RFID Design Principles

For a complete listing of titles in the
Artech House Microwave Library,
turn to the back of this book.

RFID Design Principles

Harvey Lehpamer

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the U.S. Library of Congress.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

Cover design by Yekaterina Ratner

ISBN 13: 978-1-59693-194-7

685 Canton Street
Norwood, MA 02062

10 9 8 7 6 5 4 3 2 1

To my lovely and very patient wife, Monica

Contents

1

Introduction

Radio-frequency identification (RFID) is an emerging technology and one of the most rapidly growing segments of today's automatic identification data collection (AIDC) industry. However, this emerging technology is not new; in fact, it is currently being used in numerous applications throughout the world. It was originally implemented during World War II to identify and authenticate allied planes, in an identification system known as Identification, Friend or Foe, and is still being used today for the same purposes.

RFID usage is steadily increasing, and companies across many industries are now looking at RFID to streamline operations, meet regulatory requirements, and prevent the introduction of counterfeit product into the supply chain to protect both consumer safety and company profitability. Industry experts view RFID not as competition with, but as a complement to barcode technology; in many cases, such as tracking pallets, cartons, and cases in a warehouse, both technologies are used. RFID technology, in fact, overcomes certain limitations found in some barcode applications. Because it is not an optical technology like bar coding, no inherent line of sight is required between the reader and the tagged RFID object. In addition, RFID transmits data wirelessly and is a read/write technology, so it can update or change the data encoded in the tag during the tracking cycle.

For an RFID project to be successful, it is necessary to approach any business problems that may arise and any potential RFID solutions by using a systems approach. In a design process, we need to look at all of the processes, be forward thinking, and think creatively about how to improve each operation. Implementing an RFID-based system is like implementing any new system: RFID systems should be conceived, designed, and implemented using a

systematic development process in which end users and specialists work together to design RFID systems based on the analysis of the business requirements of the organization.

One of the greatest obstacles to the wide adoption of any new technology is a standardization process. The purpose of standardization is to define the most efficient platform on which an industry can operate and advance. For example, standardization would address the question of how to ensure that a tag manufactured and installed in one part of the world will be readable and the product properly identified on the other side of the globe. Several organizations are involved in drafting standards for RFID technology, but in looking at the present status, it seems like it will be some time before all of the details are agreed on. Because RFID standardization is a very dynamic process, this book discusses only the standards that were current at the time of writing in a brief section and then provides readers with directions for pursuing further research.

Despite the considerable technical diversity of RFID technology, much of it is largely transparent to prospective users and much can be done to promote awareness of the technology's attributes without going into considerable technical detail. However, some basic technical knowledge is necessary for making an informed choice of products to meet particular application needs and to allow informed discussion among users, suppliers, systems integrators, and consultants.

This book introduces prospective users and system designers to the basics of RFID technology, including applications, benefits, technical characteristics, security and privacy, and standardization, design, and implementation of RFID's technical and economic challenges. As these technical, policy, and cost challenges are slowly mitigated, RFID will likely become the system of choice for global commerce.

Numerous issues beyond the detailed technical and sheer operational capabilities of RFID technology must be considered. Due to the large number of considerations that must be undertaken, only a few intangible and theoretical considerations, such as security, privacy, social, ethical, and future considerations, are presented in this book. In addition, this book mentions briefly a wide number of new and exciting topics and concepts, some of them, at least at this point, only marginally of interest to RFID, with the hope of piquing readers' interest in pursuing these new technologies.

This book should become a valuable resource to a wide spectrum of readers interested in exploring this new and exciting topic.

2

Comparison of Short-Range Communications Systems

2.1 Radio-Frequency Spectrum and Propagation

2.1.1 Radio-Frequency Propagation and Interference

Radio waves and microwaves are forms of electromagnetic energy we can collectively describe by the term *radio-frequency* or *RF*. RF emissions and associated phenomena can be discussed in terms of energy, radiation, or fields. We can define *radiation* as the propagation of energy through space in the form of waves or particles. Electromagnetic radiation can best be described as waves of electric and magnetic energy moving together (i.e., radiating) through space, as illustrated in Figure 2.1. These waves are generated by the movement of electrical charges, such as in a conductive metal object like an antenna. For example, the alternating movement of charge (i.e., the current) in an antenna used by a radio or television broadcast station or in a cellular base-station antenna generates electromagnetic waves. These waves that radiate away from the transmitting antenna are then intercepted by a receiving antenna, such as a rooftop TV antenna, car radio antenna, or an antenna integrated into a handheld device such as a cellular phone.

Continuous harmonic waves are typically sinusoidal in nature; thus, they are characterized by frequency, amplitude, and phase. They are also characterized by their three-dimensional shape. The energy radiated by any antenna is contained in a transverse electromagnetic wave (TEM) that is comprised of an electric and a magnetic field. These fields are always orthogonal to one another and orthogonal to the direction of propagation.

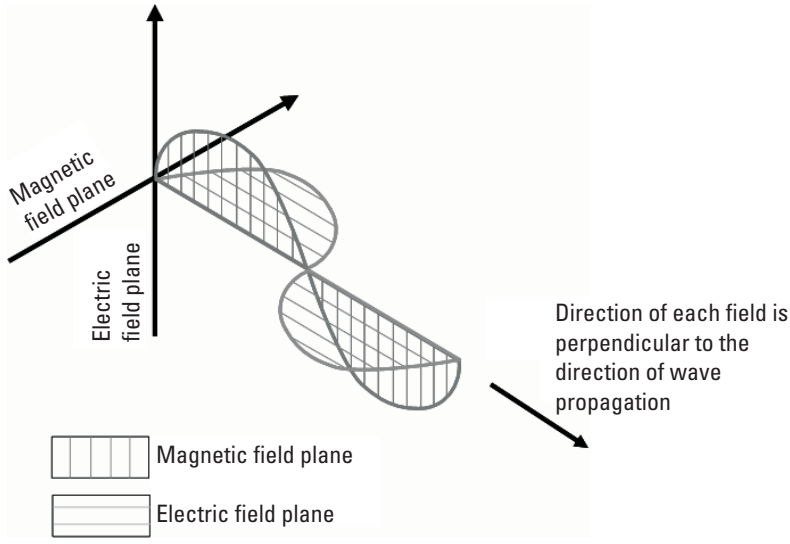


Figure 2.1 Electromagnetic wave.

The term *electromagnetic field* is used to indicate the presence of electromagnetic energy at a given location. The *RF field* can be described in terms of the electric and/or magnetic field strength at that location. Like any wave-related phenomenon, electromagnetic energy can be characterized by a *wavelength* and/or *frequency*. Frequency (f) equals the number of complete cycles occurring in 1 second; the wavelength (λ) is the distance an electromagnetic wave travels in the time it takes to oscillate through a complete cycle (1 period). Electromagnetic waves travel through space at the speed of light, and the wavelength and frequency of an electromagnetic wave are inversely related by a simple mathematical formula (2.1) connecting wavelength, speed of light (c), and frequency:

$$\lambda = \frac{c}{f} \tag{2.1}$$

Because the speed of light (3×10^8 m/s or 186,000 miles per second) in a given medium or vacuum does not change, we can see from (2.1) that the high-frequency electromagnetic waves will have short wavelengths and the low-frequency waves will have long wavelengths. The electromagnetic spectrum (the word *spectrum* literally means a range of values or a set of related quantities) includes all the various forms of electromagnetic energy from extremely low frequency (ELF) energy, with very long wavelengths, to X-rays and gamma rays, which have very high frequencies and correspondingly short wavelengths (Figure 2.2). Between these extremes are radio waves, microwaves, infrared

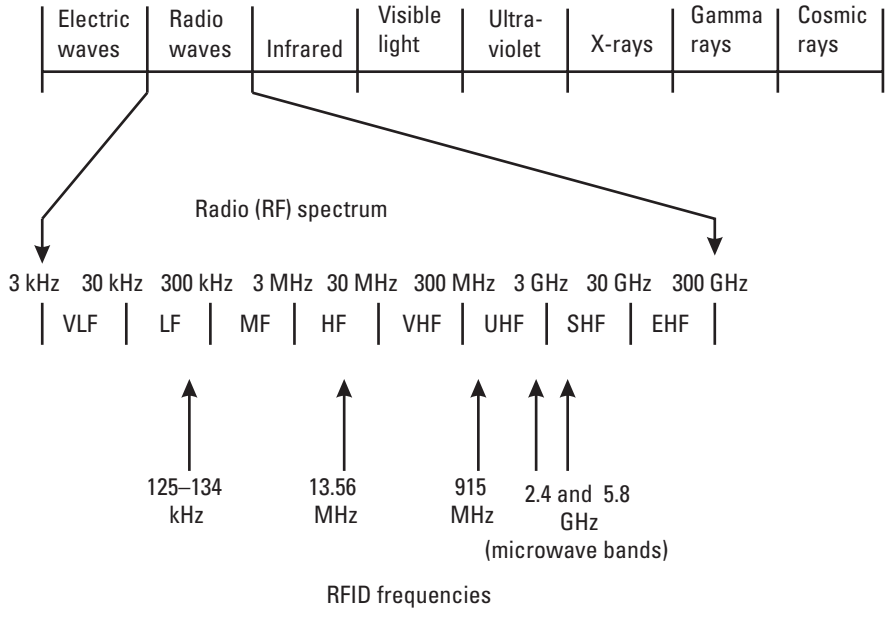


Figure 2.2 Electromagnetic spectrum.

radiation, visible light, and ultraviolet radiation, in that order. The RF part of the electromagnetic spectrum is generally defined as that part of the spectrum where electromagnetic waves have frequencies in the range of about 3 kHz to 300 GHz.

Why do we have different frequency bands? The answer is that different frequencies have different propagation characteristics. All frequencies are attenuated and reflected by materials to a greater or lesser degree, with the higher frequencies being more greatly attenuated than the lower frequencies. Low frequencies, such as the 125-kHz frequency, are attenuated very little as they propagate through materials. This allows them to have significant signal-penetration capabilities through all materials, including metal. When radiated and used in the far field, these frequencies can also have a significant communication range. For example, we listen to AM radio stations (typically operating between 580 and 1,700 kHz) that are being broadcast hundreds of miles away from us, whereas FM radio stations, typically operating between 88 and 108 MHz, have a range of maybe 20 miles.

For the regulations limiting RF emissions, the FCC distinguishes between intentional, unintentional, and incidental radiators. *Intentional radiators* are devices that intentionally emit RF energy, such as transmitters. *Unintentional radiators* are devices that intentionally generate RF energy for use only within the device or a cable system, but not for the purpose of radiation. Examples of

unintentional radiators are computer motherboards and receivers with local oscillators. *Incidental radiators* are devices that are not designed to generate RF energy at all, but for which RF radiation may occur as an unwanted side effect. Examples of incidental radiators are dc motors and mechanical switches.

As with all waves, electromagnetic waves interact with one another whenever they intersect at a point in space. Depending on the phase, amplitude, and polarization, intersecting waves may either constructively interfere or destructively interfere. This is one of the basic properties of linear waves. The observed wave at a point of intersection is the addition of all of the waves at that point. Constructive interference increases the amplitude of the detectable wave at that point. Destructive interference decreases the amplitude of the detectable wave.

Fundamental physics teaches us that at every boundary between two materials, electromagnetic waves incident upon that boundary will be both transmitted from one material to the other and reflected back into the material in which they are traveling. Conducting materials, such as metals, act similar to perfect reflectors for ultrahigh-frequency (UHF) radiation. Materials such as glass, concrete, and cardboard are effectively RF transparent for waves that are incident upon them with an angle of incidence of 90°, but they become less transparent as the angle of incidence becomes more oblique. Some materials, such as water, act as both good reflectors of electromagnetic waves and good attenuators, or absorbers, of electromagnetic energy. The partial reflection of a wave results in the energy of the wave being separated to traverse multiple paths. The result is that a partial reflection attenuates the partially transmitted wave by the amount of energy reflected at the boundary.

By passing through several materials and being reflected by several more, an electromagnetic wave traverses a path through the environment. In addition to attenuating the wave as it travels through the environment, the environment may impact the polarization of the wave. Two long parallel metal strips separated by a few inches, for example, will filter the UHF waves that are incident upon them by allowing waves that are polarized parallel to the metal strips to pass through the space between the strips, while waves polarized perpendicular to the metal strips will be reflected.

When two waves that have traversed different-length paths intersect at a point, they will be out of phase with one another. The phase difference is due to differences in the time required to traverse the different paths. Most phase differences cause destructive interference and may cause the observed wave at a point to appear to have a different frequency than what was originally transmitted.

