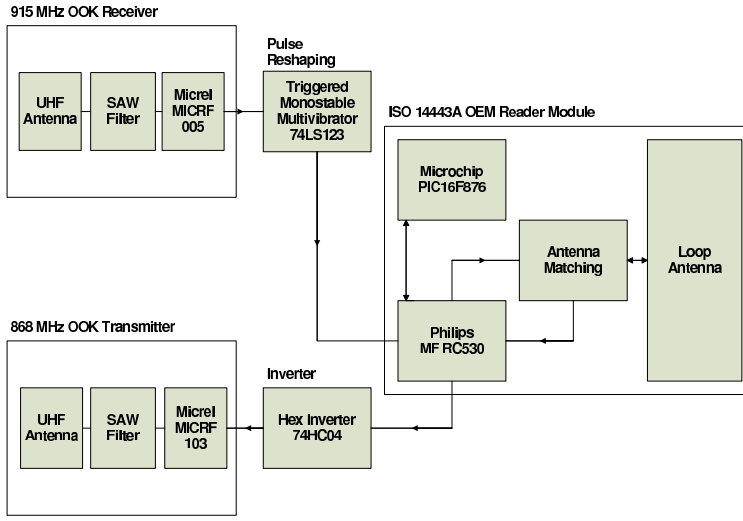
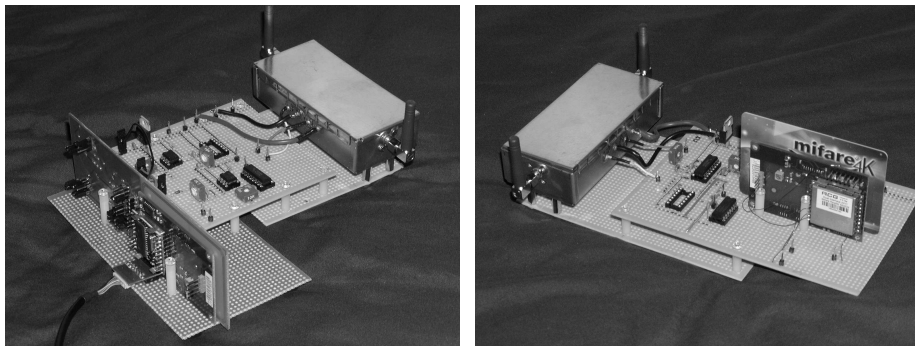


Fig. 3. Hardware diagram of the Mole

The Proxy, as shown in Figure 2, contains an ISO 14443 HF interface, an ASK demodulator, a full-duplex RF communication channel and some signal shaping functions. The HF interface interacts with the attacked system's reader. The resonant circuit is set close to 13.56 MHz for optimal power coupling and data transfer from the reader. Clock recovery is done by using a binary counter that divides the carrier signal frequency by powers of 2. This yields a 847 kHz ($f_c/16$) subcarrier signal, which is mixed with the returning 106 kbit/s Manchester-coded data. The combined signal is then used to control a switch that alters the impedance of the resonating circuit, and as a result carrier's amplitude is varied. The circuit of this component is described in [6, pp 276–278]. It was modified to accept Manchester-coded data as input, and to output the antenna signal to the ASK demodulator. The ASK demodulation is done by detecting the envelope of the 13.56 MHz carrier and deciding on a '1' or '0' using a comparator. The full-duplex communication between the two units is achieved by using two half-duplex RF channels. These channels are implemented using Micrel's QwikRadio UHF ASK/OOK integrated transmitter and receiver ICs. These ICs have a nominal filter bandwidth fixed at 300 kHz and allow for data throughput rates of up to 115 kbit/s. The RF modules were build into shielded cases to try an minimize high frequency noise from the antennas. The 3 μs Modified-Miller pulses, obtained from ASK demodulator, are stretched to decrease the required RF

bandwidth. To prevent the ACG from amplifying background noise when the carrier is keyed off for a length of time both channels transmit a '1', logic high, when idle. A '0', logic low, only then occurs during data transfer. The Modified-Miller code is already in this format but it is necessary to invert the Manchester code before transmitting it to the Proxy. An inverting comparator is used to invert the data in addition to sharpening the pulse edges deformed by the RF filtering.



