Chapter 13
RFID and Contactless Technology

Anjia Yang and Gerhard P. Hancke

Abstract An increasing number of 'contactless' systems are based on passive Radio-Frequency Identification (RFID) technology. A passive RFID token is powered by a transmitted RF carrier, which is also used for bidirectional communication. RFID technology comprises of several standards, which are suitable for different applications. Electronic Product Code (EPC) tags, contactless credit cards, e-passports, and access control are just a few examples of systems that use a subset of this technology. This chapter contains a brief explanation of RFID operating principles along with an overview of prominent implementations and industry standards.

Keywords Radio Frequency Identification (RFID) · Contactless Cards · Near-Field Communication · Over the Air (OTA) · Deployment Modes · Secure Element · Relay Attack · Cloning · ISO 18092 · ISO 14443 · ISO 15693

13.1 Introduction

RFID is a technology that increases productivity and convenience and is currently being integrated into many areas of society. RFID is flexible and the variety of standards and tokens available allows it to be tailored to any application. The technology also offers additional benefits such as reduced maintenance cost and extended product lifetime. The uses of radio-frequency identification have therefore grown remarkably, and it is reported that the cumulative number of tokens sold from 1943 to the start of 2015 is about 34 billion [1]. The total RFID market, including systems and services, is valued at $9.56 billion at 2015 and expected to increase to $14.5 billion by 2020 [1].

Despite its recent popularity, RFID technology has been around for more than half a century. It is commonly believed that the concept of radio identification started in the early 1940s with the advent of radar when IFF transponders actively modulated the

A. Yang · G.P. Hancke (✉)
City University of Hong Kong, Hong Kong, China
e-mail: ghancke@ieee.org

A. Yang
e-mail: ayang3-c@my.cityu.edu.hk

K. Mayes and K. Markantonakis (eds.), *Smart Cards, Tokens, Security and Applications*, DOI 10.1007/978-3-319-50500-8_13

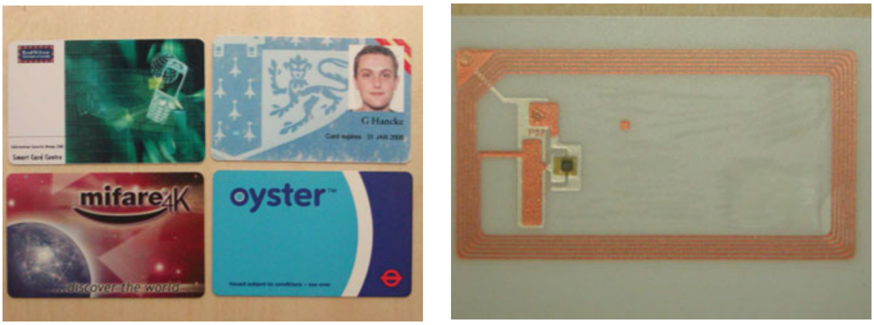
radiated ground radar signals to identify airplanes. Despite early work on RFID, such as Stockman's '*Communication by Means of Reflected Power*' in 1948 [2], recognising the potential of RFID, it would take several more years, and additional advances in electronics, before the technology was implemented in further applications. During the 1960s, stores and libraries used electronic article surveillance, an early 1-bit form of RFID, for theft control. Meanwhile, private and government research on the subject continued and in 1973 the first patents were filed that resembled modern systems: a token with rewritable memory by M.W. Cardullo and a passive token used to unlock doors by C. Walton. Scientists from the Los Alamos National Laboratory, who were asked by the US Energy Department to develop a tracking system for nuclear materials, also demonstrated the concept of modulated backscatter with 12-bit tokens operating at 915 MHz in the same year. The basic communication principle of this system is still used today by the majority of Ultra-High-Frequency (UHF) RFID tags. In the late 1980s, RFID gained widespread acceptance in automated toll collection and access control systems, which was followed by implementation in public transport payment systems and the first serious attempts at standardisation in the 1990s. In 1999, it was proposed that low-cost UHF RFID 'tags' could be used to track items in supply chains. Currently, the use of Electronic Product Code (EPC) tags for tracking at the pallet, case, and item level is probably the most prominent RFID application, driven by government agencies, such as the US Department of Defense, and various large retailers such as Tesco and Walmart. RFID technology is, however, also used on a large scale in other applications such as machine-readable travel documents (e.g., e-passports) and credit cards [3].

This chapter is intended as an introduction to RFID and summarises the aspects most relevant to contactless smart card systems. Section 13.2 gives a brief overview of existing systems, describing the general operating principles and available technology and discusses a number of high-profile implementations. The RF interface and communication theory are discussed in Sect. 13.3 and the current HF RFID standards are summarised in Sect. 13.4. The chapter concludes with an overview of Near-Field Communication (NFC) in Sect. 13.5.

13.2 Contactless Technology

Even though RFID is a collective term for a number of technologies, it is often used primarily to describe applications using low-resource devices, such as EPC tags. Devices used in identity and payment systems, like the cards shown in Fig. 13.1, are instead regularly referred to as 'contactless' or 'proximity' tokens. These devices operate in the High-Frequency (HF) radio band, contain more resources, and have a much shorter operating range than their UHF counterparts.

Examples of contactless tokens are shown in Fig. 13.1a. Each token, or Proximity Integrated Circuit Card (PICC), contains an antenna and an Integrated Circuit (IC) as shown in Fig. 13.1b. The IC performs modulation and demodulation of the RF channel and is also responsible for data storage and processing. The passive token derives its power from the RF carrier transmitted by the reader, or 'Proximity Coupling



(a) Example tokens

(b) Example of an inlay, showing RFID IC and antenna

Fig. 13.1 ‘Contactless’ tokens

Device’ (PCD). The bidirectional communication between the token and reader is also modulated onto this carrier.

The main benefit of contactless technology is its ease of use. The user does not have to physically insert his token into the reader or orientate the token in a specific way. In most cases, the token can be kept in a wallet or a purse providing some measure of personal security. All these factors combine to provide fast transactions and ensure high throughput. Furthermore, readers and tokens have no external mechanical parts that can wear out. This makes systems more durable and reliable, especially in exposed or dirty environments, and reduces maintenance costs when compared to contact or magnetic stripe systems. The fact that the token does not need to contain a power supply adds to the lifetime of the token as well. Contactless tokens can also provide the same amount of security mechanisms than contact smart cards and there are several established international standards available to aid interoperability [4].

Contactless systems differ from each other in a number of ways. It is therefore important to note the alternatives offered by the available standards and products in order to decide on the best system components for a specific application. For example, the three HF standards (ISO 14443 [5], ISO 15693 [6], and ISO 18092 [7]) allow for different data rates and operational ranges. The growth in the contactless market has also resulted in a variety of readers and tokens. In most cases, the reader’s only purpose is to act as a RF transceiver between the back-end system, which performs all the processing and security functions, and the token. In general, the only difference between readers are the standards that they support. Readers are also more expensive and are often installed as a long-term infrastructure investment. It is therefore common to find readers supporting multiple standards so that the same hardware can be used in several applications, possibly allowing for future changes and extra functionality.

In contrast, there are quite a few tokens available that can be classified in terms of resources, security, and interfaces [4]. In terms of resources, tokens can be divided into three different types:

- **Memory**: These tokens only have the ability to store information. They can perform no processing in addition to read and write functions.
- **Logic**: In addition to memory, these tokens also include some fixed processing routines that can be invoked by the reader, e.g., authenticate, increment value, decrement value.
- μ**-Controller (MCU)**: The token can run custom processing routines and might contain card operating systems such as JCOP or MULTOS.

Tokens can implement a number of security mechanisms. Security increases the token's required resources resulting in a higher system cost, so it is important that the token used provides a level of security sufficient for the application without incurring unnecessary expense. These levels can roughly be defined as follows:

- **Minimal**: Anyone can read information stored on the token. Memory can be locked so no unauthorized writing of data occurs.
- **Low**: The token implements some form of authentication mechanism. Memory is password protected or mutual authentication must be completed before data is released.
- **Medium**: The token implements a single encryption algorithm used to provide authentication and encryption of data. The algorithm could be proprietary, e.g., NXP Crypto1, or an industry standard, e.g., Data Encryption Standard (DES).
- **High**: The token can implement a number of symmetric and asymmetric industry standard algorithms for authentication, encryption, digital signatures, etc.

It may be required that a token supports several technologies. This allows for an environment where the user can carry a single token to access multiple systems. Alternatively it provides a way to migrate to, or add, a new system while still maintaining backward compatibility with existing systems. The following tokens are available that have the ability to interact with more than one system:

- **Multiple-Technology**: The token implements multiple technologies. A good example, albeit not contactless, of this is Chip and Pin cards in the UK that contain both magnetic stripe and contact smart card technologies.
- **Dual-Interface**: The token contains one integrated circuit that has more than one interface. An example would be a token containing an Integrated Circuit (IC) with both a contactless and a contact interface.
- **Hybrid**: A token with two or more integrated circuits, with their own interfaces, functioning independently. This term can be used to describe HF contactless tokens that also contain additional circuitry to support older Low-Frequency (LF) systems.

13.2.1 Applications

Contactless tokens act as an electronic credential, interacting with the rest of the system on behalf of the entity it is associated with. Initially, these tokens allowed for new applications such as contactless access control and automatic toll collection. In

Table 13.1 Summary of HF RFID tokens applications

Application	Standard	Token resources	Security
Item tracking	ISO 15693	Memory	Minimal/low
Ticketing	ISO 15693 ISO 14443	Memory/logic	Low/medium
Closed payment	ISO 14443	Logic	Low/medium
Open payment	ISO 14443	μ-controller	High
Access control	ISO 14443	μ-controller	High
Identity	ISO 14443	μ-controller	High

recent years, however, these tokens have started to replace, or supplement, established technologies such as paper tickets for travel or events, barcodes in item tracking and magnetic stripes in credit cards. This section gives an overview of prominent applications in which contactless tokens are used. Table 13.1 summarises some of these applications.

13.2.1.1 Identification

The basic function of RFID, as the name already suggests, is to assist a system in uniquely identifying an item or a person. The simplest example of this is *tracking* systems where a token, storing a Unique Identifier (UID), is attached to an item. The system can then track this item by scanning the token every time it passes a reader. Systems using HF tokens have a shorter operational range than their UHF equivalents although they provide more reliable reader coverage. NXP I-CODE and the Texas Instruments Tag-It products are examples of HF tokens used in tracking applications. In *ticketing* systems, tokens facilitate access to services for a limited time before being disposed of. A ‘ticket’ can take many forms, such as a key card to a hotel room or a day pass at the local gym. Paper tickets, containing NXP Mifare UltraLight tokens, received extensive publicity during the FIFA World Cup in 2014 and were used to gain fast access to stadiums and enhance personal security by tracking customers going off to remote locations such as ski slopes [8].