Electromagnetic waves are linear, meaning that the wave experienced at a point in space and time is the sum of the waves that intersect at that point. Because of reflections, attenuation, and different path lengths (multipath) caused by objects in the environment, the waves that arrive at a point in space

may have amplitude that sums to zero or nearly zero. Passive RFID tags are not able to harvest sufficient operating power from low-amplitude and, hence, low-power locations. When these near-zero amplitude locations are surrounded by much higher amplitude locations where passive RFID tags are able to operate, the low-amplitude location is called a *null.* The position of nulls may be changed or the nulls may be eliminated by changing the position of the objects in the environment or changing the frequency being radiated by the antenna. When the environment is static, standing waves may result. This null phenomenon is also called a *standing wave.* The most common occurrence is when the two waves intersect each other exactly half a wavelength out of phase and completely cancel the signals. This creates the null spot where a tag would not be read.

2.1.2 Basic Antenna Parameters

Antennas are a very important component of communication systems. By definition, an antenna is a device used to transform an RF signal, traveling on a conductor, into an electromagnetic wave in free space. Antennas demonstrate a property known as *reciprocity,* which means that an antenna will maintain the same characteristics regardless of whether it is transmitting or receiving. Most antennas are resonant devices, which operate efficiently over a relatively narrow frequency band. An antenna must be tuned to the same frequency band of the radio system to which it is connected; otherwise, the reception and the transmission will be impaired. When a signal is fed into an antenna, the antenna will emit radiation distributed in space in a certain way. A graphical representation of the relative angular distribution of the radiated power in space is called a *radiation pattern.*

2.1.2.1 Polarization

Polarization is a physical phenomenon of radio signal propagation and refers to the orientation of the electric field vector in the radiated wave. If the vector appears to rotate with time, then the wave is elliptically polarized. The ellipse thus described may vary in ellipticity from a circle to a straight line, or from circular to linear polarization. So, in the general sense, all polarizations may be considered to be elliptical (Figure 2.3). For linear polarization (horizontal or vertical), the vector remains in one plane as the wave propagates through space. Linear polarization has two subcategories: vertical or horizontal, and right or left handed for circular situations.

The term used to describe the relationship between the magnitudes of the two linearly polarized electric field components in a circularly polarized wave is *axial ratio.* In a pure circularly polarized wave both electric field components

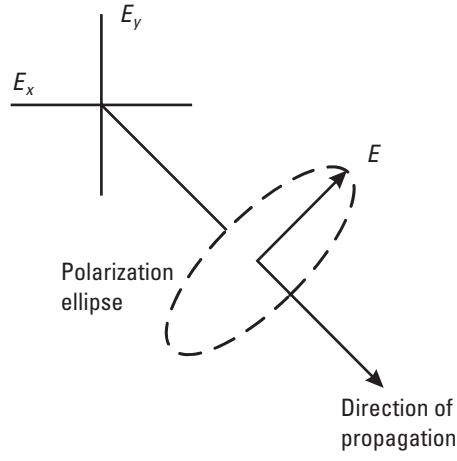


Figure 2.3 Elliptical polarization.

have equal magnitude and the axial ratio (AR) is 1 or 0 dB (10 log[AR]). In a pure linearly polarized wave the axial ratio is ∞.

Generally speaking, in most cases two antennas that form a link with each other must be set for the same polarization; intentional exceptions are made, however, that we will discuss here as well. When transmitting and receiving antennas are both linearly polarized, physical antenna misalignment will result in a polarization mismatch loss that can be approximated using the following formula:

$$\text{Polarization mismatch loss [dB]} = 20\log(\cos\theta) \tag{2.2}$$

where θ is the misalignment angle between the two antennas. For 15° we have a loss of 0.3 dB, for 30° we have 1.25 dB, for 45° we have 3 dB, and for 90° (orthogonal) we ideally have an infinite loss (no communications at all).

One of the common misconceptions regarding polarization relates to the circumstance in which one antenna in a transmit-to-receive circuit is circularly polarized and the other is linearly polarized [1]. It is generally assumed that a 3-dB system loss will result because of the polarization difference between the two antennas. In fact, the polarization mismatch loss between these two antennas will only be 3 dB when the circularly polarized antenna has an axial ratio of 0 dB. The actual mismatch loss between a circularly polarized antenna and a linearly polarized antenna will vary depending on the axial ratio of the circularly polarized antenna. When the axial ratio of the circularly polarized antenna is greater than 0 dB, this indicates that one of the two linearly polarized components will respond to a linearly polarized signal more so than the other component will. When a linearly polarized wave is aligned with the circularly polarized

linear component with the larger magnitude, the polarization mismatch loss will be less than 3 dB. When a linearly polarized wave is aligned with the circularly polarized linear component with the smaller magnitude, the polarization mismatch loss will be greater than 3 dB.

Assuming a 1-dB maximum axial ratio over the main beam, the signal loss due to polarization mismatch will be between 2.5 and 3.5 dB. We will see later that in RFID systems 3 dB will be used as an approximation and an average between the minimum and maximum polarization loss for a given axial ratio.

Multipath signals in RF systems arrive at the receiver's antenna via the reflection of the direct signal from nearby objects. If the reflecting objects are oriented such that they are not aligned with the polarization of the incident wave, the reflected wave will experience a polarization shift. The resultant or total signal available to the receiver at either end of the communications link will be the vector summation of the direct signal and all of the multipath signals. In many instances, a number of the signals arriving at the receiving site will not be aligned with the assumed standard polarization of the system antenna. As the receiving antenna rotates from vertical to horizontal, it simply intercepts or receives energy from these multiple signals. To improve or extend system performance, some system designers use receiving polarization diversity techniques in an effort to enhance signal reception. In these systems, a circularly polarized or dual linearly polarized antenna will be used at the receiving site to take advantage of the fact that many linearly polarized multipath signals, with different orientations, exist at the receiving site. These circular and dual polarized antennas theoretically have a better chance of receiving more total signal than a single, linearly polarized antenna.

Polarization of the waves in RFID systems becomes very important in tag antenna designs and deployments; antennas may be designed such that they efficiently capture and communicate with energy in one or a few different polarizations. If a reader antenna is linearly polarized and the tag antenna is linearly polarized, then the tag and the reader may communicate only when both antennas are oriented in the same linear direction. Circularly polarized antennas reduce the orientation requirements, but do not completely eliminate the orientation dependence for optimal performance.

It is difficult to predict the orientation of the electric field in the *near-field region* (i.e., very close to the antenna), because the transmitting antenna cannot be considered as a point source in this region. In the far-field region, the antenna becomes a point source, the electric and magnetic components of the field become orthogonal to the direction of propagation, and their polarization characteristics do not vary with distance. Most RF and microwave systems operate in the far-field region, with the exception of passive point-to-point microwave repeaters. Some RFID systems at lower frequencies operate in the near field and,

as we will see later, use coupling methods instead of the usual electromagnetic wave propagation.

2.1.2.2 Impedance Matching and Return Loss

Input Impedance

For an efficient transfer of energy, the impedance of the radio, of the antenna, and of the transmission cable connecting them must be the same. Transceivers and their transmission lines are typically designed for 50Ω impedance. If the antenna's impedance is different than 50Ω, then there is a mismatch and an impedance matching circuit is required. Of course, other impedances are also common as well.

Standing Waves and Voltage Standing Wave Ratio (VSWR)

In order for the antenna to operate efficiently, maximum transfer of power must take place between the transmitter and the antenna. Maximum power transfer can take place only when the impedance of the antenna is matched to that of the transmitter. According to the maximum power transfer theorem, maximum power can be transferred only if the impedance of the transmitter is a complex conjugate of the impedance of the antenna under consideration and vice versa. If the condition for matching is not satisfied, then some of the power may be reflected back, leading to the creation of standing waves, which can be characterized by a parameter called the voltage standing wave ratio (VSWR), normally written, for example, as 1.2:1 and pronounced as "1.2 to 1."

Another way to think about VSWR is that it is the amount of input power needed to get an equivalent of *unity* power out. In transmission line theory, the accepted definition is that standing waves are created by superposition of two waves traveling in opposite directions.

Because the standing wave ratio is not always calculated from the voltages, the V is sometimes dropped from VSWR, and the term is referred to as the standing wave ratio (SWR). VSWR and SWR are the same ratio and can be used interchangeably.

Return Loss

The return loss is another way of expressing impedance mismatch. It is a logarithmic ratio measured in decibels that compares the power reflected by the antenna to the power that is fed into the antenna from the transmission line. For a matched load, VSWR = 1. The relationship between VSWR and return loss is shown in the following equation:

$$\text{Return loss [dB]} = -20\log\left[\frac{(\text{VSWR}-1)}{(\text{VSWR}+1)}\right] \tag{2.3}$$

2.1.2.3 Bandwidth

The bandwidth of an antenna refers to the range of frequencies over which the antenna can operate correctly. The antenna's bandwidth is the number of hertz for which the antenna will exhibit a VSWR of less than 2:1. The bandwidth can also be described in terms of percentage of the center frequency of the band:

$$\mathrm{BW}[\%] = \frac{100(f_H - f_L)}{f_C} \tag{2.4}$$

where f_H is the highest frequency in the band, f_L is the lowest frequency in the band, and f_C is the center frequency in the band. In this way, bandwidth is constant relative to frequency. If bandwidth was expressed in absolute units of frequency, it would be different depending on the center frequency. Different types of antennas have different bandwidth limitations.

2.1.2.4 Directivity and Gain

Directivity is the ability of an antenna to focus energy in a particular direction when transmitting, or to receive energy better from a particular direction when receiving. In a static situation, it is possible to use the directivity capability of an antenna to concentrate the radiation beam in the given direction.

Gain is not a quantity that can be defined in terms of a physical quantity such as the watt or the ohm, but it is a dimensionless ratio. Gain is given in reference to a standard antenna. The two most common reference antennas are the isotropic antenna and the resonant half-wave dipole antenna. The isotropic antenna radiates equally well in all directions. Real isotropic antennas do not exist, but they provide useful and simple theoretical antenna patterns with which to compare real antennas. Any real antenna will radiate more energy in some directions than in others. Because it cannot create energy, the total power radiated is the same as that of an isotropic antenna, so in other directions it must radiate less energy. The gain of an antenna in a given direction is the amount of energy radiated in that direction compared to the energy an isotropic antenna would radiate in the same direction when driven with the same input power. Usually we are only interested in the maximum gain, which is the gain in the direction in which the antenna is radiating most of the power.

An antenna gain of 3 dB compared to an isotropic antenna would be written as 3 dBi. The resonant half-wave dipole can be a useful standard for comparing to other antennas at one frequency or over a very narrow band of frequencies. To compare the dipole to an antenna over a range of frequencies requires a number of dipoles of different lengths. An antenna gain of 3 dB compared to a dipole antenna would be written as 3 dBd.

2.1.2.5 Radiation Pattern

An antenna's *beamwidth* is usually understood to mean the half-power beamwidth. The peak radiation intensity is found and then the points on either side of the peak, which represent half the power of the peak intensity, are located. The angular distance between the half-power points is defined as the beamwidth. Half the power expressed in decibels is −3 dB, so the half-power beamwidth is sometimes referred to as the 3-dB beamwidth. Both horizontal and vertical beamwidths are usually considered. Assuming that most of the radiated power is not divided into sidelobes, then the directive gain is inversely proportional to the beamwidth (i.e., as the beamwidth decreases, the directive gain increases).

Sidelobes

No antenna is able to radiate all of its energy in one preferred direction. Some is inevitably radiated in other directions. The peaks are referred to as sidelobes, commonly specified in decibels below the main lobe.

Nulls

In an antenna radiation pattern, a null is a zone in which the effective radiated power is at a minimum. A null often has a narrow directivity angle compared to that of the main beam. Thus, the null is useful for several purposes, such as suppression of interfering signals in a given direction.

Front-to-Back Ratio

It is useful to know the front-to-back ratio, which is the ratio of the maximum directivity of an antenna to its directivity in the rearward direction. For example, when the principal plane pattern is plotted on a relative decibel scale, the front-to-back ratio is the difference in decibels between the level of the maximum radiation, and the level of radiation in an opposite direction (180°).

2.1.2.6 Antenna Modeling

The equations extracted from both the receiving and transmitting antenna models allow us to describe the behavior of a single bidirectional antenna. In both cases, antenna systems can be represented using Thevenin's equivalent circuits.

In the transmitting mode, an antenna system can be represented by a Thevenin circuit equivalent to that shown in Figure 2.4. In this figure, the antenna is represented by impedance Z_A given by:

$$Z_A = R_L + R_r + jX_A \tag{2.5}$$

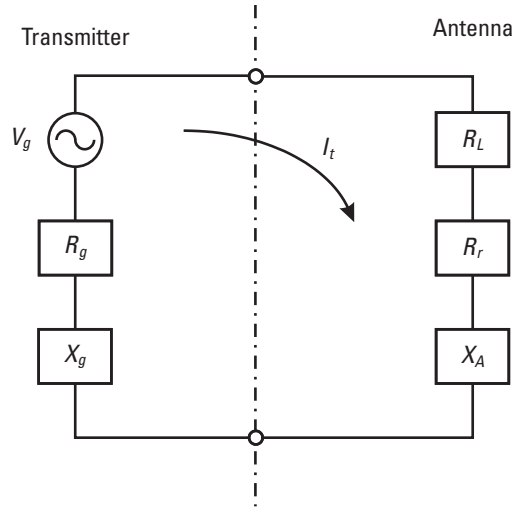


Figure 2.4 Antenna in transmitting mode.