(a) Proxy (b) Mole

Fig. 4. Implemented hardware

The Mole, as shown in Figure 3, contains a modified commercial reader and some signal shaping functions. The OEM reader is build around the Philips MF RC530 ISO 14443A compatible contactless reader IC, controlled by a 8-bit Microchip PIC16F876 microcontroller. The OEM reader served as our development platform as we could easily reconfigure its functions by reprogramming the microcontroller. The MF RC530 can be set up to transmit the data on its *MFIN* pin to the card by modulating the carrier in addition to demodulating the load modulated sidebands and providing Manchester-coded output data on its *MFOUT* pin. Therefore the MF RC530 IC performs all required modulation and demodulation of data communicated via the contactless interface. The stretched pulses received from the Proxy are reshaped to their original width and then transmitted to the card, via *MFIN*. The Manchester encoded data from *MFOUT* is then inverted and transmitted back to the Proxy. Figure 4 shows the implemented Proxy and Mole prototypes. We were not really concerned with making the prototypes compact or covert as they were meant to be proof-of-concept hardware. Size can be reduced by designing

a PCB and using surface mount components. The larger IC packaging and stripboard give the hardware a homemade appearance that only serves to illustrate its simplicity.

5 Results

We successfully executed a relay attack, using the hardware described in Section 4.2, up to a distance of 50 m. Figure 5 shows a sent *REQA* command and a received *ATQA* response, in addition to their respective data modulation schemes as measured at the reader's antenna. Waveform 2 shows the complete communication sequence: the sent ASK modulated data, response time plus system delay and the load modulated response. Waveforms 1 and 3 show the corresponding digital data.

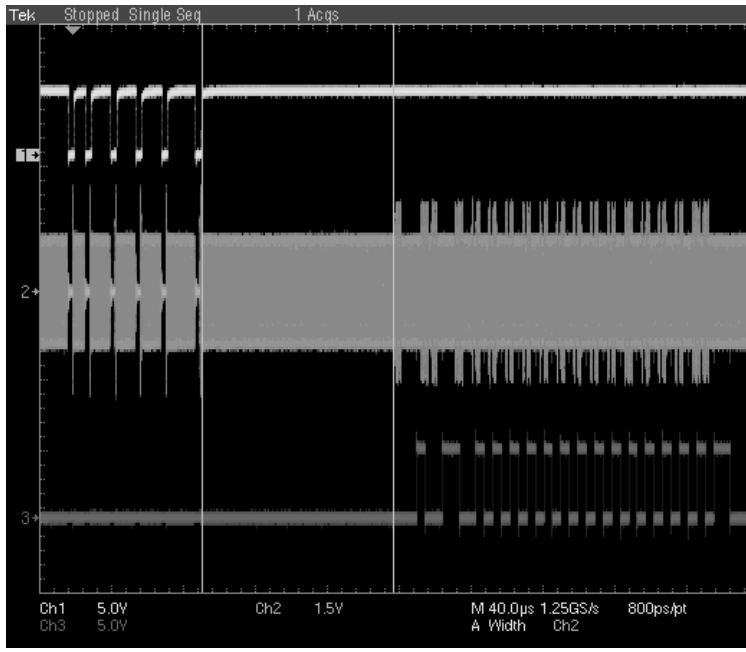


Fig. 5. Example of a relayed data sequence. Modified Miller *REQA* command (top), carrier modulation (middle) and Manchester *ATQA* response (bottom)

It turned out that the timing constraints were not a problem. Even though our system introduced a delay of 15–20 μs, the reader still functioned normally. A problem would be encountered with the anti-collision protocol if another card is placed in the reader's field in addition to our

Proxy. As shown in Figure 6 the responses from the nearby card and the Proxy are not synchronised. The misaligned bits will be interpreted by the reader as collisions and it will be unable to select a card. Within the context of our attack it should not happen as the attacker will be the the only person interacting with the reader. This problem does not extend to the situation where the victim has multiple cards within the Mole's range. Their responses are delayed by the same amount of time and are therefore still synchronised.