Contactless *access control* is popular for securing physical locations. Charles Walton first invented an RFID-based access control system in 1973. The system involved an electronic lock that opened with an RFID key card, which he sold to Schlage [3]. Since then, contactless access control systems have become widespread not only in the private sector but also with government agencies. An application closely linked to access control is that of *identity*. A token used for proof of identity must contain enough information to allow the system to verify that the person presenting it is the legitimate owner. Identity tokens therefore contain additional personal information, such as biometric data.

As an example, the Federal Information Processing Standard Publication 201 (FIPS 201) detailing Personal Identity Verification (PIV) of federal employees and contractors [9] is published by the US National Institute for Standards and Technology (NIST) at 2006 and has been updated to FIPS 201-2 [10] at 2013. It provides specifications for a standard Federal smart ID card that is to be used for access control. The cards are a requirement for all federal employees and contractors under the Homeland Security Presidential Directive 12 (HSPD-12). Other large government initiatives in the USA include:

- The Department of Defense's Common Access Card with Contactless (CAC-C) being used as identification for on duty military personnel, reserve personnel, and civilian employees.
- The Transportation Worker Identification Credential (TWIC) being issued by the Transportation Security Administration.
- The First Responder Authentication Card (FRAC) being issued in the Department of Homeland Security (DHS) pilots [11].

Contactless tokens are also used for national identity cards and some countries are planning to use contactless ID cards for their citizens, with China becoming the largest implementer. The most prominent application of contactless identity is, however, Machine-Readable Travel Documents (MRTD). By the time of writing, the USA required that 38 countries issue their citizens with e-passports in order to still qualify under the Visa Waiver Program. E-passports adhere to operational specifications as defined by the International Civil Aviation Organization (ICAO) [12], with additional guidelines for MRTD specified in ISO 7501 [13]. By 2015, ICAO wishes to have replaced all current passports with a digital version that stores encrypted biometric data on a RFID chip. The DHS also wants to use passive RFID to record who is entering or leaving the USA across land routes using a People Access Security Service (PASS) card. ICAO allows for optional security protocols that provides both authentication and encryption services. E-passports have the interesting security requirement that anyone who is presented with the passport should be able to read and verify the contents, but at the same time the user's personal details should be afforded some measure of privacy. For this reason, most e-passports implement the Basic Access Control (BAC) scheme. BAC derives a key from the passport number, expiry date, and the user's birthday, read off the Optical Character Recognition (OCR) strip inside the passport. The idea is that anyone legitimately presented with the passport can read the OCR data, derive the key and retrieve the data off the token inside. The European Union countries have already implemented Extended Access Control (EAC), involving a public key infrastructure for participating parties, for future passports including additional biometric data [14].

13.2.1.2 Payment

RFID has been used in payment systems since the 1980s, when it was first used for automatic toll collection. Since then, contactless tokens were implemented in several cashless payment systems. In *closed* systems, one organisation is in control of the entire payment process. In other words, customers will pay for the organisation's services with payment tokens issued by the same organisation. These systems often function on a 'prepaid' principle where the customer pays for credit in advance and are ideal for public transport payments. A good example of such a system is the Oyster card scheme implemented by Transport for London (TFL) using NXP Mifare Classic tokens. Closed payment systems can also be used in other environments such as service stations, e.g., the SpeedPass system implemented by ExxonMobil, or in some cafeterias and fast food outlets.

In *open* systems, an organisation issues customers with payment tokens that will be used to purchase services from other organisations. Some of these systems still operate using prepaid credit like Hong Kong's Octopus system, implemented with Sony Felica tokens. The Octopus card is not only used mainly for transport payments but can also be used to pay at convenience stores, restaurants, and other local services. The most prominent open payment system is, however, contactless credit cards. RFID credit cards have been widely deployed in the USA [15] with American Express (ExpressPay), MasterCard (PayPass), and Visa (payWave) all supporting contactless credit and debit cards.

In the UK, the Royal Bank of Scotland, working with Mastercard, and Barclays launched contactless debit cards in 2007. In particular, the Barclays' OnePulse card, developed with Visa, is intended to function as a chip-and-pin contact card, contactless debit card, and an Oyster card. The lower level communication of these cards, as specified EMV contactless specification [16] adheres to ISO 14443 while the application layer communication adheres to the same Europay, MasterCard, and Visa (EMV) framework and specifications as contact payment cards and transaction terminals [17].

13.3 Radio-Frequency Interface

The operation of contactless systems is based on the principle of inductive coupling. The token and the reader both contain antenna coils that are coupled and interact via a magnetic field. The token receives both data and power from the carrier transmitted by the reader. The token can also send data to the reader by influencing this carrier. Collectively, these methods are referred to as near-field communication since the range between the token is much smaller than one wavelength of the carrier. This section provides an overview of communication theory and physics relevant to contactless systems. The information in this section has been adapted from [18–20].

13.3.1 Communication Theory

In order to transmit information over an RF channel, the relevant data must first be encoded and then modulated onto a suitable RF carrier. Line, or data, coding changes the binary data into a signal sequence that is best suited to the transmission channel and aids the receiver in recovering the data. Modulation is the process whereby the parameters of an RF carrier is altered in relation to the resultant baseband signal. Modulation, and to a lesser extent coding, techniques can be used to shape the frequency spectrum of the communication channel. In contactless systems, coding and modulation methods have two main prerequisites: to separate the 'weak' backward channel communication from the strong forward channel carrier and to allow for data transfer from the reader to the token while ensuring that the token still receives adequate power from the HF carrier.

A line code's properties will typically be chosen to allow for the physical requirements of the transmission channel. Line codes are most often used to eliminate the DC component of the data and help the receiver with synchronisation, although some codes also provide redundancy to prevent errors. HF RFID tokens use the schemes shown in Fig. 13.2.

Non-Return-to-Zero (NRZ) coding is the most used of basic coding techniques. A '1' is represented by a logical high for one clock period and a '0' is represented by a logical low for one clock period. NRZ encoding is not ideal when used to transmit data without a maximum runlength constraint, or in other words data that contains long sequences of ones or zeros. In this case, there will be no signal transitions between high and low for a period of time, which can prevent the recovery of an accurate data clock. This also has other consequences, e.g., if a long sequence of zeros is transmitted using 100% amplitude modulation, it could disrupt the token's power supply.

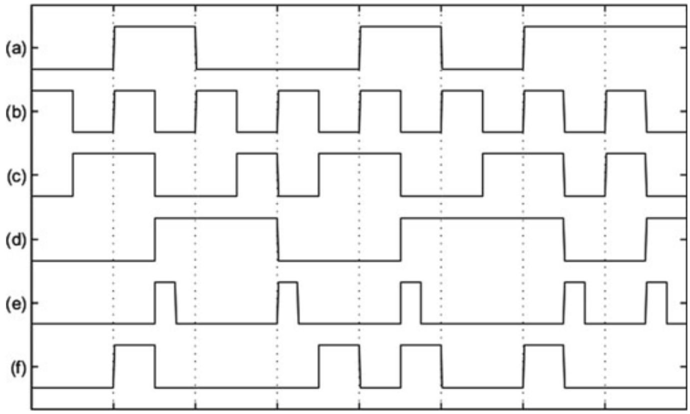


Fig. 13.2 Data coding examples: **a** Non-Return-to-Zero (NRZ), **b** data clock, **c** Manchester, **d** Miller, **e** Modified Miller, and **f** pulse-position

Manchester coded sequences have at least one transition during each bit period. This periodic transition, which occurs in the middle of the bit period, can therefore be used to recover the data clock regardless of the data values. A ‘0’ is expressed by a low-to-high transition and a ‘1’ by a high-to-low transition. Although Manchester encoded sequences contain no DC component, it requires approximately twice the channel bandwidth of NRZ.

Modified Miller coding is often used for data transmission from the reader to token. The data is encoded using Miller coding after which each transition, high-to-low and low-to-high, is replaced by a single pulse. In Miller encoding, a ‘1’ is represented by a bit period with a transition at the midpoint. A ‘0’, if preceded by a ‘1’, is represented by a bit period with no transition while a ‘0’, preceded by another ‘0’, is represented by a bit period with a transition at the start of the bit period. The advantage of Modified Miller coding is that the carrier is not interrupted for more than the duration of the coding pulse, even if a bit period is very long. This ensures a continuous power supply to the token from the reader’s carrier during data transfer. The bandwidth of Modified Miller coding depends on the duration of the coding pulse.

Pulse-Position coding, which is also sometimes referred to as Pulse-Position Modulation (PPM), represents x data bits with a single pulse in one of 2^x possible time slots. In Fig. 13.2, an example of ‘1 in 4’ PPM is given. In this case, two bit periods are divided into four time slots. The data bits are then encoded as follows: ‘00’ is represented with a pulse in the last slot, ‘01’ with a pulse in slot three, ‘10’ with a pulse in slot 2, and ‘11’ with a pulse in the first slot. PPM is also suitable for reader to token communication since the carrier is not interrupted for more than the duration of the coding pulse. As with Modified Miller coding, the required bandwidth is determined by the time width of the coding pulse.

Modulation is the process whereby a RF carrier’s parameters are changed to represent a baseband data sequence. A sinusoidal carrier can be characterised by

$$x(t) = a \cdot \sin(2\pi f_c t + \phi) \tag{13.1}$$

From the equation, it is clear that the amplitude a, frequency f_c, and phase ϕ can be varied to create distinctive carriers that are suitable to represent different data symbols. The amount that the chosen variable changes in relation to the data is referred to as the modulation index m_i. For example, m_i for amplitude modulation can be represented as

$$\frac{a_{\max} - a_{\min}}{a_{\max} + a_{\min}} \tag{13.2}$$

When modulating digital data, the chosen variable will only change between a set number of discrete values, e.g., high signal represented by $f = 10\,\text{Hz}$ and low signal represented by $f = 20\,\text{Hz}$. As a result, modulation is often referred to as ‘shift keying’ in digital systems. Examples of the modulation schemes used in HF RFID are shown in Fig. 13.3.

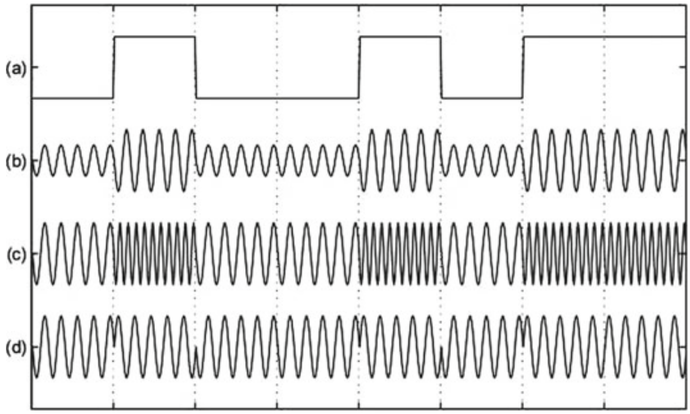


Fig. 13.3 Examples of RF modulation: **a** NRZ encoded data, **b** Amplitude-Shift Keying (ASK), **c** Frequency-Shift Keying (FSK), and **d** Phase-Shift Keying (PSK)

Amplitude-Shift Keying (ASK) changes the amplitude of the carrier to a level chosen to represent a specific data symbol. In the example shown, a ‘1’ is represented by a carrier with amplitude A and a ‘0’ is represented by a carrier with amplitude $0.5A$. In this case, the modulation index would be $0.33 \approx 33\%$. On–Off Keying (OOK) is a special case of ASK where the modulation index is equal to 100%.