The radiating element is symbolized by a radiation resistance R_r and an imaginary part X_A; R_L represents both the conduction and the dielectric losses of the antenna. The source to which the antenna is connected is represented by an ideal generator V_g having its own internal complex impedance consisting of R_g and jX_g. The energy transferred to the antenna and its environs by the reactive power flow is stored mostly in the reactive near field. The energy transferred to the antenna by the resistive power flow either heats up the antenna structure (or things in the near-field region of the antenna), or else it is radiated; most often, combinations of both of these phenomena occur. Thus, we can break down the resistive part R of the driving point impedance into the sum of a loss resistance, R_L, which gets hot, and a radiation resistance, R_r. Because the purpose of an antenna is to radiate energy, it is therefore the radiation resistance R_r that is most interesting. The radiation power delivered by the antenna is the power collected by resistance R_r, and it is given by:

$$P_r = \frac{1}{2}|I_t|^2 R_r \tag{2.6}$$

where I_t is the current through R_r.

The factor 2 arises because the average value of the square of a unit sinusoidal signal over one cycle is just 1/2. In determining the average power dissipated, we intrinsically assume the average is taken over a whole number of cycles of the ac signal.

The magnitude of I_t can be calculated using the following expression:

$$|I_t| = \frac{|V_g|}{\sqrt{(R_r + R_L + R_g)^2 + (X_A + X_g)^2}} \tag{2.7}$$

Maximum power is delivered to the antenna under conjugate matching, meaning:

$$\begin{aligned} R_L + R_r &= R_g \\ X_A &= -X_g \end{aligned} \tag{2.8}$$

The maximum power at maximum efficiency will be transferred when the impedances are complex conjugate matched throughout the power chain, from the transmitter output, through the transmission line (which may be a balanced pair, a coaxial cable, or a waveguide), to the antenna system, which consists of an impedance matching device and the radiating element(s). For maximum power, $Z_{load} = Z_{source}^*$ (where * indicates the complex conjugate).

If we consider a lossless antenna ($R_L = 0$), the ideal amount of power collected by R_r is calculated by combining (2.6) and (2.7) and given in:

$$P_r = \frac{|V_g|^2}{8R_r} \tag{2.9}$$

Power collected induces a voltage V_r on the receiving antenna, which is analogous to V_g of the transmitting antenna model. The Thevenin equivalent circuit of the receiving antenna and its load is shown in Figure 2.5. The load to which the receiving antenna is connected (receiver and transmission line) is represented by the receiver's input complex impedance, R_{rec} and X_{rec}. As previously shown for the transmitting model, we can derive the functional equations from this receiving antenna model as well.

Typically, the gain in any given direction and the impedance at a given frequency are the same when the antenna is used in transmission or in reception, due to *reciprocity*.

2.1.3 Theory of Electromagnetism and Maxwell's Equations

James Clerk Maxwell (1831–1879) was a Scottish physicist and mathematician whose major discovery of the ether described the vast sea of space that made possible the transmission of light, heat, and radio waves. Maxwell's discovery of the ether (or his metaphor) led to many advances in electronic communications. His extension of the electromagnetic theory of light led directly to Heinrich Hertz's

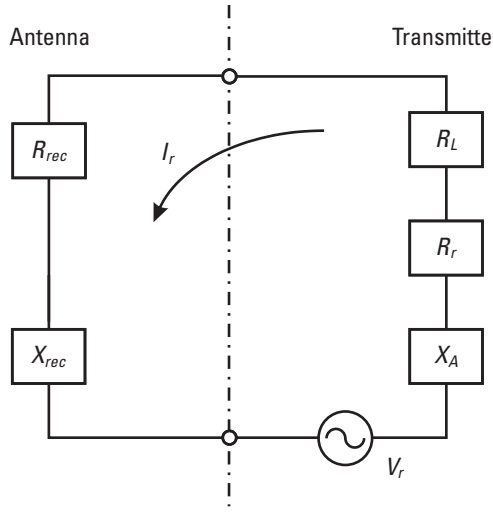


Figure 2.5 Antenna in a receiving mode.

discovery of radio waves and to the related advances in science and technology of today.

Maxwell's mathematical equations, expressing the behavior of electric and magnetic fields and their interrelated nature, were valid, even though his theory of the ether was not. His calculations were scientific observations resulting in his conclusion that the speed of propagation of an electromagnetic field is approximately that of the speed of light (300 million m/s) Maxwell's proposal that the phenomenon of light is therefore an electromagnetic phenomenon seemed to fit what he and other scientists could observe of the world around them. Maxwell concluded that visible light forms only a small part of the entire spectrum of possible electromagnetic radiation (EMR).

The "Complete Laws of Electrodynamics" define the relationship between the electric field quantities and the magnetic field quantities, and although a detailed explanation of these laws is beyond the scope of this book, they deserve to be mentioned here at least briefly. Maxwell was the first to correctly assemble the complete laws of electrodynamics in his classic text in 1873. Modern electromagnetism theory is based on the four fundamental equations known as *Maxwell's equations.* Before Maxwell, the laws of electrodynamics, including Gauss's law, Ampere's law of magnetostatics, and Faraday's law, were laws of electrostatics, and did not predict waves. These laws correctly described what is known as the *near field* (i.e., the electrostatic field of an electric charge and the magnetostatic field of a current loop). These laws described the observable impact of electric charges and magnetic fields close to the source, but failed to describe the distant impact of these forces. In the static case, when all electric

charges are permanently fixed or if they all move at a steady state, the electric field and the magnetic field are not interconnected. This allows us to study electricity and magnetism as two distinct and separate phenomena. Up until Maxwell challenged conventional wisdom, the separation of electricity and magnetism was the accepted state of the world. He corrected Ampere's law of magnetostatics to become Ampere's law as corrected by Maxwell, so that consistency with the law of conservation of charge now occurred. Maxwell added a term indicating that the vortices of magnetic fields can be displacement current density (time-varying electric flux density) as well as conduction current density. The resulting corrected equations define the complete laws of electrodynamics and predict electromagnetic waves. Heinrich Rudolf Hertz confirmed experimentally that these waves exist.

Maxwell showed that any conductor (e.g., antenna) supplied with an alternating current produces a varying magnetic field (H-field) that in turn produces electric field lines (E-field) in space. This is termed the *near field*. In the near field both the E- and H-fields are relatively static, with no propagation. They vary only in strength as the current varies, with the magnetic flux of the H-field coming out from the antenna, and going back in, and the E-field emanating outward. Maxwell also proved that beyond this quasistatic near field, both the E-fields and H-fields at a certain distance detached themselves from the conductor and propagated into free space as a combined wave, moving at the speed of light, with a constant ratio of E/H = 120π or 377Ω. (Ohms are used because the E-field is measured in volts per meter and the H-field in amps per meter.) The point at which this happens is called the *far field*. By applying Maxwell's equation to magnetic dipoles, we can identify that the distance $r = \lambda/(2\pi)$ is of significance in determining the nature of the fields surrounding the dipoles. Within this distance, we have the near-field region; beyond this distance, the far-field region starts. Maxwell's beautiful equations correctly describe both the energy storage field and the energy propagation field.

It was the physicist Ludwig Boltzmann who said that "there is nothing more practical than a good theory," and that is absolutely true in the case of Maxwell's equations. Although highly theoretical and mathematical in their nature, they provided an unprecedented contribution to our understanding of the world around us, as well as a very practical side to the development of radio communications.

2.1.4 Range of a Radio Communications System

For any given radio transmitter and receiver, there is a maximum distance (or range) over which communications can work reliably. If the separation between transmitter and receiver were increased beyond this distance, the receiver would no longer be able to correctly recreate the information being transmitted. It is

prudent to operate a radio communications system with a certain margin, that is, not right at the maximum range, because the exact range is likely to vary from moment to moment, and this must be accommodated to achieve reliable performance. The exact range will be affected by a number of factors, but in simple terms there are four:

1. The power contained in the wave transmitted;
2. The sensitivity of the receiving equipment;
3. The environment through which the wave travels;
4. The presence of interference.

These factors are largely self-explanatory, but of particular note is the relationship between power and distance. As the radio wave travels away from the transmitting antenna, it disperses in all directions. This means that every time the distance from the transmitter doubles, the proportion of the original wave that is available for reception is quartered (the so-called inverse square relationship.) Also note the effect that the environment has on radio communications: When electromagnetic radiation passes through materials, it may be absorbed to a certain extent, depending on the properties of the material and the type of radiation. This absorption results in a reduction of the strength of the radiation, a process known as *attenuation*. This attenuation increases with the thickness of the material. Visible light is absorbed relatively easily, whereas radio waves are more likely to pass through materials (especially gases in the atmosphere, such as nitrogen and oxygen, and also paper, cardboard, and certain plastics) with only little attenuation of the radiation. Other materials (metal and liquids, for example) have a stronger attenuating effect, although such attenuation varies, depending on the frequency of the wave. Once again, the different frequencies within a single band will display properties similar to those of each other.

In addition to attenuation by absorption, certain frequencies of radio waves are also susceptible to *multipath fading*. This occurs when waves are reflected by objects in the environment, and the reflections interfere with the original waveform, making it much more difficult for the radio-wave receiver to determine the original wave. In the worst case there are positions (called nulls) where reception is not possible, even though the transmitter and receiver are relatively close to each other. Similarly, if any other radio waves are being transmitted on a similar frequency, these will cause interference in the same manner.

2.2 Overview of Short-Range Communications Systems

Human society is entering an era of ubiquitous computing, when networks are seamlessly interconnected and information is always accessible when needed. The concept of *ubiquitous computing* (meaning that computers are everywhere) was introduced by Xerox's Palo Alto Research Center (PARC) in 1988 for the first time. This is a computing environment in which all objects and targets in the physical environment become intelligent and exchange information by linking to each other. The practical implementation of ubiquitous services requires three levels of connectivity: wide-area networks (WANs), typically via the Internet, to remotely connect all types of servers and terminals; local-area networks (LANs), typically via Ethernet or WiFi connectivity among all information and communication appliances in offices and homes; and human-area networks (HANs) for connectivity to personal information, media, and communication appliances within the much smaller sphere of ordinary daily activities (i.e., the last 1m). In the future, short-range wireless technology will play a key role in scenarios in which people are connected anywhere and anytime by different types of communication links.

Short-range radio communications devices (SRDs) have been used for many years to provide low-cost services such as short-range telemetry, voice and video communications, radio LANs, and security systems (Figure 2.6). They are defined as radio transmitters which have a low capability of causing interference to other radio equipment. In recent years, there has been rapid growth in the use of SRDs, driven by new technologies and international coordination on specifications. Designers of SRD wireless systems need to use great care when choosing the radio's communication frequency. In most cases, the choice is limited to

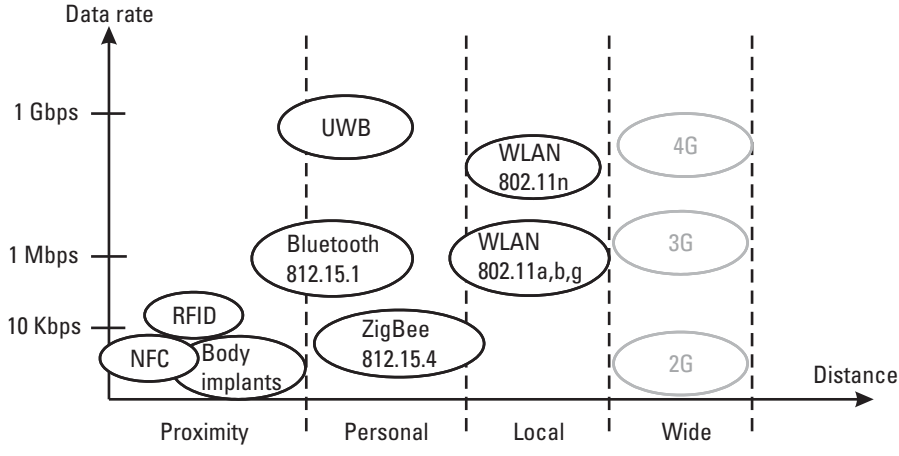


Figure 2.6 Short-range communications systems.

those portions of the spectrum that allow license-exempt operation, given that certain specifications and conditions on usage are met. The international coordination has included attempts to provide common spectrum allocations in the major trading regions to ensure mass markets and minimize the chance of nonstandard equipment appearing on the market.