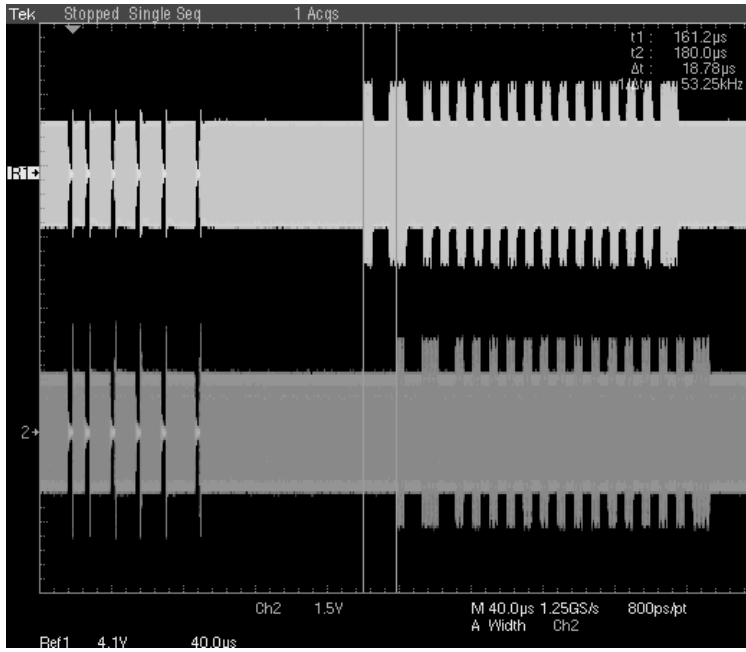


Fig. 6. Timing comparison of a card (top) and Proxy (bottom) response.

6 Implications and Related Work

The attack is difficult to defend against, as it is not strictly a failure of cryptographic protocols. Contactless cards are equipped with cryptographic algorithms to implement confidentiality, authentication and integrity services. Cards provide well known algorithms such as AES, DES, RSA and SHA-1 while some manufacturers provide proprietary solutions e.g. Philips Crypto1. There is work being done on RFID tag security [9]

with solutions focusing on privacy, using cryptographic locking [10] and anti-collision schemes [11][12], but these are not applicable to contactless cards. All these methods are still circumvented by a relay attack as they are at the application layer. New mechanisms are therefore needed that would detect or prevent this type of attack.

Solutions should therefore focus on the crucial elements of the attack: Added time delay and unnoticed access to the card. Physical shielding of the device from radio signals by enclosing it in a Faraday cage, consisting of a metal mesh or foil, could prevent access by unauthorized parties [13]. This is a viable option in some cases and could prevent unauthorised access to the card. The problem is that the card still needs to be extracted at certain intervals to enable the user to access services, which still provides a window of opportunity for attackers and inconveniences the user. Strict time checking might be able to detect our 20 μs delay, although this is only possible with *REQA* and *SELECT* signals where timing is defined. Delays in data transfers are not specified and variations in processing time are too great. Even if high resolution time stamps were used in the anti-collision phase the system could circumvent it. There is no security provided by the anti-collision scheme and the Mole can read the ID of the remote card and send it to the Proxy, who then stores it and responds to the reader within the required time. Brands et al. [14] suggested placing an upper-bound on the physical distance to another party in an attempt to prevent man-in-the-middle attacks in identification schemes.

Wireless networks and context-aware computing have driven research into location-aware systems and services. Possibilities include using Time Of Arrival (TOA), Difference in Time Of Arrival (DTOA) and Angle Of Arrival (AOA) with triangulation for positioning. Such have been demonstrated in an indoor environment [15] [16] but require hardware capable of high sampling rates and complex DSP capability. Ultrasound [17] and Received Signal Strength (RSS) [18][19] are also used for distance measurement. Sound travels slower than light, so greater resolution can be obtained in simple transponders using ultrasound frequencies. Sound has been applied in some positioning applications such as the Bat system [20]. The Echo protocol [21] implements a distance bounding protocol to verify location claims using ultrasound. However, it would be possible to insert a wireless relay section between two nodes, the communication propagation speed will be greater than that of sound, which makes the distance appear to be less. Similarly attackers can increase or decrease transmission strength in order to spoof location.

7 Conclusions

We implemented a practical relay attack against a contactless smart card system using self-built hardware. Foreseen limitations, such as timing constraints were not as strict as defined in the standards, and allowed sufficient time to relay messages. The necessary hardware parts were easily obtained and the cost of the whole system was well under £100, with most of the cost being the OEM reader.