Frequency-Shift Keying (FSK) changes the frequency of the carrier to represent a specific data symbol. In the example shown, a ‘0’ is represented by a carrier with $f_c = f_1$ and a ‘1’ is represented by a carrier with $f_c = 2 \cdot f_1$.

Phase-Shift Keying (PSK) represents each specific data symbol by a shift in the carrier’s phase. In the example shown, a ‘0’ is represented by a carrier with phase equal to 0° and a ‘1’ is represented by a carrier with phase equal to 180°.

The modulation process also changes the frequency-domain representation of the data. ASK and PSK cause the power spectrum of the data to move from the baseband to f_c, while the spectrum power components for FSK data, as per our example, will be at f_1 and $2 \cdot f_1$.

The modulation process can be represented mathematically as follows:

$$x(t) = d(t) \cdot \sin(2\pi \cdot f_c t)$$
$$X(f) = D(f) * (\delta(-f_c) + \delta(f_c))$$
$$X(f) = D(f + f_c) + D(f - f_c)$$

where $d(t)$ is the baseband data sequence with frequency spectrum $D(f)$.

This is useful in RFID systems where both the forward and backward channel data are transmitted using the same carrier. The signal power of the backward channel can be up to 80 dB smaller than that of the carrier, which means that the reader would find it difficult to distinguish this ‘weak’ data from the ‘strong’ carrier if it cannot effectively isolate the data of interest and attenuate the carrier. Separating the two channels in the frequency domain simplifies the recovery of the backward channel

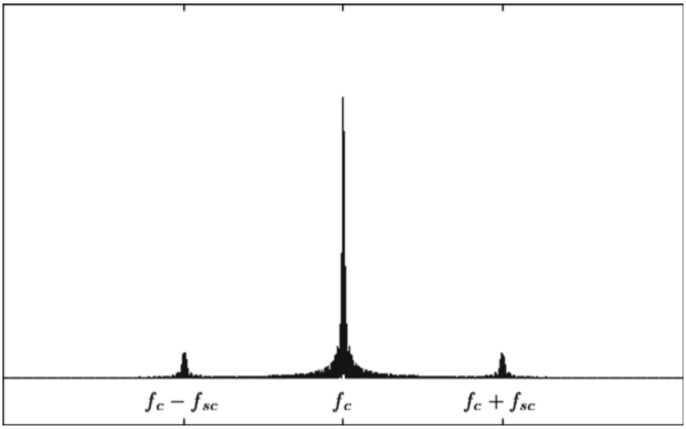


Fig. 13.4 The theoretical positive-frequency spectrum of the forward and backward channel modulated using a carrier with frequency f_c and a sub-carrier with frequency f_{sc}

data. In HF RFID, the forward channel data is modulated directly onto the main carrier. The backward channel, however, is first modulated onto a sub-carrier before being modulated onto the main carrier. Modulating with a sub-carrier creates two data sidebands separated by f_{sc} from the operational frequency f_c. The backward channel's modulation process can be represented mathematically as follows:

$$x(t) = (d_B(t) \cdot \sin(2\pi \cdot f_{sc}t)) \cdot \sin(2\pi \cdot f_c t)$$
$$X(f) = (D_B(f) * (\delta(-f_{sc}) + \delta(f_{sc}))) * (\delta(-f_c) + \delta(f_c))$$
$$X(f) = D_B(f + f_c + f_{sc}) + D_B(f + f_c - f_{sc}) + D_B(f - f_c + f_{sc}) + D_B(f - f_c + f_{sc})$$

where $d_B(t)$ is the backward channel data with frequency spectrum $D_B(f)$. The resultant positive-frequency spectrum of the forward and backward channels is shown in Fig. 13.4. The data at $f_c + f_{sc}$ is referred to as the upper-sideband while the data at $f_c - f_{sc}$ is referred to as the lower-sideband. The backward channel data can now be recovered, despite the presence of a strong operational carrier, by bandpass filtering one of the sidebands.

13.3.2 *Inductive Coupling*

HF RFID systems work on the principle of inductive coupling, where one device transfers energy to another by means of a shared magnetic field (H). This operational model is true as long as the token is placed in the near field of the reader, since the high-frequency electromagnetic field (E) generated by the reader acts primarily as a magnetic field if the distance between the reader and token is less than $\lambda_{f_c} \cdot \frac{1}{2\pi}$. In simple terms, an RFID system acts in a similar same way to a transformer, with the primary coil contained in the reader and the secondary coil contained in the token. Current flowing through the reader's coil generates a magnetic field, which in turn induces a proportional current flow in the coil of the token. The high-frequency carrier can therefore be used for both data and power transfer between the reader and the token.

13.3.2.1 Power Transfer

Moving charge, such as the flow of current in a conductor, generates a magnetic field. The magnitude of this field at a specific point is described by the magnetic field strength H. In HF RFID systems, the reader uses conductor loops to generate a magnetic field. The magnitude of the magnetic field generated by these loop antennas depends on the current I, number of loops N, the radius of the loops R, and the distance from the loop antenna d. The following equation can be used to calculate the field strength along the axis of the loop antenna:

$$H = \frac{I \cdot N \cdot R^2}{2\sqrt{(R^2 + d^2)^3}} \tag{13.3}$$

In this case, d is the distance from the centre of the coil along the coil's axis. In general, the field strength is almost uniform at short distances (d < R) where smaller loop antennas also have a higher field strength. The larger antennas, however, have higher field strength at greater distances. A token will specify the minimum field strength it needs to function. Designing the antenna is then a trade-off between making the antenna small enough to generate a strong enough magnetic field and also making it large enough to achieve the system's required operational range.

In an HF RFID system, both the reader and the token contain loop antennas in close proximity, as shown in Fig. 13.5. If a second loop antenna, with area A_2, is located close to another loop antenna, with area A_1, then the second antenna will be affected by a proportion of the total magnetic flux Ψ flowing through the first antenna. The two circuits are connected together by this partial transfer of flux and is therefore said to be coupled. The magnitude of the coupled flux Ψ_{21} depends on the characteristics of the loop antennas, the position of the antennas in relation to each other and the magnetic conductivity, or permeability, of the medium between the antennas.

The ratio of the total flux that is generated in an area, to the current in the conductor that encloses that area, is described by inductance L:

$$L = \frac{\Psi}{I} = \frac{N \cdot \mu \cdot H \cdot A}{I} \tag{13.4}$$

The constant μ is equal to $\mu_0 \cdot \mu_r$, where μ_0 is the magnetic field constant ($4\pi \times 10^{-6}$) and describes the permeability of a vacuum, while μ_r is the relative permeability, indicating the ratio of the permeability of a material relative to μ_0.

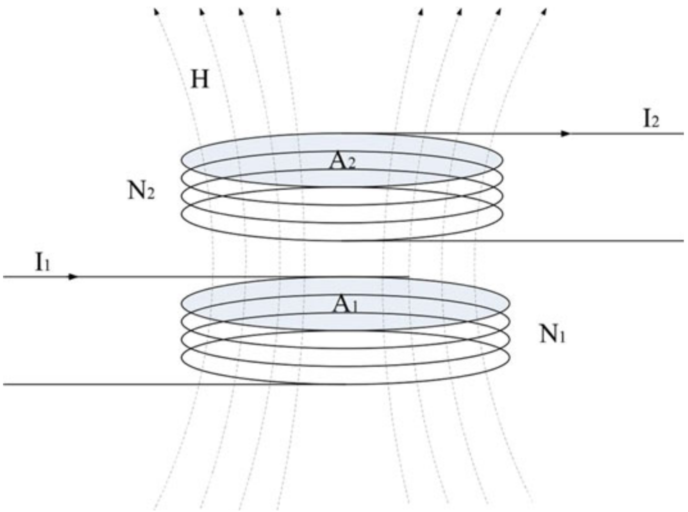


Fig. 13.5 Inductive coupling

The concept of mutual inductance is used to describe the coupling of two antennas by means of a magnetic field. The mutual inductance M_{21} is defined as the ratio of coupled flux Ψ_{21} enclosed by a second loop antenna to the current I_1 in the first loop:

$$M_{21} = \frac{\Psi_{21}(I_1)}{I_1} \tag{13.5}$$

Similarly, there is also a mutual inductance M_{12} although $M_{21} = M_{12} = M$. The mutual inductance between two loop antennas can be calculated using Eqs. 13.4 and 13.5 and can be approximated as follows:

$$M_{12} = \frac{\mu_0 \cdot H(I_1) \cdot N_2 \cdot A_2}{I_1} \tag{13.6}$$

Replacing $H(I_1)$ with Eq. 13.3 and substituting $R^2\pi$ for A the final result is:

$$M_{12} = \frac{\mu_0 \cdot N_1 \cdot R_1^2 \cdot N_2 \cdot R_2^2 \cdot \pi}{2\sqrt{(R_1^2 + d^2)^3}} \tag{13.7}$$

In practice, the HF carrier transmitted by the reader is a sinusoidal alternating current. A time-variant current $I_1(t)$ flowing in a loop antenna generates a time-variant magnetic flux $\mathrm{d}\Psi(I_1)/\mathrm{d}t$. According to Faraday's law, a voltage will be induced in the loop antenna that encloses some of this changing flux. Figure 13.6 shows a simplified circuit diagram for a coupled RFID system, where L_1 is the

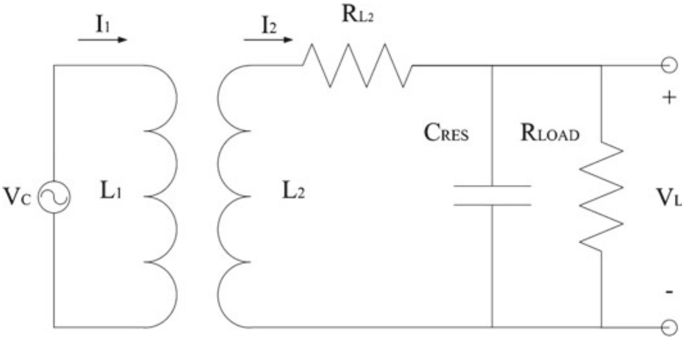


Fig. 13.6 Simplified circuit diagram of coupled token

antenna of the reader, L_2 is the antenna of the token, R_{L_2} is the resistance of the token's antenna, and R_{LOAD} represents the load. C_{RES} is ignored for now and is equal to 0.