Products typically operating in the frequency range between 300 MHz and 2.5 GHz are often referred to as ISM-band products in the United States and SRD products in the European Union (EU). (ISM is an acronym for industrial, scientific, and medical frequency bands.) Both the U.S. and EU regulatory agencies place limitations on operating frequencies, output power, spurious emissions, modulation methods, and transmit duty cycles, among other things. The 2.4-GHz band is widely used by designers who want to build systems that can operate worldwide, and in fact, it has become the frequency band of choice for such standards as Bluetooth, wireless LAN (WLAN), and ZigBee. The 5.8-GHz band has also attracted some attention in cordless phones or the 802.11a version of WLAN, for example.

In the United States, the regulations governing unlicensed wireless products fall into two broad categories: periodic devices and ISM-band devices. The type of application and the communications range determine the appropriate band and Federal Communications Commission (FCC) classification to use. In the European Union, the regulations governing low-power wireless devices are essentially defined by two separate bodies. One group defines the allocation of frequency bands and their use and another group defines the test methodologies and general transceiver specifications.

Spread-spectrum systems were originally developed by the military to counter attempts to detect, decode, or block signal transmissions. The most important peacetime characteristics of spread-spectrum systems are that they facilitate radio communications in a manner that minimizes the potential to cause harmful interference to other services, and they are able to withstand higher levels of interference than other technologies. Hence, spread-spectrum systems have a significant potential to share a common spectrum with other services. Spread-spectrum devices operate on a license fee–exempt basis if certain technical conditions are met and type approval is mandatory. Two main types of spread-spectrum systems are commercially available: *direct sequence* and *frequency hopping*. Prior to the latest EN 300 220 specification (1996), however, the U.S. and European bodies took vastly different regulatory approaches. The United States adopted a frequency-hopping approach, whereas Europe applied duty-cycle limits in each of the sub-bands, as described in the ERC REC-70 document. Both of these implementations are useful in minimizing interference, but manufacturers who were designing systems for both regions needed to completely rewrite the media access layer (MAC) in the system's communication protocol. Fortunately, the latest European EN 300 220 regulations have

extended the frequency bands to allow for frequency-hopping spread-spectrum (FHSS) or direct-sequence spread-spectrum (DSSS) technologies, thus making the MAC implementations more similar to those designed for the U.S. market.

Aside from the two most popular spread-spectrum systems described below, an interesting aspect of the new European regulations is that they provide for other wideband spread-spectrum modulation schemes in addition to FHSS and DSSS. FSK/GFSK (Gaussian frequency-shift-keying) modulation, with an occupied bandwidth greater than 200 kHz, is considered wideband modulation under the European regulations.

2.2.1 Frequency-Hopping Spread-Spectrum Systems

The FHSS transmission technology spreads energy in the time domain by dividing the spectrum into a number of channels, switching between them in a pseudorandom sequence, or hopping code, that is known by both the receiver and transmitter. U.S. and European standards both specify a similar number of hopping channels, and a maximum dwell time (the time spent at a particular frequency during any single hop) of 400 ms. Bandwidths of up to 7 MHz are available once either the listen-before-talk (LBT) or duty-cycle limits are met, as compared to the 2-MHz range available previously. Listen-before-talk, a polite communication protocol, scans the channel for activity before initiating a transmission. Also called *clear-channel-assessment* (CCA), systems using LBT with frequency hopping have no duty-cycle limitations.

2.2.2 Direct-Sequence Spread-Spectrum Systems

Besides FHSS, DSSS is also addressed in the new European regulations. In a DSSS system, a narrowband signal is multiplied by a high-speed pseudorandom number (PRN) sequence to generate a spread signal. Each PRN pulse is called a *chip,* and the rate of the sequence is called the *chip rate.* The extent to which the original narrowband signal is spread is referred to as the processing gain; it is the ratio of the chip rate to the narrowband data symbol rate. At the receiver, the incoming spread-spectrum signal is multiplied with the same PRN code to despread the signal, allowing the original narrowband signal to be extracted. At the same time, any narrowband interferers at the receiver are spread and appear to the demodulator as wideband noise. The allocation of different PRN codes to each user in the system allows isolation between users in the same frequency band. This is known as *code-division multiple access* (CDMA).

A few examples of systems using DSSS modulation include IEEE 802.15.4 (WPAN), IEEE 802.11 (WLAN), and the global positioning system (GPS). The main advantages of DSSS are:

- Interference resilience; the essence of the interference-rejection capability of DSSS is that the useful signal gets multiplied twice (spread and despread) by the PRN code, while any interferers are multiplied just once (spread);
- Low-power spectral density; introducing minimal interference with existing narrowband systems;
- Security; very resistant to jamming because of spreading/despreading;
- Mitigation of multipath effects.

2.3 Wireless LANs

2.3.1 Basics of WLANs

Wireless LAN access technology provides a perfect broadband complement for the operators' existing or new third generation (3G) services in an indoor environment. Three standards dominate the WLAN marketplace; 802.11b has been the industry standard for several years. Operating in the unlicensed portion of the 2.4-GHz RF spectrum, it delivers a maximum data rate of 11 Mbps and boasts numerous strengths. The 802.11b standard enjoys broad user acceptance and vendor support because many vendors manufacture compatible devices, and this compatibility is ensured through the WiFi certification program. Thousands of enterprise organizations that typically find its speed and performance acceptable for their current applications have deployed 802.11b technology. In the United States, a number of wireless Internet service providers (ISPs) have emerged that are offering public access services using IEEE 802.11b equipment operating in the 2.4-GHz band. These providers are targeting public areas, such as hotels or coffee shops, where business travelers may wish to access corporate intranets or the Internet. In some parts of Europe, a number of service providers, both mobile operators and ISPs, now offer wireless Internet services based on 802.11b technology in the 2.4-GHz band.

Another WLAN standard, IEEE 802.11a, operates in the uncluttered 5-GHz RF spectrum. With a maximum data rate of 54 Mbps, this standard offers a fivefold performance increase over the 802.11b standard. Therefore, it provides greater bandwidth for particularly demanding applications. The IEEE ratified the 802.11a standard in 1999, but the first 802.11a-compliant products did not begin appearing on the market until December 2001. The 802.11a standard delivers a maximum data rate of 54 Mbps and eight nonoverlapping frequency channels, resulting in increased network capacity, improved scalability, and the ability to create microcellular deployments without interference from adjacent cells.

Operating in the unlicensed portion of the 5-GHz radio band, 802.11a is also immune to interference from devices that operate in the 2.4-GHz band, such as microwave ovens, cordless phones, and Bluetooth (a short-range, low-speed, point-to-point, personal-area network wireless standard). The 802.11a standard is not compatible with existing 802.11b-compliant wireless devices. The 2.4- and 5-GHz equipment can operate in the same physical environment without interference.

The IEEE 802.11g standard is the latest widely adopted high-performance standard, and it delivers the same 54-Mbps maximum data rate as 802.11a and operates in the same 2.4-GHz band as 802.11b. Because this spectrum is unlicensed, even more uses for it are expected to develop in the future. As the band becomes more widely used, radio interference will increase. Bluetooth uses FHSS, is a shorter range and lower bandwidth technology than 802.11b, and uses frequently changing narrowbands over all channels. Not only do microwave ovens operate within this range, but other RF communications technologies as well, most notable of which is IEEE 802.11b. A new 802.11 Task Group (TGn) is developing a new amendment to the 802.11 standard for WLANs with the real-data throughput estimated to reach a theoretical 540 Mbps (most likely about 120 Mbps real world). The 802.11n standard will build on previous 802.11 standards by adding multiple-input multiple-output (MIMO) technique. MIMO uses multiple transmitting and receiving antennas to allow for increased data throughput through spatial multiplexing and increased range by exploiting the spatial diversity and new coding schemes.

The Enhanced Wireless Consortium (EWC) was formed to help accelerate the IEEE 802.11n development process and promote a technology specification for interoperability of next generation WLAN products. On January 19, 2006, the IEEE 802.11n Task Group approved the joint proposal's specification, based on EWC's specification as the confirmed 802.11n proposal. According to the IEEE 802.11 Working Group project timelines, the 802.11n standard was not due for final approval, called publication date, until the end of 2008.

2.3.2 WLAN Components

A wireless LAN is made up of two key components; an access point (AP) or base station that is usually, but not necessarily, physically connected to a LAN and a wireless card that is either built into or added to a computer device. This can be a handheld personal digital assistant (PDA), tablet, laptop, or desktop computer (Figure 2.7). With a wireless LAN in place, portable computers can remain connected to the network while on the move. Any device with a wireless adaptor within range of an access point can potentially connect to the WLAN. This provides greatly increased freedom and flexibility compared to a wired network.

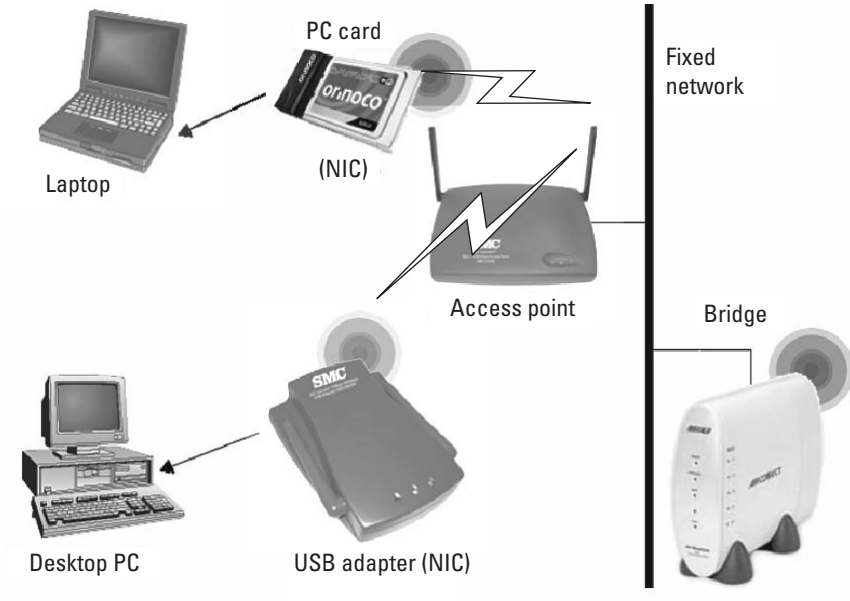


Figure 2.7 WLAN components.

Extending the WLAN to include additional users often only requires that the user have a wireless-enabled computer device and be in range of an access point. Increasing the overall network coverage of the WLAN can often be achieved by adding further access points.

2.4 Wireless Personal Area Network

A wireless personal area network (WPAN) is defined as a network of a very limited radius, up to 10 feet (3.3m), occupying a very small amount of space. Specification 802.15 is a wireless specification defined by IEEE for WPANs. It has characteristics such as short range, low power, low cost, small networks, and communication of devices within a personal operating space. As radios decrease in cost and power consumption, it becomes feasible to embed them in more types of electronic devices, which can be used to create smart homes, sensor networks, and other compelling new applications. Two radio technologies have emerged to support this trend, Bluetooth (IEEE 802.15.1) and ZigBee (IEEE 802.15.4). WPANs come in another variation that targets high data rates for image and multimedia applications, known as a high-rate wireless personal area network, described by the IEEE 802.15.3 standard.

2.4.1 Bluetooth

The Bluetooth standard is named after Harald Bluetooth, the king of Denmark between 940 and 985 A.D. who united Denmark and Norway. Bluetooth technology proposes to unite devices via radio connections, hence the inspiration for its name.

Bluetooth radios provide short-range connections between wireless devices along with rudimentary networking capabilities. The Bluetooth standard is based on a tiny microchip incorporating a radio transceiver that is built into digital devices. The transceiver takes the place of a connecting cable for devices such as cell phones, laptop and palmtop computers, portable printers and projectors, and network access points. Bluetooth is mainly used for short-range communications, for example, from a laptop to a nearby printer or from a cell phone to a wireless headset. Its normal range of operation is 10m (at 1 mW of transmitting power). This range can be increased to 100m by increasing the transmitting power to 100 mW.

The system operates in the unlicensed 2.4-GHz frequency band; hence, it can be used worldwide without any licensing issues. For carrier frequencies, 79 RF channels are available in this frequency band and the following relation is valid:

$$f[\text{MHz}] = 2{,}402 + k, \text{ where } k = 0,1,2,\ldots,78 \tag{2.10}$$

The RF channel bandwidth is 1.0 MHz. To comply with out-of-band regulations, the guard bands are defined. The lower guard-band bandwidth is 2.0 MHz, and the upper guard-band bandwidth is 3.5 MHz.