With relay attacks against contact-based smart cards, the user has to insert his card into the attacker's hardware or allow the card to be taken away for 'payment'. In our attack the user is unaware that his contactless card is being accessed. As a result it might be some time before the effects of an attack, or multiple attacks, are discovered. And as new contactless payment schemes are devised, with greater transaction values than a just bus fare, this type of attack could become more attractive to criminals.

This attack is invisible to application layer security and therefore new protection measures should focus on the physical layer. An attacker wishing to forward data between a card and reader that are a distance apart will be unable to avoid causing a delay in the system. Distance bounding protocols would be an effective way of preventing relay attacks, although the practical implementation for a low-cost passive token with limited resources remains a problem to be solved.

References

1. J. Collins. *Dexit Turns RFID Cards into Cash.* `http://www.rfidjournal.com/article/view/673`
2. Mastercard PayPass. `http://www.paypass.com/`
3. M. Roberti, *Fear of Big Brother*, RFID Journal. `http://www.rfidjournal.com/article/view/276`
4. ISO 14443. *Identification cards – Contactless integrated circuit cards – Proximity cards.*
5. ISO 15693. *Identification cards – Contactless integrated circuit cards – Vicinity cards.*
6. K. Finkenzeller, *RFID Handbook: Radio-frequency identification fundamentals and applications*, Wiley, 1999.
7. ISO 9798. *Information Technology – Security techniques – Entity authentication*
8. J. Sorrels. *Optimizing read range in RFID systems*, EDN. `http://www.edn.com/article/CA84480.html`
9. S.E. Sarma, S.A. Weis and D.W. Engels, *RFID Systems, Security & Privacy Implications*, Auto-ID Labs, 2002. `http://www.autoidlabs.org/whitepapers/`
10. A. Juels. *Minimalist cryptography for RFID tags*, 2003. `http://www.rsasecurity.com/rsalabs/staff/bios/ajuels/publications/minimalist/index.html`

11. S.A. Weis, S.E. Sarma, R.L. Rivest and D.W. Engels, *Security and Privacy Aspects of Low-Cost Radio Frequency Identification Systems*, First International Conference on Security in Pervasive Computing, 2003. `http://theory.lcs.mit.edu/sweis/spc-rfid.pdf`
12. A. Juels, R.L. Rivest and M. Szydlo. *The Blocker Tag: Selective Blocking of RFID Tags for Consumer Privacy*, Proceedings of the 10th ACM conference on Computer and communication security, pp 101–111, October 2003.
13. *Foiling the Oyster Card*, Spy Blog, February 2004. `http://www.spy.org.uk/spyblog/archives/000198.html`
14. S. Brands and D. Chaum. *Distance Bounding Protocols*. Advances in Cryptology EUROCYPT '93, Springer-Verlag LNCS 765, pp 344–359, May 1993.
15. J. Werb and C. Lanzl. *Designing a positioning system for finding things and people indoors*. IEEE Spectrum, Volume: 35, Issue: 9, pp 71–78, September 1998.
16. P. Bahl and V.N. Padmanabhan. *RADAR: an in-building RF-based user location and tracking system*. IEEE Proceedings Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, pp 775–784, March 2000.
17. A. Ward, A. Jones and A. Hopper. *A New Location Technique for the Active Office*. IEEE Personal Communications Magazine, Vol. 4, No. 5, pp 42–47, October 1997.
18. K.P. Fishkin and S. Roy, *Enhancing RFID Privacy via Antenna Energy Analysis*, RFID Privacy Workshop, 2003. `http://www.rfidprivacy.org/papers/fishkin.pdf`
19. J. Krumm and E. Horvitz. *LOCADIO: Inferring Motion and Location from Wi-Fi Signal Strengths*. First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, Mobiquitous 2004, August 22–26, 2004. `http://research.microsoft.com/~horvitz/locadio.pdf`
20. A. Harter, A. Hopper, P. Steggles, A. Ward and Paul Webster. *The Anatomy of a Context-Aware Application*. Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking, MOBICOM'99, pp 59–68, August 1999.
21. N. Sastry, U. Shankar and D. Wagner. *Secure verification of location claims*. Proceedings of the 2003 ACM Workshop on Wireless Security, pp 1–10, September 2003.