The total time-variant flux in L_1 induces a voltage V_{L1} in L_2 due to the mutual inductance M. There is a voltage drop across R_{L_2} and I_2 also induces magnetic flux in L_2, which opposes $\Psi(I_1)$, so the voltage across the load can be approximated by:

$$V_L = \frac{\mathrm{d}\Psi_2}{\mathrm{d}t} = M\frac{\mathrm{d}I_1}{\mathrm{d}t} - L_2\frac{\mathrm{d}I_2}{\mathrm{d}t} - I_2 R_{L_2} \tag{13.8}$$

V_L can now be rectified and used as a power supply for the token. In order to improve the efficiency of the coupling, an additional capacitor C_{RES} can be added in parallel with the antenna L_2 to form a parallel resonant circuit with a resonant frequency corresponding to the operating frequency of the RFID system. The resonant frequency can be calculated as follows:

$$f = \frac{1}{2\pi\sqrt{L_2 \cdot C_{RES}}} \tag{13.9}$$

When operating at the resonant frequency, the voltage V_L induced in a system with a resonant circuit increases by more than a factor of ten compared to a system using the antenna by itself. The influence of the resonant circuit's R_{L_2} and R_{LOAD} on voltage V_L can be characterised by the Q, or quality, factor. The Q-factor is discussed in more detail later with regard to data transfer.

A further practical constraint that effects the coupling efficiency is the orientation of the token in relation to the reader. In the equations above, it was assumed that the antenna in the reader and the antenna in the token have a common axis. Although this is often the case, the token can also be displaced or tilted in such a way that the magnetic flux enclosed by its antenna is decreased. As a rule of thumb, maximum voltage is induced in the token's antenna when it is perpendicular to the magnetic field lines, while no voltage is induced if it is parallel. This results in a specific interrogation

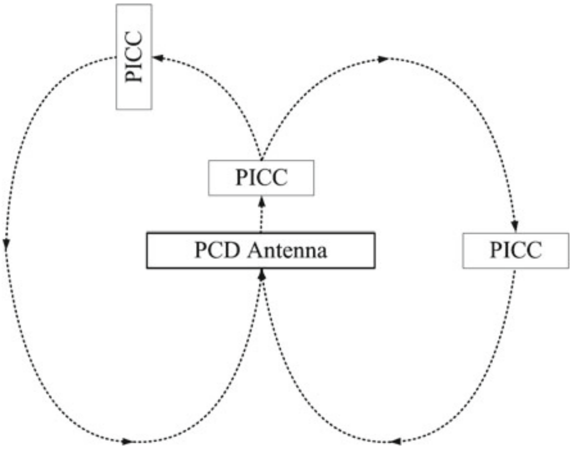


Fig. 13.7 Orientation of token-to-reader antenna for maximum coupling

zone around the reader's antenna, as illustrated by an example in Fig. 13.7. Following the expected path of the field lines in this case, it can be seen that a token parallel to the reader's antenna would be read if directly in front, or to the side, of the antenna, while a token that is perpendicular to the antenna could be read on the diagonal corners.

13.3.2.2 Data Transfer

The reader and token use different techniques to transmit data. As a result reader-to-token communication is referred to as the forward channel, while token-to-reader communication is referred to as the backward channel. The forward channel is relatively simple as the reader can directly modulate the data onto the carrier it transmits. The backward channel, however, requires that the token send data even though it is a passive device. The token must therefore modulate the reader's carrier, which in most cases is done using *load modulation*.

Load modulation works on the principle that the token's impedance Z_T can be altered by changing the parameters of the resonant circuit. Changing the impedance not only influences the voltage induced in the token's antenna L_2 but also the magnitude of the voltage across the reader's antenna L_1. The token can therefore amplitude modulate the voltage on the reader's antenna. It is only possible for the token to alter the load resistance R_{LOAD} or the parallel capacitor C_{RES} of the resonant circuit. Load modulation is therefore either resistive or capacitive. In resistive load modulation, a resistor R_{MOD} is added in parallel to R_{LOAD}, so the impedance is switched between $Z_T(R_{LOAD})$ and $Z_T(R_{LOAD}||R_{MOD})$ during the modulation process. In capacitive load modulation, an additional capacitor C_{MOD} is added, which changes the resonant frequency when switched into the circuit. Detuning the resonant circuit greatly influences the token's impedance, which causes the desired modulation effect.

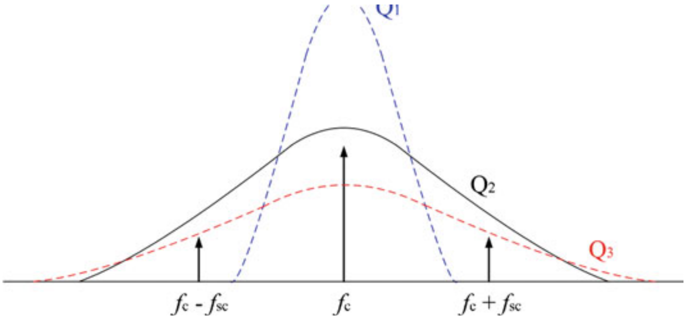


Fig. 13.8 The effect of the Q-factor

As mentioned before, the Q-factor is a measure of the voltage in the token when operating at the resonant frequency. It can be approximated as follows:

$$Q = \left(\frac{R_{L_2}}{2\pi f_c L_2} + \frac{2\pi f_c L_2}{R_{LOAD}} \right)^{-1} \tag{13.10}$$

Generally, the value of Q should be maximised to allow for the maximum power transfer and operational range. It should be kept in mind that Q also influences the bandwidth of the communication channel. The frequency response of Q peaks around the resonant frequency and rolls off to either side, as shown by the examples in Fig. 13.8. This indicates that the resonant circuit acts as a crude bandpass filter centered around f_c with bandwidth $BW = f_c/Q$. The value of Q must therefore be chosen in such a way that it allows sufficient power transfer, while still allowing for the backward channel's modulation sidebands. In Fig. 13.8, the value of Q_1 is too high since it excludes the modulation side bands, whereas the value of Q_3 is too low since it decreases the center frequency needed for power transfer.

13.4 Standards

RFID technology encompasses a range of systems from multiple vendors. To ensure interoperability between different RFID systems, several standards have been defined to which these systems must adhere. In the HF band, there are three main standards acknowledged by ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) that deal with HF RFID technology operating at 13.56 MHz. These are ISO/IEC 14443 [5], ISO/IEC 15693 [6] and finally ISO/IEC 18092 [7]. Another standard that should be mentioned is ISO/IEC 18000, which defines RFID communication interfaces for several operating frequencies, including 13.56 MHz. The standards define, among other things, the RF interface, the initialization sequence, and the data format. It is not feasible to describe each standard in its entirety within this chapter, so only a summary of the key technical aspects of the interaction between the reader and the token is presented. ISO 14443 and ISO 15693 are discussed in more detail since it is the most relevant to the applications described in Sect. 13.2.1, with ISO 18000 and ISO 18092 discussed only briefly. The information in this section has been adapted from [5–7, 18, 21] and the reader should consult these sources for a more complete description.

13.4.1 ISO 14443

ISO 14443, titled *Identification Cards—Proximity Integrated Circuit Cards*, is commonly used in systems using logic and μ-controller tokens. This means that it is the standard of choice for e-passports, credit cards, and most access control systems. The popular Mifare range of products by NXP also adhere to Part 1–3 of ISO 14443 Type A.

13.4.1.1 Part 1—Physical Characteristics

The first part of the standard states that the token shall have physical characteristics according to the requirements for the card type ID-1 specified in ISO/IEC 7810, i.e., 85.72 mm × 54.03 mm × 0.76 mm. It also specifies tolerance levels for the token with regard to ultraviolet light, X-rays, dynamic bending stress, dynamic torsional stress, alternating magnetic fields, alternating electric fields, static electricity, static magnetic fields, and operating temperature.

13.4.1.2 Part 2—Radio-Frequency Power and Signal Interface

The second part of the standard describes the signal characteristics of two types of radio interfaces between the token and the reader. These interfaces allow for both power transfer and bidirectional communication. The token's power is provided by an alternating magnetic field at a frequency of 13.56 MHz and the reader must ensure that the magnetic field is within the range $1.5\,\text{A/m} \leq H \leq 7.5\,\text{A/m}$. The operational range for systems using ISO 14443 usually extends up to 10 cm.

ISO 14443 defines two different methods for data transfer. In *Type A*, the forward channel data uses Modified Miller coding with a coding pulse of 2–3 µs modulated onto the 13.56 MHz carrier with 100% ASK. The backward channel uses Manchester encoding, which is 100% ASK modulated onto a 847 kHZ sub-carrier before being load modulated onto the 13.56 MHz carrier. In *Type B*, the forward channel uses NRZ encoding modulated onto the 13.56 MHz carrier with 10% ASK. The backward channel uses NRZ encoding, which is first modulated onto a 847 kHZ sub-carrier using PSK (0°, 180°) before being load modulated onto the 13.56 MHz carrier. The basic data rate for both the forward and backward channels in Type A and Type B is

106 kbps. Some ISO 14443 tokens and readers support higher data rates, 212/424/848 kbps, which can be selected after the anti-collision process has finished.

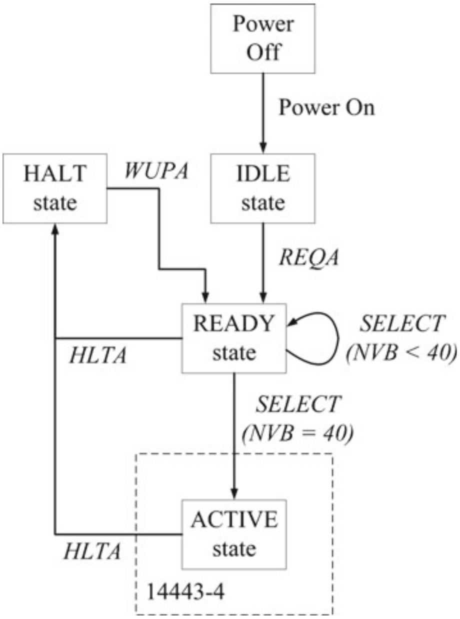
13.4.1.3 Part 3—Initialization and Anti-collision

This part of the standard describes byte and data frame formats and the initial commands used to detect and initialise communication with the token. This includes the ability to poll for new tokens in the interrogation field and choosing a token, even if multiple tokens are present, through an anti-collision process.