The Bluetooth standard provides one asynchronous data channel at 723.2 Kbps. This mode, also known as asynchronous connection-less, or ACL, has a reverse channel with a data rate of 57.6 Kbps. The specification also allows up to three synchronous channels each at a rate of 64 Kbps. This mode, also known as synchronous connection oriented, or SCO, is mainly used for voice applications such as headsets, but can also be used for data. These different modes result in an aggregate bit rate of approximately 1 Mbps.

Routing of the asynchronous data is done via a packet-switching protocol based on frequency hopping at 1,600 hops per second. There is also a circuit-switching protocol for the synchronous data. Bluetooth uses frequency hopping for multiple access with a carrier spacing of 1.0 MHz, and up to 80 different frequencies are typically used, for a total bandwidth of 80 MHz. At any given time, the bandwidth available is 1 MHz, with a maximum of eight devices sharing the bandwidth. Different logical channels (different hopping sequences) can simultaneously share the same 80-MHz bandwidth. Collisions will occur when devices in different *piconets*, on different logical channels, happen to use

the same hop frequency at the same time. As the number of piconets in an area increases, the number of collisions increases, and performance degrades.

The Bluetooth standard was developed jointly by 3Com, Ericsson, Intel, IBM, Lucent, Microsoft, Motorola, Nokia, and Toshiba. The standard has now been adopted by a majority of manufacturers, and many consumer electronic products incorporate Bluetooth, including wireless headsets for cell phones, wireless USB or RS232 connectors, wireless PCMCIA cards, and wireless set-top boxes.

Bluetooth is considered to be a point-to-multipoint system, although it can also be a point-to-point system, depending on the application. Bluetooth technology allows for the replacement of the many proprietary cables that connect one device to another with one universal short-range radio link. For instance, Bluetooth radio technology built into both the cellular telephone and the laptop would replace the cumbersome cable used today to connect a laptop to a cellular telephone. Printers, PDAs, desktops, fax machines, keyboards, joysticks, and virtually any other digital device can be part of the Bluetooth system.

But beyond replacing the cables, Bluetooth radio technology provides a universal bridge to existing data networks, a peripheral interface, and a mechanism to form small, private, ad hoc groupings of connected devices a distance away from fixed network infrastructures. Designed to operate in a noisy RF environment, the Bluetooth radio uses a fast acknowledgment and frequency-hopping scheme to make the link robust. Bluetooth radio modules avoid interference from other signals by hopping to a new frequency after transmitting or receiving a packet.

The spread of Bluetooth, although initially slower than early industry predictions, has already had an impact on the availability of embedded car telephones. In addition, there will be opportunities for hybrid systems, which combine an embedded telephone with a Bluetooth connection to the SIM card in the driver's mobile. Although the integration of mobile telephones has the highest priority, vehicle manufacturers are also exploring the potential of Bluetooth to be used for the connection of digital music players and portable navigation systems.

Compared with other systems operating in the same frequency band, Bluetooth radio typically hops faster and uses shorter packets, thus making it more robust than other systems. Short packages and fast hopping also limit the impact of domestic and professional microwave ovens. Use of forward error correction (FEC) limits the impact of random noise on long-distance links. The encoding is optimized for an uncoordinated environment, since Bluetooth radios operate in the unlicensed ISM band at 2.4 GHz. A frequency-hop transceiver is applied to combat interference and fading. Shaped, binary FM modulation is applied to minimize transceiver complexity and a time-division duplex scheme is used for full-duplex transmission, with a gross data rate of about 1

Mbps. The Bluetooth baseband protocol is a combination of circuit and packet switching. Slots can be reserved for synchronous packets. Each packet is transmitted in a different hop frequency. A packet nominally covers a single slot, but it can be extended to cover up to five slots. Bluetooth can support an asynchronous data channel, up to three simultaneous synchronous voice channels, or a channel that simultaneously supports asynchronous data and synchronous voice. Each voice channel supports a 64-Kbps synchronous (voice) link. The asynchronous channel can support an asymmetric link of maximally 721 Kbps in either direction while permitting 57.6 Kbps in the return direction, or a 432.6-Kbps symmetric link.

The Bluetooth system supports both point-to-point and point-to-multipoint connections; a piconet is an ad hoc computer network of devices using Bluetooth technology protocols to allow one master device to interconnect with up to seven active slave devices. (A 3-bit address space limits the number of devices in any piconet to eight.) Up to 255 further slave devices can be inactive, or parked; the master device can make these slaves active at any time. Several piconets can be established and linked together in an ad hoc manner, in which each piconet is identified by a different frequency-hopping sequence, and all users participating on the same piconet are synchronized to this hopping sequence. The topology can best be described as a *multiple piconet structure*.

Voice channels use the continuous variable slope delta modulation (CVSD) voice coding scheme, and never retransmit voice packets. The CVSD method was chosen for its robustness in handling dropped and damaged voice samples. Rising interference levels are experienced as increased background noise: Even at bit-error rates up 4%, the CVSD-coded voice is quite audible. The Bluetooth air interface is based on a nominal antenna power of 0 dBm. The air interface complies with the FCC rules for the ISM band at power levels up to 0 dBm. Spectrum spreading has been added to facilitate optional operation at power levels up to 100 mW (20 dBm) worldwide. Spectrum spreading is accomplished by frequency hopping in 79 hops displaced by 1 MHz, between 2.402 and 2.480 GHz. Due to local regulations the bandwidth is reduced in Japan, France, and Spain. This is handled by an internal software switch.

The maximum frequency-hopping rate is 1,600 hops per second. The nominal link range is 10 cm to 10m, but can be extended to more than 100m by increasing the transmitting power. The link type defines what type of packets can be used on a particular link. The Bluetooth baseband technology supports two link types: the SCO type, which is used primarily for voice, and the ACL type, which is used primarily for packet data.

Different master/slave pairs on the same piconet can use different link types, and the link type may change arbitrarily during a session. Each link type supports up to 16 different packet types. Four of these are control packets and are common to both SCO and ACL links. Both link types use a time-division

duplex (TDD) scheme for full-duplex transmissions. Three error-correction schemes are defined for Bluetooth baseband controllers:

- One-third rate FEC;
- Two-thirds rate FEC;
- Automatic repeat request (ARQ) scheme for data.

The purpose of the FEC scheme on the data payload is to reduce the number of retransmissions. However, in a reasonably error-free environment, FEC creates unnecessary overhead that reduces the throughput. Therefore, the packet definitions have been kept flexible as to whether or not to use FEC in the payload. The packet header is always protected by a one-third rate FEC; it contains valuable link information and should survive bit errors. An unnumbered ARQ scheme is applied in which data transmitted in one slot are directly acknowledged by the recipient in the next slot. For a data transmission to be acknowledged, both the header error check and the cyclic redundancy check must be okay; otherwise, a "negative acknowledge" message is returned.

2.4.2 ZigBee

The ZigBee radio specification is designed for lower cost and power consumption than Bluetooth. ZigBee takes its name from the dance that honeybees use to communicate information about newfound food sources to other members of the colony. Yet another protocol was needed because other short-range protocols, such as 802.11 and 802.15 (Bluetooth), use too much power and the protocols are too complex—and thus too expensive—to be embedded in virtually every kind of device imaginable. Potential applications are sensors, interactive toys, and remote controls. In addition, ZigBee technology makes it possible to control upcoming home networks, such as those used for controlling electrical appliances, checking temperature and humidity, and sending mobile messages to alarms in cases of trespass.

The goal of ZigBee is to provide radio operation for months or years without recharging, thereby targeting applications such as sensor networks and inventory tags. The IEEE 802.15 WPAN Task Group 4 (TG4) was chartered to investigate a low-data-rate solution with multimonth to multiyear battery life and very low complexity; the new standard has been published as IEEE 802.15.4-2006. The specification operates in the 2.4-GHz (ISM) radio band, the same band used for the 802.11b standard, Bluetooth, microwaves, and some other devices. It is capable of connecting 255 devices per network [2]. The data rate of ZigBee is 250 Kbps at 2.4 GHz, 40 Kbps at 915 MHz, and 20 Kbps at 868 MHz, whereas that of Bluetooth is 1 Mbps. ZigBee's data rates are slower

than 802.11b (11 Mbps) and Bluetooth (1 Mbps), but it consumes significantly less power. The transmitting power levels for 802.15.4 radios are very low, typically −3 dBm (0.5 mW). The receiving sensitivity is −80 to −100 dBm, depending on the 802.15.4 radio.

ZigBee allows small, low-cost devices to quickly transmit small amounts of data, such as temperature readings for thermostats, on/off requests for light switches, or keystrokes for a wireless keyboard. ZigBee devices, typically battery powered, can actually transmit information much farther than 20m (60 feet) because each device within listening distance passes the message along to any other device within range, and only the intended device acts on the message. ZigBee networks are primarily intended for low-duty-cycle sensor networks (<1%). A new network node may be recognized and associated in about 30 ms. Waking up a sleeping node takes about 15 ms, as does accessing a channel and transmitting data. ZigBee applications benefit from the ability to quickly attach information, detach, and go to deep sleep, which results in low power consumption and extended battery life.

The ZigBee Alliance, which totals nearly 100 companies, including Honeywell, Mitsubishi, Motorola, Philips, and Samsung, is hoping to penetrate the home-automation market, which has thus far been confused by a variety of proprietary technologies.

2.5 Wireless Body Area Networks

2.5.1 About WBANs

Thus far we have talked about WLANs and WPANs. Now we focus on the wireless networks with a range of less than 1 foot, the so-called wireless body area networks (WBANs). In the future, devices such as implantable biomedical systems will become really miniature computers that employ sensitive, low-voltage, low-power, application-specific integrated circuits (ASICs) to measure, monitor, and regulate physiological parameters, and control the delivery of electrical impulses to different organs in the human body (Figure 2.8). The implanted medical devices and on-body sensors are mainly connected with monitoring tools to provide patient health data in real time using body area networks (BANs), also called body sensor networks (BSNs).

The body network is composed of tiny portable devices equipped with a variety of sensors (such as heart rate, heart rhythm, temperature, pulse oximeter, and accelerometer sensors) and performs biophysical monitoring, patient identification, location detection, and other desired tasks [3, 4]. The energy consumption of these miniature sensors is also optimized so that the battery does not need to be changed regularly; they may use kinetic recharging. Actuators notify the wearer of important messages from an external entity. For example, an

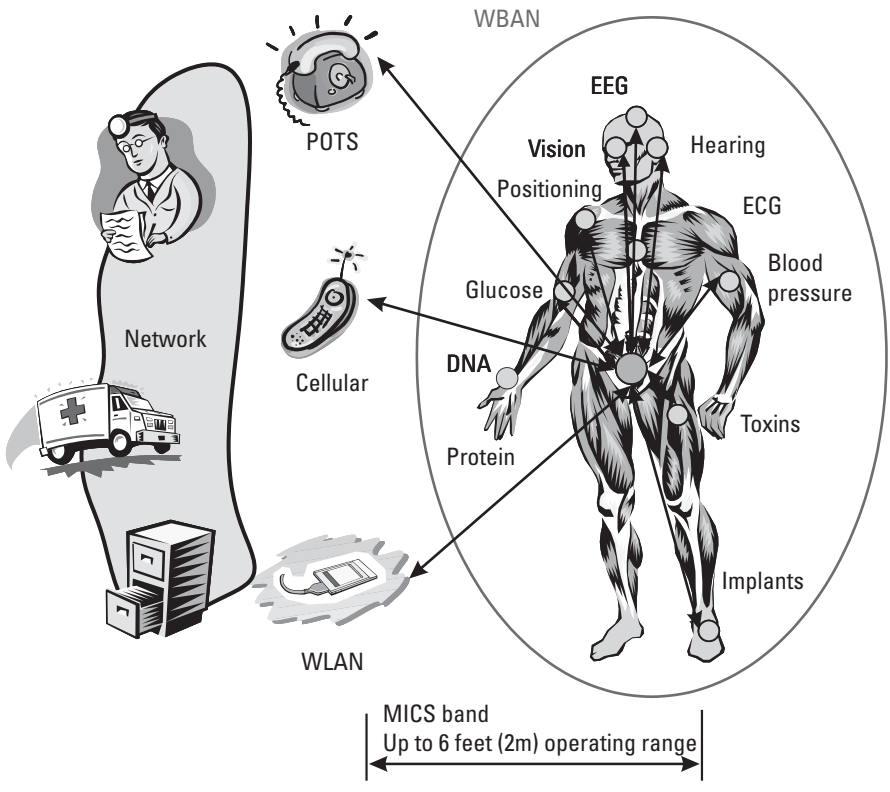


Figure 2.8 Wireless body implants.

actuator can remind an early-stage Alzheimer's patient to check the oven because sensors detect an abnormally high temperature, or a tone may indicate that it is time to take medication. A node in the body network is designated as the gateway to the emplaced sensor network. Due to size and energy constraints, nodes in this network have small processing and storage capabilities.