Type A: A data frame is identified by Start-of-Frame (SoF) and End-of-Frame (EoF) symbols and may contain multiple bytes. Each byte of data is followed by an odd-parity bit. Each frame also contains a 2-byte Cyclic Redundancy Check (CRC) that is appended at the end of the data. During the initialization phase, the token acts like a state machine. The reader then issues commands to change the state of the token as required. A simplified Type A state machine is shown in Fig. 13.9. The reader uses the following commands:

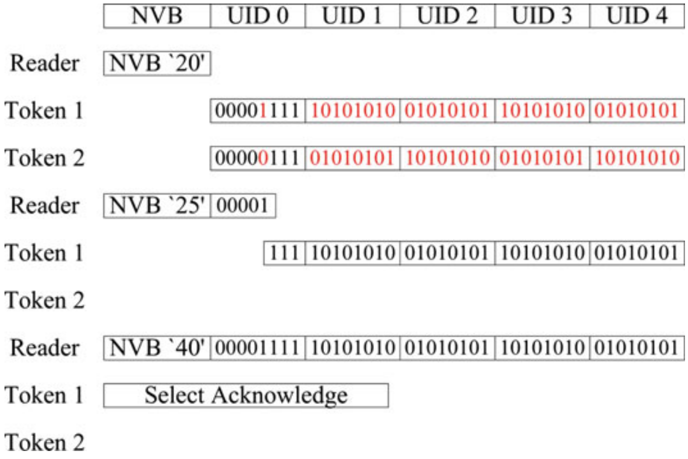
- *REQA/WUPA*: The Type A Request (*REQA*) and the Type A Wake-Up (*WUPA*) commands are periodically sent by the reader to poll its interrogation field for tokens. The *WUPA* command can also be used to put tokens that have entered the HALT state into the READY state.

Fig. 13.9 Type A: token state machine



- *SELECT (NVB < 40)*: If the Number of Valid Bits (NVB) is less than the number of bits in the Unique Identifier (UID), then the *SELECT* command is used for anti-collision; thus, it is also referred to as the *ANTICOLLISION* command.
- *SELECT (NVB = 40)*: When the reader has determined the unique identifier of the token, it wishes to communicate with it using the *SELECT* command to place that token in the ACTIVE state.
- *HLTA*: The Type A Halt (*HLTA*) puts the token into the HALT state.

When the token enters the interrogation zone of the reader, it powers up and enters the IDLE state. In this state, the token will not respond to any commands from the reader except *REQA* and *WUPA*, which will put it into the READY state. Readers will periodically transmit a *REQA* command to see whether there are tokens within its interrogation zone. If it receives a Type A Answer to Request (*ATQA*) response, it knows that a token is present and will proceed to the next step of initialisation. It is often the case that multiple tokens will be presented to the reader at once, e.g., travel and credit cards in the same wallet, so the standard must allow the reader to select a specific token. This process of selection is known as anti-collision, which in Type A is implemented using a binary search tree algorithm. The *SELECT* command is a bit-oriented frame containing the length of the current search (NVB) and a search pattern. If the least significant bits of a token's unique identifier match the search pattern for the specified search length, that token will respond with the rest of its unique identifier. An example of the anti-collision process is shown in Fig. 13.10. In the first step, the reader sets the search length to 0, which results in both tokens transmitting their whole 4-byte unique identifier and a checksum (BCC). The first collision that the reader detects is in the fifth bit. The reader therefore sets the search length to 5 and transmits a search pattern that indicates to the token whether the fifth bit should be a '1' or a '0'. Only the first token's unique identifier matches the search



	NVB	UID 0	UID 1	UID 2	UID 3	UID 4
Reader	NVB `20'					
Token 1		00001111	10101010	01010101	10101010	01010101
Token 2		00000111	01010101	10101010	01010101	10101010
Reader	NVB `25'	00001				
Token 1		111	10101010	01010101	10101010	01010101
Token 2						
Reader	NVB `40'	00001111	10101010	01010101	10101010	01010101
Token 1	Select Acknowledge					
Token 2						

Fig. 13.10 Type A: example of anti-collision sequence

string, so it alone responds with the rest of its identifier. Since the reader detects no collisions, it knows that it has identified a single token. Finally, the reader sends the same *SELECT* command with maximum search length and full unique identifier to which the tokens respond with a Select AcKnowledge (SAK). This also results in the token being put in the ACTIVE state.

For the anti-collision to work, the token's responses must be closely synchronised with the reader's commands. Type A expects the token to behave in a synchronous manner, so it prescribes a fixed bit grid which defines when a response must be sent. This grid is defined by specifying a Frame Delay Time (FDT) as follows: FDT $=$ $(n \cdot 128 + 84)/f_c$ if the last bit sent by the reader is a '1' and FDT $= (n \cdot 128 + 20)/f_c$ if the last bit sent by the reader is '0'. n is equal to 9 for *REQA*, *WUPA* and *SELECT* commands, while $n \geq 9$ for all other commands.

Type B: A data frame, marked with start-of-frame and end-of-frame symbols, contains multiple characters. Each character consists of a start bit, eight data bits, and a stop bit, which must be followed by a set Extra Guard Time (EGT) before the next character starts. Each frame also contains a 2-byte cyclic redundancy check that is appended at the end of the data.

As with Type A, the token acts like a state machine during initialisation. A simplified Type B state machine is shown in Fig. 13.11. The reader uses the following commands:

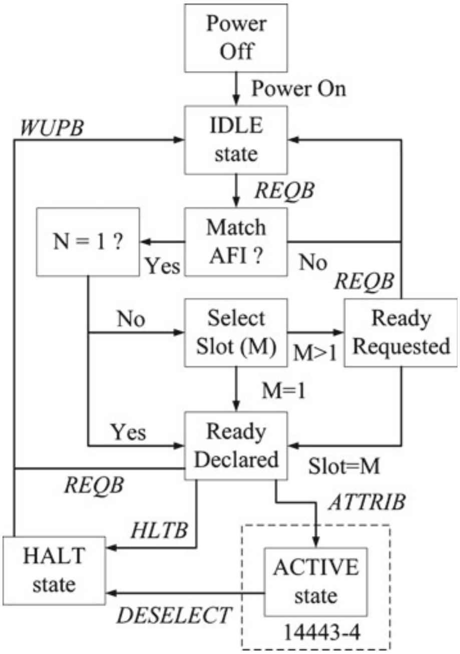


Fig. 13.11 Type B: token state machine

- *REQB/WUPB*: The Type B Request (*REQB*) and the Type B Wake-Up (*WUPB*) commands are periodically sent by the reader to poll its interrogation field for tokens and initiates the anti-collision procedure. The *WUPB* command can also be used to put tokens that have entered the HALT state into the IDLE state. *REQB* and *WUPB* commands contain an Application Family Identifier (AFI) that indicates the type of application targeted by the reader. This field is used to preselect tokens participating in the anti-collision since only tokens with an application of the type indicated by the AFI may answer with a Type B Answer To Request (*ATQB*).
- *SLOT-MARKER*: During anti-collision, the reader may send up to $N - 1$ *SLOT-MARKER* commands to indicate each time slot available for the tokens' *ATQB* responses. The commands can be sent after the end of an received *ATQB* message to mark the start of the next slot or earlier if no *ATQB* is received and it is known that the slot will be empty.
- *ATTRIB*: The *ATTRIB* command is used by the reader to select a single token. Upon receiving an *ATTRIB* command containing its identifier, a token enters the ACTIVE state where it only responds to commands defined in ISO/IEC 14443-4 that includes the Card Identifier (CID) assigned to it in the *ATTRIB* command parameters.
- *HLTB*: The Type B Halt (*HLTB*) puts the token into the HALT state.

When the token enters the interrogation zone of the reader, it powers up and enters the IDLE state. The anti-collision procedure, based on a dynamic slotted ALOHA algorithm, is started when the *REQB* command is transmitted. The token checks to see whether it has an application that matches the received AFI parameter and if this is the case it calculates the number of anti-collision slots N from the parameters in the *REQB* command. If $N = 1$, the token responds immediately with its *ATQB* response. Alternatively, the token is put in the READY REQUESTED state and randomly calculates a slot in which to send its response. In a probabilistic system, which does not use time slots, a token responds only if its randomly chosen slot is equal to 1. If it chooses any other slot, it returns to the IDLE state and waits for the anti-collision procedure to start again. In a pseudo-deterministic system that uses multiple slots, the tokens respond within the time slot corresponding to its randomly chosen slot. Once a token has responded with an *ATQB*, it is in the READY DECLARED state. The *ATQB* response contains information that the reader can use to identify and select the token. Once the reader has received a collision-free *ATQB* from the card it wants to select, it uses the *ATTRIB* command to place the chosen token into the ACTIVE state.

13.4.1.4 Part 4—Transmission Protocol

The final part of the standard specifies a half-duplex block transmission protocol for communication between the reader and the token. In Type A tokens, additional setup parameters need to be exchanged between the token and the reader to configure the protocol, such as frame size and card identifier. In Type B tokens, this is not necessary as these parameters have already been exchanged in the *ATQB* response and the *ATTRIB* command. When the Type A token is selected, the select acknowledgment will contain information about whether the token implements a proprietary protocol, or whether it supports ISO 14443-4. If a protocol adhering to ISO 14443-4 is available, the reader will transmit a Request for Answer To Select (*RATS*) command to which the token replies with an Answer To Select (*ATS*). This is followed by Protocol and Parameter Selection (*PPS*), if supported, that allows for the change of data rate between the token and the reader. The transmission protocol itself allows for the transmission of an Application Data Unit (APDU) that can contain any required data. The structure of the protocol is based on the $T = 1$ protocol specified in ISO 7816-3 for contact cards and is therefore often referred to as $T = CL$. This simplifies integration of contactless applications into smart card operating systems that are already available, especially in dual-interface tokens.

13.4.2 ISO 15693

ISO 15693, titled *Identification Cards—Contactless Integrated Circuit Cards—Vicinity Cards*, is most often implemented in systems using memory tokens for tracking or simple identification. Part 1 of the standard is very similar to the corresponding part in ISO 14443, so this section only discusses Parts 2 and 3.

13.4.2.1 Part 2—Air Interface and Initialization

This part of the standard describes the signal characteristics of the radio interface between the token and the reader, which allows for both power transfer and bidirectional communication. The token's power is provided by an alternating magnetic field at a frequency of 13.56 MHz and the reader must ensure that the magnetic field is within the range 115 mA/m $\leq H \leq$ 7.5 A/m. The operational range for systems using ISO 15693 can extend up to 1 m.