Up until recently, communication was achieved over an inductive link. For example, a small coil was placed inside the case of the pacemaker, and a larger coil was placed on the chest of the patient, directly on top of the pacemaker. The inductive coupling between these two coils was then used to transfer data to and from the pacemaker. The link was usually at half-duplex, meaning that transmission is in only one direction at any given time. The speed was typically low, a few hundred bits per second, and although higher speeds are achievable, the low carrier frequency limits the available data bandwidth severely.

With the increased sophistication of medical implants, there is a growing need for flexible high-speed communication with the implant from outside the body. Today, the communication is done by an inductive link between the implant and an external coil at a low carrier frequency. Extended range and

communication speed can be achieved by increasing the carrier frequency and the bandwidth; even the 2.45-GHz ISM band is a possibility, but has the drawback of being heavily used by other applications, such as wireless computer networks and microwave ovens. A number of advantages accrue if communication with the implant can be moved to a higher carrier frequency. The first one is an increase in bandwidth, which makes it possible to achieve a higher bit rate. The second one is that a higher frequency gives rise to a propagating electromagnetic wave, which makes the system usable at longer ranges. A longer communication range makes a number of new user scenarios possible.

Today, a number of other implants are in use and under development. Examples are brain pacemakers for the treatment of Parkinson's disease, implantable drug pumps, cochlea implants, artificial eyes, muscle stimulators, and nerve-signal recorders for use with robotic prostheses. All of these implants need some kind of data transfer, either in one or two directions. Neither inductive nor RF is the best for all of them because the power requirements, range, and speed are different from application to application.

Body-worn sensor networks (BSN) could have a major impact on how health care is conducted in the future. By continuously monitoring the life signs of patients and analyzing the signal patterns, dangerous medical conditions could be detected earlier, which would lead to more effective treatment and shorter hospital stays. In addition, long-term life-sign data could improve the quality of diagnoses when a person becomes ill. Many other applications of on-body sensors have been explored by research groups. More and up-to-date information on the new products and applications can be found at http://medicalconnectivity.com.

2.5.2 Inductive Coupling Theory

Inductive links are widely used for implanted biomedical applications, and with amplitude modulation their use can be expanded, for example, to the transmission of pulse trains for deep brain stimulation (DBS). In general, inductive coupling is the transfer of energy from one circuit to another through a shared magnetic field. An electrical current passing through the coil of a primary conductor creates a magnetic field that induces an electrical current in the coil of a secondary conductor exposed to the magnetic field. A complete inductive powering system consists of two major parts: a drive coil and a pickup coil, and their associated impedance-matching networks. The drive coil (reader) is designed to maximize the magnetic field within the desired enclosure at the drive frequency. Similarly, the pickup coil (tag) is designed to maximize the amount of magnetic flux density converted to power for the implant, while minimizing its own dimensions. The following is a description of the reader/tag operation that can be used in a telemetry system for implanted devices in humans or animals.

Previous studies have shown that RF energy between 1 and 10 MHz penetrates the body with minimum energy loss.

Without getting into the details of the principle of the operation, we can say that a transformer is a couple of coils of wire that transfer power from one to the other by a changing magnetic field. By having different numbers of windings, or turns of wire, a transformer can step up or step down an ac voltage. According to (2.11), the proportion of energy captured by the secondary coil can be represented by the coupling coefficient, k:

$$k = \frac{M}{\sqrt{L_1 L_2}} \tag{2.11}$$

where k is the coefficient of coupling and $0 \le k \le 1$; L_1 is the inductance of the first coil; L_2 is the inductance of the second coil; and M is the mutual inductance, which depends only on the geometry of the two coils and is independent of the current in the coil. The coefficient of coupling is always between 1 and 0, and is an important factor in the operation of any inductively coupled system; $k = 1$ means that all of the magnetic flux produced by one coil passes through the other coil, which is practically never the case. In other words, this coefficient describes how closely linked the two inductors are magnetically; the better these two inductors are magnetically coupled, the more efficient the energy transfer between them should be.

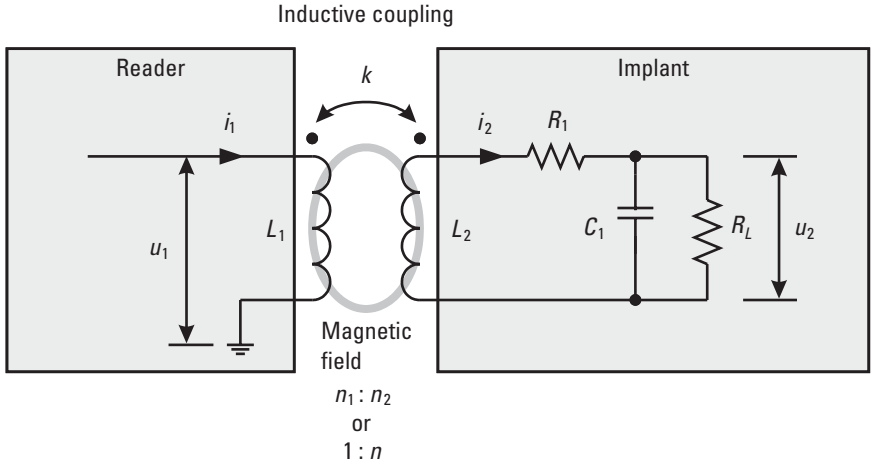
The basic principles behind transferring power and data through an inductive link are the same as those used in transformer circuits; the major difference here is that in this case the two coils are fairly weakly coupled, that is, through the air. Typical values for k in inductively powered system are between 0.01 and 0.1. The coupling coefficient between the two coils, where the radius of the reader coil is much larger than the radius of the transponder (implanted) coil, can be determined empirically using for the air coupling case:

$$k = \frac{a_{\text{implant}}^2 \cdot a_{\text{reader}}^2}{\sqrt{a_{\text{implant}} \cdot a_{\text{reader}}}\left(\sqrt{r^2 + a_{\text{reader}}^2}\right)^3} \cdot \cos\theta \tag{2.12}$$

Equation (2.12) is based on the radii of the two coils (a_{implant} and a_{reader}) and the distance r between them. (See the sample calculation in the Table 2.1.) We assume that the two coils are parallel ($\theta = 0° \rightarrow \cos\theta = 1$) and center aligned, with only air between the two coils. While not as accurate as finite element modeling, this still provides a good approximation of the system coupling coefficient [5]. This value can then be used in a simplified model of the complete inductively coupled system shown in Figure 2.9.

Table 2.1
Coupling Coefficient Values

Implant Radius [in]	Reader Radius [in]	Distance [in]	Coupling Coefficient, *k*
0.3	5.0	1.0	0.0139
0.3	7.0	1.0	0.0086
0.4	5.0	1.0	0.0213
0.4	7.0	1.0	0.0133
1.0	5.0	2.0	0.0716
1.0	7.0	2.0	0.0480



Note: The dot on each of the two coils indicates both voltages have same polarity.

Figure 2.9 Simplified model of the inductively coupled system.

The two windings of the transformer were labeled rather generically 1 and 2. It is also quite common to refer to the two windings as primary and secondary. This convention is often used when a generator is connected to a primary winding, and a load is connected to a secondary winding. In that case, the energy flow is into the primary and out of the secondary; however, all transformers are bidirectional, so there is nothing inherently primary about either of the two windings. The transformer turns ratio was indicated on the symbol as "n_1:n_2" or "1:n." In recognition of the fact that the ideal transformer is a model

for a real transformer, the numbers n_1 and n_2 may be the actual physical turns count of a transformer, such as 363:33. For circuit analysis purposes it is equivalent to give the turns ratio as 11:1, or as n = 1/11 = 0.0909. It might be noted that the turns ratio of a physical transformer is always a rational number, that is, the ratio of two integers. Therefore, a turns ratio of $\sqrt{3}$:1 is not possible, although 173 turns and 100 turns on a physical transformer would do a good job of approximating it.

The left side of this model represents the outside components of the system reader, whereas the right side includes a basic model of the implanted system. Here, R_1 represents the parasitic resistance in the coil, C_1 is the tuning capacitance used to raise the coil voltage, and R_L is the load on the system. The weakly coupled transformer is used here to represent the two discrete coils, L_1 and L_2. The primary (reader's) coil, L_1, is driven by an RF amplifier supplying current at frequency ω. In the real system, R_L is time varying and complex, whereas in this model it is represented as a real resistor.

In circuit analysis it is quite common to use the concept of an *ideal transformer*, which does not generate, dissipate, or store energy (Figure 2.10). Therefore, the instantaneous power leaving the transformer is the same as that entering. This could be said in other words by saying that if one were to draw a box around an ideal transformer and sum the power flows into (or out of) the box, the answer is zero at every moment in time.

According to the basic equations governing ideal transformer behavior (i.e., the magnetizing current is negligibly small), the current out of the transformer, i_2, is shown here:

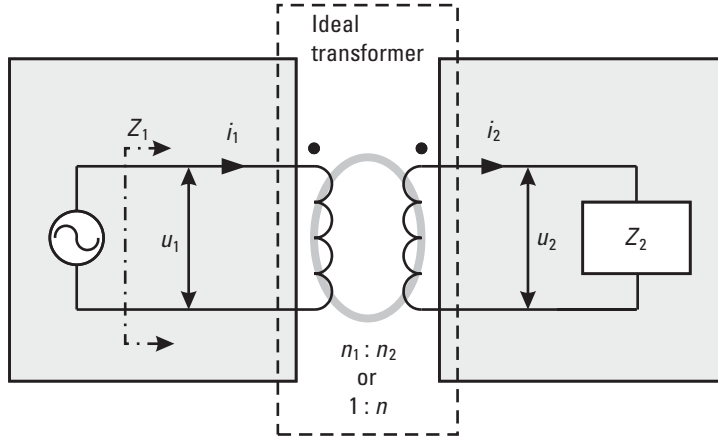


Figure 2.10 Ideal transformer.

$$\begin{aligned} i_1 n_1 &= i_2 n_2 \\ i_2 &= \frac{n_1}{n_2} i_1 \\ i_2 &= \frac{i_1}{n} \end{aligned} \tag{2.13}$$

The ideal transformer has the voltage relationship shown in:

$$\frac{u_1}{u_2} = \frac{n_1}{n_2} \tag{2.14}$$

The voltage across the secondary of the transformer (the output windings) can be calculated as follows:

$$u_2 = \frac{n_2}{n_1} u_1 = n \cdot u_1 \tag{2.15}$$

Equation (2.16) illustrates the *impedance-scaling* property of the ideal transformer:

$$Z_1 = \frac{u_1}{i_1} = \frac{u_2}{n^2 i_2} = \frac{Z_2}{n^2} \tag{2.16}$$

The impedance-scaling property may be interpreted by saying that if a given impedance is connected to one winding of an ideal transformer, it will appear the same at the other winding, scaled in magnitude by the turns-ratio squared (n^2). The impedance appears greater at the winding having the greater number of turns and smaller at the winding having fewer turns. The impedance-scaling property of an ideal transformer allows circuit elements to be moved from one winding to another by scaling their impedances according to the square of the transformer turns ratio.

Now, the weakly coupled transformer can be replaced by an approximation (Figure 2.11). In this model, the new ratio n' (actually, a transfer function) is approximated by:

$$n' \approx k\sqrt{\frac{L_2}{L_1}} \tag{2.17}$$

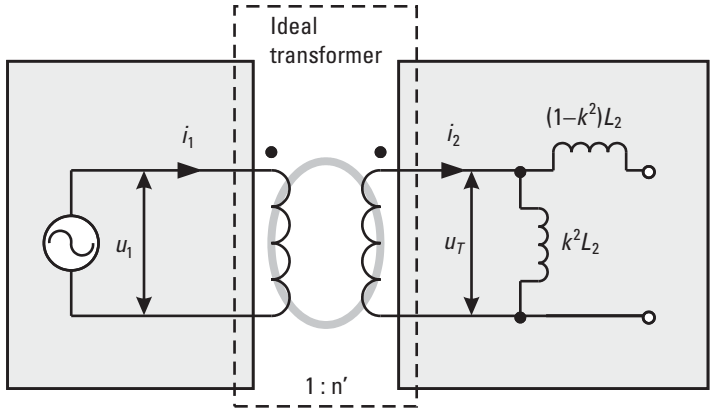


Figure 2.11 Model of a weakly coupled transformer.

Not only the data signal, but also other external spurious magnetic fields are transformed by the primary coil. The magnitude of these transformed voltages depends on the coil geometry. Generally speaking, increasing primary inductance L_1 will cause an increasing influence of the external fields. With a low coupling coefficient k, the impedance seen by current i_2 is approximately equal to that of an inductor with value k^2L_2 (the impedance of this inductor is much lower than that of the other inductor in the circuit; therefore, nearly all of the current will flow through this inductor). Using the impedance equation of the inductor at a known frequency, the voltage induced by this current is:

$$u_T = jX_L \cdot i_2 = j\omega L_2 k^2 \frac{1}{n'} i_1 = j\omega k\sqrt{L_1 L_2}\, i_1 \tag{2.18}$$

Here u_T is the voltage induced by the current i_1 across the L_2 component of the transformer. We can now replace the weakly coupled transformer with a voltage source, u_T, in series with an inductor. With a small k value, we can approximate the value of this inductor by L_2 (Figure 2.12).

In this simplified case, a basic equation for the voltage across the load, u_2, is given by:

$$u_2 = \frac{u_T}{j\left(\frac{\omega L_2}{R_L} + \omega R_1 C_1\right) + \left(1 - \omega^2 L_2 C_1 + \frac{R_1}{R_L}\right)} \tag{2.19}$$

This equation makes use of the impedances of the various components at a known frequency of operation (the UHF band). Substituting in the equation for

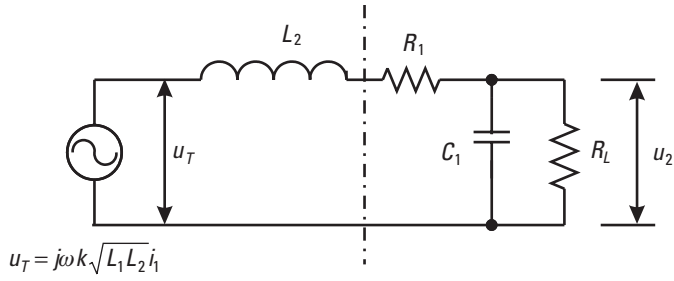


Figure 2.12 Equivalent model of the inductively coupled system.

the transformer voltage and solving for the real part of the solution leads to a final answer with respect to the known parameters of the system:

$$u_2 = \frac{\omega k\sqrt{L_1 L_2}\, i_1}{\sqrt{\left(\frac{\omega L_2}{R_L} + \omega R_1 C_1\right)^2 + \left(1 - \omega^2 L_2 C_1 + \frac{R_1}{R_L}\right)^2}} \tag{2.20}$$

This creates a linear scale factor B, or the gain of the system, with units of ohms. The new equation for this particular system can be written as follows:

$$u_2 = B \cdot k \cdot i_1 \tag{2.21}$$

where we require u_2 to be within a certain range; it is a factor of the distance between the two coils. Therefore, by changing the current through the primary coil of the system, the voltage on the implanted coil could be adjusted for a fixed coupling factor. This gives a required coil (rms) current on the order of 100 mA for the system with weak coupling.

Inductive coupling has many different applications: implanted electronic circuits, RFID, or even as a low-cost, low-power, high-bandwidth interfaces for interconnection of the modules on a chip [6].

2.5.3 Medical Implant Communication Service and Wireless Medical Telemetry Service Bands

Implantable medical devices (IMDs) have a history of outstanding success in the treatment of many diseases, including heart diseases, neurological disorders, and deafness. Today's aging population is a driving force for more advanced health care treatments, including wireless implant devices that can deliver ongoing and cost-effective monitoring of a patient's condition. New ultralow-power RF technologies are spurring the development of innovative medical tools, from

endoscopic camera capsules that are swallowed to implanted devices that wirelessly transmit patient health data. Communication links between external programming devices (or base stations) and medical implants are critical to the success of IMDs. The communication link enables a clinician to reprogram therapy and obtain useful diagnostic information. Historically, low-frequency inductive links (introduced in the early 1970s) have been the most prevalent method of communication. They typically operate in the tens-to-hundreds-of kilohertz range, with data rates of 1 to 30 Kbps. These low-power systems can accommodate a small coiled antenna in the IMD and have proven to be robust and suitably reliable.

However, antenna size and power limitations in implants result in a very low magnetic field strength for an IMD that is communicating with an external programmer. Therefore, inductive links are short range and often require the external programmer to have contact with the skin of the patient directly over the implant. To overcome these operating-range and low-data-rate limitations, new ultralow-power RF technologies are being developed that operate at much higher frequencies, such as in the 433- and 915-MHz ISM bands and the more recently allocated 402- to 405-MHz medical implant communication service (MICS) band. RF integrated circuit (IC) technology can now offer a low-power, reduced external component count and higher levels of integration, which will open new markets for medical device manufacturers (see http://www.zarlink. com).

From a regulatory viewpoint, the establishment of the MICS band began in the mid-1990s when Medtronic petitioned the FCC to allocate a spectrum dedicated to medical-implant use. After gaining wider industry support, the 402- to 405-MHz MICS band was recommended for allocation by ITU-R Recommendation SA1346 in 1998. The FCC established the band in 1999, and similar standards followed in Europe. The MICS band, located in the frequency range of 402 to 405 MHz, is reserved specifically for wireless data communications between implanted medical devices and external equipment. The FCC set aside this band because the signal propagation characteristics in the band are particularly well suited for implantable applications, due to signal-propagation characteristics in the human body, the relative dearth of other users in the band, and the ability to apply the band internationally. The MICS use of this band is secondary to the primary users of this spectrum, that is, the Meteorological Aids Service, the Meteorological Satellite Services, and the Earth Satellite Service.

Technical rules were established to minimize interference and ensure the safe coexistence of multiple MICS devices. The maximum transmitting power is very low, EIRP = 25 μW, in order to reduce the risk of interfering with other users of the same band. The MICS band is broken into 300-kHz-wide channels. The rules specify that devices must listen for other devices before transmitting [i.e., listen-before-talk (LBT)]. If interference is encountered, the radio switches

channels and listens again, a process known as *frequency agility.* The rules also allow MICS devices to transmit without prior frequency monitoring in response to a non-RF actuation signal generated by a device external to the body (i.e., manual activation) or in response to a medical implant event (i.e., alert or alarm condition). Relevant MICS standards are:

- FCC Rules and Regulations, "MICS Band Plan," Part 95, January 2003;
- 47 CFR 95.601-95.673, Subpart E, FCC, 1999;
- ETSI EN 301 839-1, "Electromagnetic Compatibility and Radio Spectrum Matters (ERM);Radio Equipment in the Frequency Range 402 MHz to 405 MHz for Ultra Low Power Active Medical Implants and Accessories; Part 1: Technical Characteristics, Including Electromagnetic Compatibility Requirements, and Test Methods," European Telecommunications Standards Institute, 2002.

The MICS regulations require the system to perform a clear-channel assessment (CCA), in which the user scans all 10 of the 300-kHz channels and is allowed to transmit on the channel with the lowest ambient signal level (i.e., the least noisy channel). The user can also choose to transmit on the first available channel with an ambient power below a certain threshold (as defined in the standard).The MICS standard requires that the external programmer carry out the scanning process. For this reason, the IMD transceiver should support a low-power method of sniffing for the presence of an external programmer signal. MICS regulations provide an exception to the CCA procedure in the event of an emergency medical event. For clinically significant medical emergencies, the IMD may transmit immediately on any channel. For example, if an implanted ECG monitor or pacemaker detects a cardiac arrest, the device could transmit immediately to a monitoring base station, which, in turn, calls an emergency response service.

Wireless medical telemetry enables monitoring equipment to remotely and unobtrusively observe several patients at one time. Such telemetry systems transmit real-time physiologic data, so it is critical to ensure that data are not lost or delayed. More and more radios for nonmedical applications are operating in the ISM bands, increasing the likelihood of signal loss and interference. In response to growing concerns about interference resulting from new digital television transmitters, low-power television transmitters, and greater use of private land mobile radio equipment, the FCC established the Wireless Medical Telemetry Service (WMTS), dedicating bands of frequencies for interference-free operation of medical telemetry systems. The WMTS bands are 608 to 614 MHz, 1,395 to 1,400 MHz, and 1,427 to 1,432 MHz. All transmitters

operating in the WMTS bands must be registered in the database to ensure interference- free operation. Prior to operation, authorized health care providers who desire to use wireless medical telemetry devices must register all devices with a designated frequency coordinator.

The 608- to 614-MHz band is shared with the radio astronomy service. There are 13 radio astronomy sites located throughout the United States. These sites have a protected radius of up to 50 miles. If the proposed WMTS deployment should fall within the protected radius of any of these sites, it is necessary to coordinate with the National Science Foundation (NSF).

The 1,395- to 1,400-MHz band is shared with military radar systems. There are 17 radar sites located throughout the United States; these systems can have a protected radius of up to 55 miles. If the proposed WMTS deployment should fall within the protected radius of any of these sites, it is necessary to coordinate with the NTIA.

A key element of an RF-linked implant is the in-body antenna, which must meet stringent biocompatible and size-limit requirements. An implanted transceiver also faces numerous RF challenges. Unlike free-air performance, the human body is often an unpredictable and hostile environment for a wireless signal. Improving therapy and diagnoses, an implanted pacemaker will regularly transmit performance data and the patient's condition to a doctor's office. If the pacemaker detects a cardiac arrest, the device could signal a base station to alert an emergency response team. Integrated communications from different in-body implants and on-body sensors will also allow hearing for the deaf, sight for the blind, and mobility for the disabled. Using functional electrical stimulus (FES), implants can stimulate muscles or nerves in response to movement detected by sensors elsewhere on the body, allowing a paralyzed patient to walk again. Similarly, a radio-controlled valve for the urinary tract is in development that will be operated on demand to restore bladder control.

There is a growing need for implants, particularly heart implants, to communicate over greater distances than the current rules, at 25 μW EIRP (–16 dBm), allow. The power permitted under MICS accommodates, at most, 6 to 8 feet of separation. For instance, it is increasingly difficult, if not impractical, to position implant monitoring equipment near patients in operating environments when physicians and nurses require unfettered access to patients at all times. In addition, in operating theaters, implant monitoring equipment must be located outside the sterile field, which often means an estimated 30 feet or more between implant and reader. Furthermore, where multiple patients reside in common areas (e.g., nursing homes or hospital wards), independent sessions with individual patients become increasingly economical and convenient, as the distance between implants and programmers increases.

In addition, the spectrum currently allocated to MICS (3 MHz) may not be insufficient to support the variety of implants and data rates that will be

demanded by doctors (and patients) in the years ahead. Because of the reliance on spectrum for functionality of medical devices, and in anticipation of even greater usage, the FCC initiated a proceeding to make more spectra available for medical devices.

2.5.4 Wireless Body Implant Networks

Wireless receivers and transmitters are often arranged in a star configuration, in which a centrally located transceiver communicates with one remote location (a point-to-point topology) or with several remote locations (a point-to-multipoint topology) simultaneously. In a point-to-point architecture, only one transmitter-and-receiver pair is communicating at any given time on a specific carrier frequency. The central node takes the role of a master coordinator, while the remote location is a subordinate. Point-to-point communications can be simplex (one way only), half-duplex (first one direction and then the other, sequentially), and full duplex (simultaneous communication in both directions).

Networks that are more complex may have multiple transmitters and receivers that can communicate with one another as peers. Within a network, one of the central nodes is designated as a coordinator. The coordinator is tasked with waking up other subordinate devices on the network out of a low-current sleep mode just before data is to be transmitted. Coordinator transceivers can also talk to one another as peers. Mesh networks allow wireless devices to talk indirectly to one another, even when the two devices cannot see each other. A transmitting device can pass data to its neighbor, which in turn can pass data onto its next neighbor, and so on.

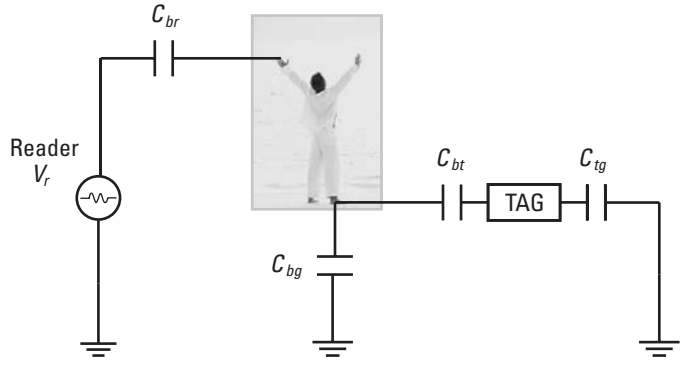
It has been observed that a successful attack on an implanted medical device could disrupt its life-critical functions. Fortunately, the most common forms of vulnerability are due to software-coding errors and thus are potentially preventable. Replay-attack resistant symmetric cryptographic protocols may be sufficient for this application because implanted medical devices will generally only need to communicate with a small number of designated scanner-equipped systems that can be explicitly given copies of an implantee-specific shared secret key.

The U.S. Food and Drug Administration (FDA) has recently approved a surgically implantable RFID device for medical applications in the United States; it contains a microchip transponder, encoded with a unique verification number that can be surgically implanted in a human patient. The implanted microchip can then be scanned to identify the patient and allow access to his or her medical records via the Internet. Its widespread adoption will have to wait until serious privacy concerns are addressed.