ISO 15693-2 defines both 'long distance' and 'fast' communication modes. In the 'fast' mode, the forward channel uses 1 of 4 pulse-position coding that is modulated onto the 13.56 MHz carrier with 100% ASK. One symbol comprises 8 time slots each of duration 9.44 μs and the modulation pulse can only be transmitted at an uneven time slot. The value n of the symbol can be determined by pulse slot $= (2 \cdot n) + 1$ and can be 0, 1, 2, or 3. One symbol takes 75.53 μs $=$ 8×9.44 μs to transmit, but each symbol can convey two bits of data, so the data rate is 26.48 kbps. For the 'long distance' forward channel data 1 of 256 pulse-position coding is used, which is then 10% ASK modulated onto the 13.56 MHz carrier. One symbol now comprises of 512 time slots and takes 4.833 ms to transmit. Since the symbol can have any value between 0 and 255, it can represent 8 bits of data so the effective data rate is 1.65 kbps. The backward channel uses Manchester coding, which can either be ASK modulated onto a 423 kHz sub-carrier or FSK modulated with a 423/485 kHz sub-carrier before being load modulated onto the 13.56 MHz carrier. For the 'long distance' mode, the data rate is 6.62 kbps, and for the 'fast' mode, the data rate is 26.48 kbps. Data is transmitted in frames marked by start-of-frame and end-of-frame symbols.

13.4.2.2 Part 3—Anti-collision and Transmission Protocol

The final part of the standard provides details on the anti-collision procedure, token initialisation, and possible commands. It should be noted that a large section of the guidelines given in this part is optional. As in ISO 14443 the token acts like a state machine during initialisation. A possible state machine diagram, given as an example in the standard, is shown in Fig. 13.12. The standard's transmission protocol specifies the format for command requests and responses, including a format for a number of optional commands such as *SELECT*, *RESET*, *READ* and *WRITE*. The standard only specifies two mandatory commands:

- *INVENTORY*: The token shall perform anti-collision when receiving the *INVENTORY* request.
- *STAY-QUIET*: If the token receives the *STAY-QUIET* request, it shall enter the QUIET state.

When the token enters the interrogation zone of the reader, it enters into the READY state. The reader will poll for tokens by transmitting an *INVENTORY* request. If a token is present, it will participate in the anti-collision procedure. The token compares the mask value from the *INVENTORY* request, varying in length from 0 to 8 bytes, with the corresponding least significant bits in its 64-bit unique

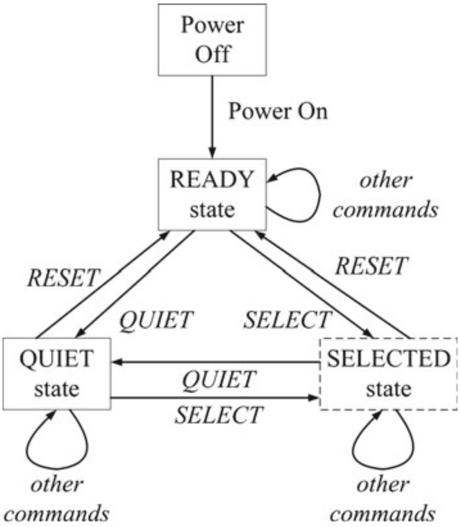


Fig. 13.12 ISO 15693: possible token state machine

identifier. If it is a match, the token will reply with its identifier during one of sixteen slots marked by the reader. It is assumed that the token does not have the necessary resources to randomly choose a slot. It therefore uses the four least significant bits not compared with the mask to determine the slot number. For example, if the 2 least significant bytes of the tokens UID is *4FAC* and the mask was *FAC*, then the token will respond in slot number 4. Once the reader has the token's unique ID, it can put the token in the QUIET state, or in the SELECTED state, if supported, where further commands can be issued.

13.4.3 ISO 18000

ISO 18000, titled *Information Technology AIDC Techniques—RFID for Item Management—Air Interface*, defines a generic structure for use in item management applications (Part 1), along with air interfaces for operation at 135 kHz (Part 2), 13.56 MHz (Part 3), 2.45 GHz (Part 4), 5.8 GHz (Part 5), 860–930 MHz (Part 6), and 433 MHz (Part 7). It is expected that all the parts of ISO 18000 will still be revised to include fixes and allow for the extra capabilities, such as active tokens and sensors [22]. This section only gives a brief overview of the standard in comparison with Part-2 of the ISO 14443 and ISO 15693 standards.

13.4.3.1 Part 3—Parameters for Air Interface Communications at 13.56 MHz

ISO 18000-3 defines a physical layer, collision management system, and protocol values, in accordance with ISO 18000-1, for RFID systems operating at 13.56 MHz. It specifies two modes of operation intended for use with different applications. Mode 1 is based on ISO 15693 with additional specifications to allows for item management applications and improved vendor compatibility. The reader to token data rate is 1.65, or 26.48 kbps. The token-to-reader data rate is 26.48 kbps. A protocol extension allows for data rates of 53 and 106 kbps. Mode 2 specifies a high-speed interface where the reader to tag data rate is 423.75 kbps.

13.5 Near-Field Communication

Near-Field Communication (NFC) is a contactless technology which enables electronic devices to communicate with each other by bringing them into proximity to a distance of several centimeters or less. NFC was initially established by Sony NXP (previously Philips Semiconductors) and Nokia in 2002, with the purpose of integrating RFID technology into mobile and embedded devices. NFC operates at the same frequency (13.56 MHz) and the radio interface shares several properties with existing HF RFID systems. Compared with RFID technology, NFC's bidirectional communication ability is ideal for establishing connections with other technologies by a simple touch, while RFID is usually used for a reader identifying or authenticating a tag which acts as a transponder. In addition, as a subset of HF RFID, NFC has the advantage of the short read range limitations of its radio frequency, which makes it a popular choice for intuitive communication between consumer devices, such as smartphones. Given its compatibility with existing systems, and offering new application possibilities, NFC has been widely deployed in our daily life. The following are some examples of the applications of NFC:

- *Ticketing*: Using our smartphone as a travel card, NFC works with most contactless smart cards and readers, which means it could be integrated into the transportation systems. For example, in 2008, German rail operator Deutsche Bahn launched an NFC-ticketing pilot program in which the travellers only needed to touch their phones to an NFC tag when boarding the train and then to another when getting off [23]. After that more and more countries choose to deploy NFC-enabled devices in public transportation. Among them, Iran trailed NFC-ticketing systems in 2012 [24].
- *Sharing*: two smartphones supporting NFC can share information with each other such as business cards. For example, Google announced NFC-based Android beam for sharing between phones in 2011 [25].
- *Service Discovery*: advertisers and marketers can use NFC chips in posters to promote their information, and people only need to touch their phones to these 'smart' posters to obtain the information.
- *Payment*: NFC works in a short range (less than 10 cm), which makes it a good choice for secure transactions such as contactless credit card payments. Examples include Google Wallet [26] and Apple Pay [27].

13.5.1 Standards—ISO 18092

ISO 18092, which is also referred to as NFCIP-1 (*Near-Field Communication Interface and Protocol*) or ECMA 340, specifies an RF interface and transmission protocol for communication between two inductively coupled devices operating at a frequency of 13.56 MHz. This standard allows for an active device, such as a mobile phone or PDA, to access RFID applications by acting as a reader, or a passive token, and it can also be used for short-range peer-to-peer communication. Although relatively new, NFCIP has strong support from industry. A NFCIP device can operate in three different ways, and it has more resources than a passive token, thereby allowing it to interface with its environment in a number of ways. The standard itself and the additional specifications for data exchange formats, record types, and compatible tokens are therefore quite comprehensive. This section only briefly discusses the different modes of operation and the sections of the standard corresponding to Part-2 of the

ISO 14443 and ISO 15693 standards. The reader is encouraged to read the ECMA 340 documentation [28] or visit the NFC Forum [29] for more details.

This standard defines 'active' and 'passive' communication modes between two entities, referred to as the target and the initiator. In active mode, both the initiator and the target generate a RF field. The initiator will start communication and transmit data by modulating its own carrier. Once it has finished transmitting, it will switch off the carrier and wait for a response. This is similar to two readers transmitting data to one another. The target then switches on its carrier and transmits a response. In the passive communication mode, only the initiator generates an RF field and starts the communication by modulating data on its own carrier. The target then responds to the initiator using a load modulation scheme. In this mode, one device acts like a reader while the other device emulates a passive token. Each device must ensure that it generates a magnetic field, at a frequency of 13.56 MHz, that is, within the range 1.5 A/m $\leq H \leq$ 7.5 A/m. The operational range for NFCIP systems is in the order of a few centimetres.

Both active and passive modes are defined for communication rates of 106, 212, and 424 kbps. The method for transmitting data at 106 kbps in passive mode is the same as for ISO 14443A. 106 kbps data transmission in active mode uses the same modulation scheme as the forward channel in ISO 14443A. For 212 and 424 kbps 'passive' and 'active' modes, both the forward and backward channel use Manchester code that is ASK modulated onto the 13.56 MHz carrier with a modulation index of 8–30% [7]. It should be noted that the backward channel does not use sub-carrier modulation. Given that NFCIP technology was initially developed by Nokia, NXP and Sony, the lower layer communication is compatible with NXP Mifare (106 kbps) and Sony FeliCa (212/424 kbps) products.

13.5.2 NFC Forum Specifications

To promote implementation and standardisation of NFC technology, the NFC Forum [29] is created by NXP, Sony, and Nokia at 2004, currently owning over 190 member companies. The NFC Forum has developed various technical specifications which can help achieve full interoperability between existing technologies and devices to enable new services. There are five types of specifications in total.

- *Protocol Technical Specifications*: This type of specifications deals with communication protocol between NFC devices. It includes Logical Link Control Protocol (LLCP) Technical Specification, Digital Protocol Technical Specification, Activity Technical Specification, Simple NFC Data Exchange Format (NDEF) Exchange Protocol Specification, and Analog Technical Specification. The details of the functionalities of each specifications can be found at [29].
- *Data Exchange Specifications*: This type of specification consists of the NDEF Technical Specification, which specifies a common data format for NFC Forum-compliant devices and tokens.

- *NFC Forum Tag-Type Technical Specifications*: This type of specification defines types of NFC tokens, which can be divided into four tag types. The four different tag operation specifications provide the technical information that is needed to implement the NFC devices. All the four types of tag are based on existing contactless products. Both Type 1 tag and Type 2 tag are based on ISO/IEC 1443A, and they can be read and rewrite capable. The memory availability of Type 1 and Type 2 tags is from 96 bytes to 2K bytes, and from 48 bytes to 2K bytes, respectively. The Type 3 tag, also known as FeliCa, is based on the Japanese Industrial Standard X6319-4. The theoretical memory availability is up to 1M bytes. Type 4 tags are fully compatible with the ISO/IEC 14443 standard series. Their memory availability is up to 32K bytes.
- *Record-Type Definition Technical Specifications*: This type of specification provides the technical specifications for Record-Type Definitions (RTD) and four specific RTDs including Text, URL, smart poster, and generic control. These specifications specify the format and rules for building standard record types and allow users to create their own applications based on these NFC Forum specifications. Readers can refer to [29] for the details.
- *Reference Application Technical Specifications*: This type of specification specifies the reference applications using NFC. Examples given in the NFC Forum include Connection Handover Technical Specification and Personal Health Device Communication Technical Specification. The former enables developers to define information needed for the connection setup to be carried in NFC Data Exchange Format messages. The latter supports an interoperable data transport for personal health devices conforming to the ISO/IEEE Standard 11073-20601 Optimized Exchange Protocol and NFC Forum specifications.

13.5.3 Mobile NFC Architecture

NFC-enabled mobile devices all have the three required components of the basic NFC architecture. The three basic components are the application execution environment (the normal execution and storage area of the device), the NFC module (radio frontend), and the secure element (trusted, secure execution and storage area). Figure 13.13 shows an example of a secure element and NFC module, together with an integrated antenna, in a phone circuit layout.

13.5.3.1 NFC Module

The NFC module handles the underlying communication functions. As introduced in Sect. 13.5.1, there are active and passive communication modes for NFC. In active mode, both initiator and target devices communicate by alternately generating their own fields, while in passive mode, the target device draws power through electromagnetic induction from the carrier field provided by the initiator device. Both active

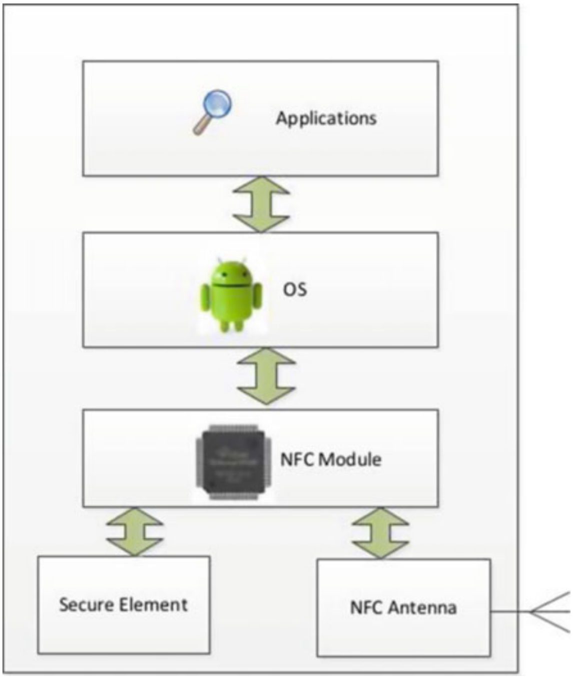
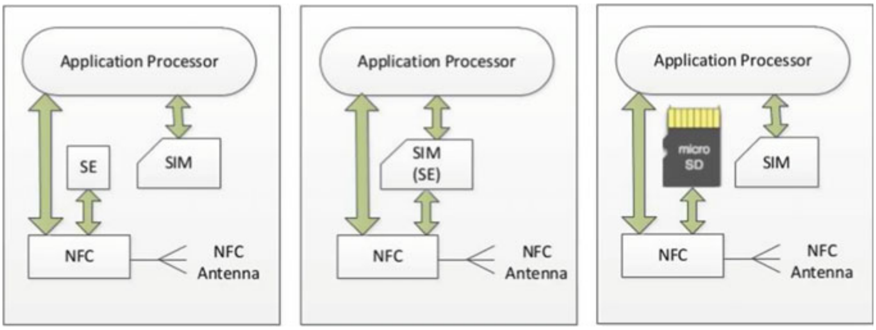


Fig. 13.13 Mobile NFC architecture

and passive modes are defined for communication rates of 106, 212, or 424 kbps. NFC-enabled devices can work in three operation modes: card emulation mode, peer-to-peer mode, and reader/writer Mode.

- *Card Emulation Mode*: In this mode, NFC-enabled devices act like smart cards, allowing users to perform transactions such as ticketing and purchases.
- *Peer-to-Peer Mode*: This mode enables two NFC-enabled devices to interact with each other such as exchanging information and sharing files.
- *Reader/Writer Mode*: In this mode, NFC-enabled devices can read information stored on inexpensive NFC tags embedded in smart posters and displays.

The NFC module has multiple interfaces to the Secure Element (SE), including ETSI 613 SWP (Single Wire Protocol) [30] and ETSI 622 HCI (Host Controller Interface) [31]. SWP is an interface between the Contactless Frontend (CLF) and the UICC (SIM card chip), only running on data link layer for communication. The CLF acts as a master and the UICC as a slave which can be powered by the CLF. HCI provides the network, transport, and session layers of the logical interface which runs over SWP. It supports multiple SEs controlled by a single NFC controller. An SE running on an application is referred to as a host while the NFC controller is referred to as a host controller.



(a) Embedded SE. (b) SIM as SE. (c) MicroSD SE

Fig. 13.14 Three basic SE architectures

13.5.3.2 Secure Element (SE)

The SE is for NFC-enabled mobile devices to perform secure transactions and store sensitive data in a trusted environment. It provides a secure means to establish trust between service provider and the device. Currently, SEs are commonly deployed in mobile devices with three different architectures, or a hybrid architecture that is a combination of the basic three architectures.

- *Embedded SE*: In this architecture as shown in Fig. 13.14a, the SE exists as an independent embedded hardware module (i.e., a stand-alone integrated chip) and is built into the phone. It could be seen as a mobile Trusted Platform Module (TPM) and, at the time of writing, is only used as card emulation.
- *SIM SE*: In this architecture, the SE is implemented within the existing SIM card (see Fig. 13.14b). There is no extra hardware. There are two variations: DIF-SIM where all functionality is on SIM with antenna in phone, and SIM-Flex where all functionality is on SIM with attached antenna.
- *micro-SD SE*: In this architecture, the SE is built in a removable memory component such as a micro-SD memory slot as shown in Fig. 13.14c. It can be added to any handset with an SD slot.

Although the SE is for providing secure NFC transactions, there are also some problems that need to be addressed. First, only card emulation mode really uses the SE, i.e., if the phone is acting as token, then the application is in the SE but if acting like a reader or in peer-to-peer mode, the SE is mostly not involved. Second, there is no standard security specifications, i.e., security is application specific. Finally, the SE memory is limited, since it is essentially a smart card. In addition, the card emulations are meant to run in an attack/tamper-resistant SE chip, and there are business/practical problems for service providers to get access to SEs, which inhibits services. To solve these problems, a solution first offered by BlackBerry and later by Android is Host Card Emulation (HCE).

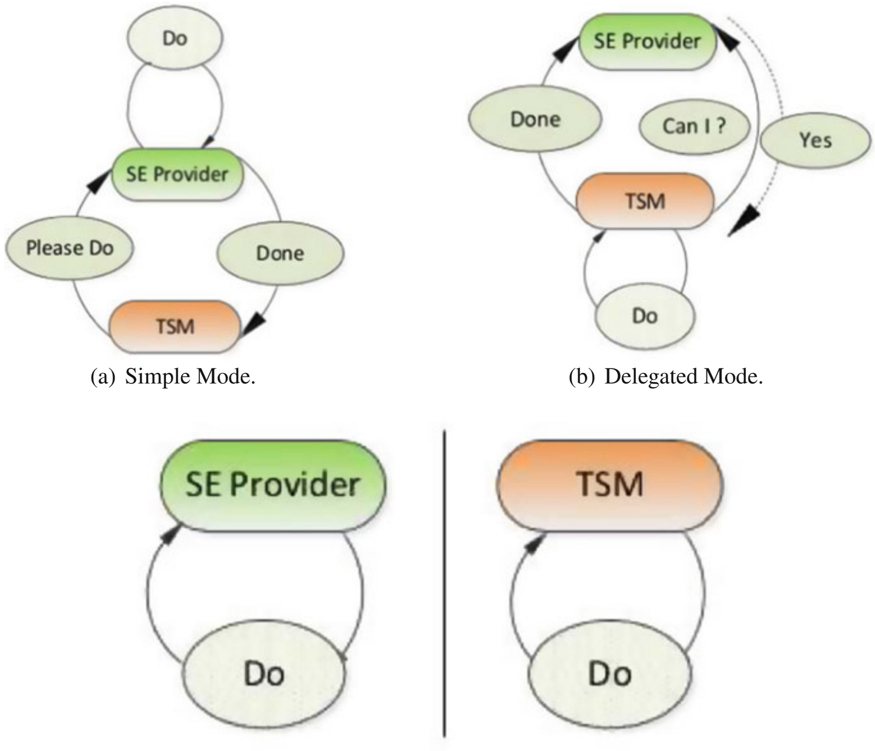
HCE is effectively a means to have a software SE running on the main phone processor, without the need for a physical secure element embedded as hardware in the phone. Instead of putting the SE inside the mobile device, HCE enables the SE to be placed in the phone application area or in a remote and hosted cloud environment and provides exact virtual representation of electronic identity cards using only software. In particular, with HCE, mobile applications can run on supported operating systems which offer payment card and access card solutions independently of third parties.

The term 'Host Card Emulation' was first introduced by Doug Yeager and Ted Fifelski, the founders of SimplyTapp, Inc., in 2012 [32] to describe the ability to provide a communication channel between a contactless payments terminal and a remotely hosted secure element which contains financial payment card data, allowing financial transactions to be conducted at a Point of Sale (POS) terminal. The first-known mobile handset to support HCE was the BlackBerry Bold 9900. With the release of Android 4.4, Google introduced a platform for secure NFC-based transactions through HCE, for payments, loyalty programs, card access, transit passes, and other custom services [33]. After the adoption by Google, HCE has also obtained support from many other big parties such as Visa, MasterCard and Microsoft. For normal users, however, HCE on Android is only available for application development when using an aftermarket Android release, such as *cyanogenmod* [34].

13.5.4 Basic Deployment Modes

NFC applications need to be securely placed into the secure element of devices OTA (Over the Air). This is different from a smart card. For example, the detailed information will be put onto the card by the smart card provider during the card personalisation once the bank arranges a bank card for a specific person. However, if a person buys a phone and then wishes to use it to pay, then the bank needs to put their application onto the phone while the phone is not under their control. The same trusted methods to personalise smart cards are therefore not available for NFC devices. For NFC-based applications, there are three different deployment modes [35].

- *Simple Mode*: as shown in Fig. 13.15a, it is an issuer-centric model, where the Trusted Service Manager (TSM) requests the SE provider to perform a Card Content Management (CCM) operation and return the execution result to the TSM.
- *Delegated Mode*: in this mode, CCM is delegated to the TSM. Each operation requires preauthorisation from the SE provider. The execution result of the operation is optionally sent to the SE provider.
- *Dual Mode*: in this mode shown in Fig. 13.15c, CCM is fully delegated to the TSM on a dedicated area of the SE. Dual mode is characterised by the presence of at least two security domains with the authorised management privilege in the secure element.



(a) Simple Mode.

(b) Delegated Mode.

(c) Dual Mode

Fig. 13.15 Basic deployment modes [35]

13.5.5 NFC Security

Although NFC helps to enable various new services, it is faced with serious security issues. It is possible to unlock the SE in the embedded SE architecture mode. In particular, NFC is actually a nice attack platform for relay attacks, especially with HCE since an HCE can run any user application and does not need to be unlocked to be used for user-generated applications. NFC allows an attacker to have a programmable card for cloning applications and act as a proxy reader and proxy token in relay. In the following sections, a relay attack [36] and a cloning attack are illustrated that used NFC-enabled mobile devices as an attack platform.

13.5.5.1 Relay Attack

In a normal NFC communication system, two devices are in close physical proximity up to 10 cm. However, in a relay attack, the communication can be relayed

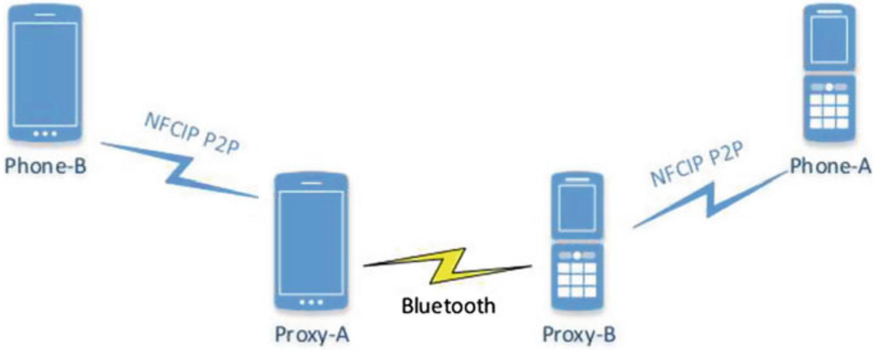


Fig. 13.16 Relay attack in NFC-enabled mobile devices communication [36]

over an extended distance by placing a proxy device within communication range of each legitimate participant and then forwarding the communication using another communication channel. The two legitimate participants receive valid transmissions from each other and therefore assume that they are in close physical proximity communicating with each other.

Figure 13.16 shows the relay attack against two NFC-enabled mobile phones operating in peer-to-peer mode and participating in a legitimate transaction demonstrated by Francis et al. [36]. The attack functionality can be implemented using only software which is written into the mobile devices Proxy-A and Proxy-B via publicly available APIs in a standard Mobile Information Device Profile (MIDP) using JSR 118 API [37]. In the presented experiment, Phone-A and Phone-B communicate with each other with the mobile-based proxy platforms, i.e., Proxy-A and Proxy-B. Proxy-B acts as the server and Proxy-A acts as the client, both of them establishing the relay channel using Bluetooth. Phone-A sends a message which is received by Proxy-B and relayed to Proxy-A over the Bluetooth data bearer. Then, Proxy-A transfers the payload onto Phone-B as well as relays the response from Phone-B to Proxy-B. Finally, Proxy-B transmits the response to Phone-A. In effect, Phone-A is made to believe that the response message is originating from Proxy-B while actually it is from Phone-B. Similarly, Phone-B is made to believe that the message is from Proxy-A which is actually from Phone-A. Fortunately, the authors in [36] proposed a countermeasure against the relay attack by using location information. By using NFC-enabled devices, it is also possible to relay communication between a normal contactless token and reader [38].

13.5.5.2 Cloning Attack

The first-generation contactless cards have rudimentary security mechanism, i.e., card is authenticated based on static data. In this case, the attacker can develop a mobile pick pocketing tool running on a NFC-enabled mobile phone and read data off the card. Then, the attacker can clone the card and use payment-reserved AID (Application ID) to communicate with a POS for a payment transaction. As a practical example, Roland and Langer [39] presented a cloning attack on EMV contactless payment cards where the adversary can create functional clones of a card that contains necessary credit data as well as preplayed authorisation codes. The card clones can then be used to perform a limited number of EMV Mag-Stripe transactions at any EMV contactless payment terminal.

13.6 Conclusion

Contactless tokens act as an electronic credential, interacting with the rest of the system on behalf of its owner. Contactless operation can increase productivity and convenience, while it offers additional benefits such as reduced maintenance cost and extended product lifetime. A choice of industry standards and tokens also allows this technology to be tailored to many applications. As a result, applications for contactless tokens have extended beyond access control and fare collection, with contactless tokens starting to replace, or supplement, established technologies such as paper tickets, barcodes, and magnetic stripe cards. This trend is expected to continue, especially in the identification and payment sectors, and with the introduction of NFC-enabled devices, the immediate future of contactless technology looks promising.

It is recommended that anybody wishing to learn more about contactless technology should consult literature by the Smart Card Alliance [11] and the '*RFID Handbook*' by Finkenzeller [18]. A number of open source RFID projects also provide hardware and software thatcan facilitate better understanding by means of practical experimentation [40–42].

References

1. R. Das. *RAIN RFID 2015–2020: Market size, growth opportunities and trends*. IDTechEx, 2015.
2. H. Stockman. *Communication by Means of Reflected Power*. Proceedings of the IRE, pp 1196–1204, October, 1948.
3. *The History of RFID Technology*. RFID Journal, December, 2006. http://www.rfidjournal.com/article/articleview/1338/1/129/.
4. Smart Card Alliance. *Contactless Technology for Secure Physical Access: Technology and Standards Choices.* Publication No. ID-02002, October, 2001.
5. ISO/IEC 14443. *Identification cards – Contactless integrated circuit cards – Proximity cards.*, 2011.
6. ISO/IEC 15693. *Identification cards – Contactless integrated circuit cards – Vicinity cards.*, 2009.
7. ISO/IEC 18092 (ECMA-340). *Information technology–Telecommunications and information exchange between systems–Near Field Communication – Interface and Protocol (NFCIP-1)*, 2013.

8. 2014 FIFA World Cup with RFID Ticketing Presents Cautionary Tale. June, 2014. https://www.tsl.com/2014/06/2014-fifa-world-cup-rfid-ticketing-presents-cautionary-tale/.
9. Federal Information Processing Standards. *Publication 201-1: Personal Identity Verification (PIV) of Federal Employees and Contractors*. March, 2006.
10. Federal Information Processing Standards. *Publication 201-2: Personal Identity Verification (PIV) of Federal Employees and Contractors*. August, 2013.
11. Smart Card Alliance. http://www.smartcardalliance.org/.
12. International Civil Aviation Organization (ICAO). *Document 9303 Machine Readable Travel Documents (MRTD). Part I: Machine Readable Passports, Seventh Edition*, 2015.
13. ISO/IEC 7501 *Identification cards – Machine readable travel documents*, 2008.
14. Bundesamt für Sicherheit in der Informationstechnik. *Advanced Security Mechanisms for Machine Readable Travel Documents' Extended Access Control (EAC).* Technical Guideline TR-03110, September, 2007.
15. T.S. Heydt-Benjamin, D.V. Bailey, K. Fu, A. Juels and T. O'Hare. *Vulnerabilities in first-generation RFID-enabled credit cards.* Technical report, University of Massachusetts Amherst, October 2006.
16. EMVCo. *EMV Contactless Communication Protocol Specification v2.5*. June, 2015.
17. EMV Integrated Circuit Card Specifications for Payment Systems-Book 1: Application Independent ICC to Terminal Interface Requirements. v4.3, November, 2011. https://www.emvco.com/specifications.aspx?id=223.
18. K. Finkenzeller, *RFID Handbook: Radio-frequency identification fundamentals and applications*, Wiley, 1999.
19. J.G. Proakis. *Digital Communications*, 5th Edition, McGraw-Hill, 2007.
20. J.G. Proakis and M. Salehi. *Communication Systems Engineering*, 2rd Edition, Prentice-Hall, 2002.
21. ISO/IEC 18000. *ISO/IEC 18000 Information Technology AIDC Techniques-RFID for Item Management – Air Interface*, 2010.
22. Institute for Prospective Technological Studies. *RFID Technologies: Emerging Issues, Challenges and Policy Options*, Technical Report EUR 22770 EN, 2007.
23. Marcus Gemeinder. *Touch & Travel C NFC based automatic fare collection using a passive infrastructure*. NFC Forum: Transport and City Life focus group, 2008.
24. Irancell demonstrates NFC payments and ticketing. http://www.nfcworld.com/2012/01/13/312430/irancell-demonstrates-nfc-payments-and-ticketing/, 2012.
25. Google announces NFC-based Android Beam for sharing between phones (video). http://www.engadget.com/2011/10/18/google-announces-nfc-based-android-beam-for-sharing-between-phon/, 2011.
26. Google Wallet. https://en.wikipedia.org/wiki/Google_Wallet.
27. Apple Pay. https://en.wikipedia.org/wiki/Apple_Pay.
28. ECMA-340 *Near Field Communication Interface and Protocol (NFCIP-1)* 3nd Edition, June, 2013. http://www.ecma-international.org/publications/standards/Ecma-340.htm.
29. NFC Forum. http://www.nfc-forum.org/.
30. ETSI TS 102 613-V11.0.0. *Smart cards: UICC– contactless Front-end (CLF) Interface; Part1: Physical and data link layer characteristics.* September, 2012.
31. ETSI TS 102 622-V11.1.0 *Smart cards: UICC– Contactless Front-end (CLF) Interface; Host Controller Interface (HCI).* October, 2012.
32. SimplyTapp proposes secure elements in the cloud. http://www.nfcworld.com/2012/09/19/317966/simplytapp-proposes-secure-elements-in-the-cloud/, 2012.
33. Host-based Card Emulation. https://developer.android.com/guide/topics/connectivity/nfc/hce.html.
34. CyanogenMod. https://en.wikipedia.org/wiki/CyanogenMod.
35. GlobalPlatforms Proposition for NFC Mobile: Secure Element Management and Messaging. Online whitepaper, GlobalPlatform, April, 2009.
36. Lishoy Francis, Gerhard Hancke, Keith Mayes and Konstantinos Markantonakis. *Practical NFC Peer-to-Peer Relay Attack Using Mobile Phones*. In: Radio Frequency Identification: Security and Privacy Issues, pages 35–49, LNCS, Springer, 2010.

37. Sun Microsystems, JSR-000118 Mobile Information Device Profile 2.0, https://jcp.org/aboutJava/communityprocess/final/jsr118/index.html, 2010.
38. Lishoy Francis, Gerhard Hancke, Keith Mayes and Konstantinos Markantonakis. *Practical Relay Attack on contactless Transactions by Using NFC Mobile Phones*. In: Journal of Radio Frequency Identification System Security RFIDsec'12 Asia Workshop Proceedings, pages 21–32, IOS Press, 2013.
39. Michael Roland and Josef Langer, *Cloning Credit Cards: A Combined Pre-play and Downgrade Attack on EMV Contactless*. In 7th Usenix Workshop on Offensive Technologies, Washington, D.C., 2013.
40. OpenPCD Project. http://www.openpcd.org.
41. rfdump Project. http://www.rfdump.org/.
42. rfidiot Project. http://www.rfidiot.org/.