2.5.5 Passive Wearable Electrostatic Tags

Passive wearable electrostatic tags, also called *bodytags,* have also been under development for some time. The wearable tag exploits the human body's natural ability to conduct electric fields and allows the wearer to present tags to tag readers through natural motions, such as the grasp of a doorknob or the push of a button. The tag and reader imbue the user's physical gesture with digital meaning [7]. The bodytag is also less expensive than other conventional inductive tags because it contains no magnetic flux coupling coil. The body may be modeled as a conductor surrounded by an insulator. Power and data signals may be coupled electrostatically to the body's interior and sent through it.

The human body acts as a poor conductor connecting the tag and the reader (Figure 2.13). However, displacement current, not dc current, passes through the user's body, allowing the tag and reader to exchange data and power through the body. We call this type of communication *intrabody signaling*. At low frequencies, the human body appears to be a capacitive load; at higher frequencies, the body radiates RF energy. It is possible to send power and data through the body by capacitively coupling displacement current into the body and using the ambient ground reference provided by our environment as the current-return path. The human body is modeled as a solid ideal conductor (the briny interior) surrounded by an ideal insulator (the skin). It is not a good idea to send dc current through the body because it is surrounded by an insulator, not to mention that it could be hazardous to present a constant voltage drop across the interior of the body. However, ac current can be sent through the



C_{rb} = Capacity between the tag reader and the body
C_{bg} = Capacity between the body and ambient ground
C_{bt} = Capacity between the body and the tag
C_{tg} = Capacity between the tag and ambient ground

Figure 2.13 Bodytag circuit model.

body by capacitively coupling to its interior and using it as a single low-impedance node in a network of capacitors.

Each of these capacitances is on the order of 10 to 100 pF. Note that the body couples to one electrode on the tag, while the tag's other electrode couples to the ambient ground. If we put the body tag into the shoe, these electrodes could be the top and the bottom of an inserted pad. Generally speaking, *wireless body-centric networks* consist of a number of nodes and units placed on the human body or in proximity, such as on everyday clothing. Currently it is used to receive or transmit simple information that requires very low processing capabilities. However, some high-performance and complex units are needed in the future to provide the facilities for powerful computational processing with high data rates, for applications such as video streaming and heavy data communications. These have led to increasing research and development activities in the field of body area network applications for many purposes, with the main interest being health care and patient monitoring and task-specific/fully compatible wearable body networks (e.g., wearable computers) that have been applied in fields such as construction and medicine.

Three primary criteria are applied to wireless modules for wireless body-centric networks. First, they must support high data rates and, second, they must be small, both of which suggest the use of high frequencies. Third, they must consume a minimum of power, which implies highly efficient links. In terms of antennas and propagation, efficient design requires good understanding of the properties of the propagation channel involved and the development of optimized antennas.

2.6 Ultrawideband (UWB) Technology

2.6.1 About UWB

Ultrawideband (UWB) is a recently approved technology that relies on extremely short pulses that generate signals with very wide bandwidths, sometimes up to several gigahertz. UWB signals go undetected by most conventional receivers, minimizing their threat as harmful interferers. UWB technologies are currently being used in a variety of applications, such as ground-penetrating radar, and are likely to be used in a variety of emerging applications, such as through-wall imaging and high-speed data transmission.

Gerald F. Ross first demonstrated the feasibility of UWB waveforms for radar and communications applications in the late 1960s and early 1970s [8]. Originally developed by the Defense Advanced Research Projects Agency, the technology was called *baseband, carrier-free, impulse communications or time-domain signaling,* until the U.S. Department of Defense named it *ultra-wideband* in 1989. UWB radios are extremely wideband radios with very

high potential data rates (Figure 2.14). The concept of ultrawideband communications actually originated with Marconi's spark gap transmitter, which occupied a very wide bandwidth. However, because only a single low-rate user could occupy the spectrum, wideband communication was abandoned in favor of more efficient communication techniques. The renewed interest in wideband communications was spurred by the FCC's decision in 2002 to allow operation of UWB devices beneath existing users over a 7-GHz range of frequencies. These systems can operate in the 3.1- to 10.6-GHz range.

The FCC defines a UWB signal as any signal that occupies more than 500 MHz in the 3.1- to 10.6-GHz band and that meets the spectrum mask. This definition replaced a previous one that expressed UWB in terms of fractional bandwidth. The new definition opened up a new way of thinking for several leaders in the UWB community. Given the recent spectral allocation and the new definition of UWB adopted by the FCC, UWB is not considered a technology anymore, but an available spectrum for unlicensed use. This means that any transmission signal that meets the FCC requirements for the UWB spectrum can be considered UWB technology. This, of course, is not just restricted to the impulse radios or high-speed spread-spectrum radios pioneered by companies so far, but to any technology that utilizes more than 500 MHz of spectrum in the allowed spectral mask and with the current emission limit's restrictions.

In theory, the system could interfere with all the systems in that frequency range, including critical safety and military systems, unlicensed systems, such as 802.11 wireless and Bluetooth, and cellular systems, for which operators paid billions of dollars for dedicated spectrum use. The FCC's ruling was quite controversial given the vested interest in the interference-free spectrum of these users. To minimize the impact of UWB on primary band users, the FCC put in

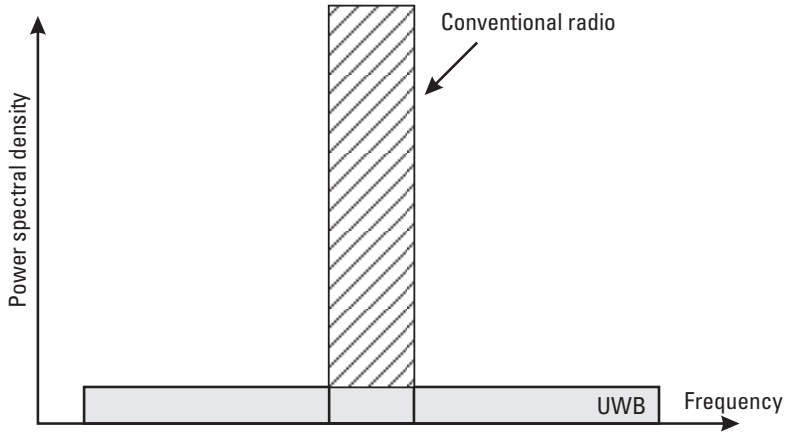


Figure 2.14 Conventional and UWB radio transmission.

place severe transmitting power restrictions. These restrictions require UWB devices to be within proximity of their intended receiver [9].

UWB radios come with unique advantages that have long been appreciated by the radar and communications communities. Their wideband nature allows UWB signals to easily penetrate through obstacles and provides very precise ranging capabilities. Moreover, the available UWB bandwidth has the potential for very high data rates. Finally, the power restrictions dictate that the devices can be small with low power consumption. Initial UWB systems used ultrashort pulses with simple amplitude or position modulation. Multipath can significantly degrade the performance of such systems, and proposals have been put forth to mitigate the effects of multipath equalization and multicarrier modulation. Precise and rapid synchronization is also a big challenge for these systems. Although many technical challenges remain, the appeal of UWB technology has sparked great interest both commercially and in the research community to address these issues.

UWB has several features that differentiate it from conventional narrowband systems:

- The large instantaneous bandwidth enables fine time resolutions for network time distribution, precision location capability, or use as radar.
- Short-duration pulses are able to provide robust performance in dense multipath environments by exploiting more resolvable paths.
- Low power spectral density allows coexistence with existing users and has a low probability of intercept (LPI).
- The data rate may be traded for power spectral density and multipath performance.

On February 14, 2002, the FCC issued a First Report and Order, which classified UWB operation into three separate categories, and each category was allocated a specific spectral mask:

1. Communication and measurement systems;
2. Vehicular radar systems;
3. Imaging systems, including ground-penetrating radar, through-wall imaging and surveillance systems, and medical imaging.

The FCC ruling, however, did not specifically address a precision location for asset tracking or inventory control. These applications, known as *location-*

aware communication systems, are a hybrid of radar and data communications that use UWB pulses to track the two-dimensional and three-dimensional position of an item to accuracies within a few tens of centimeters, as well as transmitting information about the item, such as its contents, to a centralized database system. Note that the FCC has only specified a spectral mask and has not restricted users to any particular modulation scheme. As discussed previously, a number of organizations are promoting multicarrier techniques, such as orthogonal frequency-division multiplexing, as a potential alternative for high-data-rate communications. Beyond the United States, other countries have been using a similar approach toward licensing UWB technology: In both Europe and Japan, initial studies have been completed, and regulations are expected to be issued in the near future that are expected to harmonize with the FCC mask.

UWB systems are approved for unlicensed use within the United States under FCC Part 15, specifically Subpart F and Part 15.250, permitting both indoor and outdoor use. The FCC Part 15.250 band spans from 5,925 to 7,250 GHz. In March 2007, the European Commission (EC) formally adopted a UWB frequency range from 3.4 to 4.8 GHz and 6 to 8.5 GHz, for use in EC-member countries, which will establish several frequency limitations requiring UWB vendors to alter their technology to meet those limits. The EC opinion mandates that all 27 participating European countries accept UWB frequencies for devices used within their borders. Member countries have 6 months to ratify the decision. The EC decision designates the frequency bands of 3.4 to 4.8 GHz and 6 to 8.5 GHz for use by UWB RFID tags and interrogators, as well as for other applications, such as data networking. UWB devices utilizing frequencies between 4.2 and 4.8 GHz will be permitted only until December 31, 2010, by which time they must convert to the 6- to 8.5-GHz band, according to requirements of the Radio Spectrum Committee (RSC), which assists the EC in the development and adoption of measures aimed at ensuring harmonized conditions for the available and efficient use of radio spectrum.

There are two common forms of UWB. One is based on sending very short-duration pulses to convey information and the other uses multiple simultaneous carriers. Each approach has its relative technical merits and demerits. The most common form of multicarrier modulation, orthogonal frequency-division multiplexing (OFDM), has become the leading modulation for high-data-rate systems, and much information on this modulation type is available in recent technical literature. Pure impulse radio, unlike classic communications, does not use a modulated sinusoidal carrier to convey information. Instead, the transmitting signal is a series of baseband pulses. Because the pulses are extremely short (commonly in the nanosecond range or shorter), the transmitting signal bandwidth is on the order of gigahertz.

2.6.2 Orthogonal Frequency-Division Multiplexing

The basic idea behind OFDM is the use of a large number of parallel narrowband subcarriers instead of a single wideband carrier to transport information. OFDM refers to the use, by a single transmitter, of a set of frequency multiplexed signals with the exact minimum frequency spacing needed to make them orthogonal so that they do not interfere with each other. The advantages are that such a system is very efficient in dealing with multipath situations and robust against narrowband interference. Disadvantages include sensitivity to frequency offset and phase noise. Also, the peak-to-average problem reduces the power efficiency of the RF amplifier at the transmitter. Multicarrier communications were first used in the late 1950s and early 1960s for higher-data-rate HF military communications. Since that time, OFDM has emerged as a special case of multicarrier modulation using densely spaced subcarriers and overlapping spectra, and was patented in the United States in 1970. However, the technique did not become practical until several innovations occurred. Fortunately, the apparently very complex processes of modulating (and demodulating) thousands of carriers simultaneously are equivalent to discrete Fourier transform operations, for which efficient fast Fourier transform (FFT) algorithms exist. Thus, integrated circuit implementations of OFDM demodulators are feasible for affordable mass-produced receivers.

Throughout the 1980s and 1990s, other practical issues in OFDM implementation were addressed, such as oscillator stability in the transmitter and receiver, linearity of the power amplifiers, and compensation of channel effects. Doppler spreading caused by rapid time variations of the channel can cause interference between the carriers and held back the development of OFDM until Cimini developed *coded multicarrier modulation.*

The concept of using parallel data transmission by means of frequency-division multiplexing (FDM) was published in mid-1960s while some early development can be traced back to the 1950s. A U.S. patent was filed and issued in January 1970. The idea was to use parallel data streams and FDM with overlapping subchannels to avoid the use of high-speed equalization and to combat impulsive noise and multipath distortion, as well as to fully use the available bandwidth. The initial applications were in the military communications. In the telecommunications field, the terms *discrete multitone* (DMT), *multichannel modulation,* and *multicarrier modulation* (MCM) are widely used and sometimes they are interchangeable with OFDM. In OFDM, each carrier is orthogonal to all other carriers (Figure 2.15); however, this condition is not always maintained in MCM. We could say that the OFDM is an optimal version of multicarrier transmission schemes.

OFDM is also used in asymmetric digital subscriber line (ADSL) services, digital audio broadcast (DAB), digital terrestrial television broadcast (DVB) in

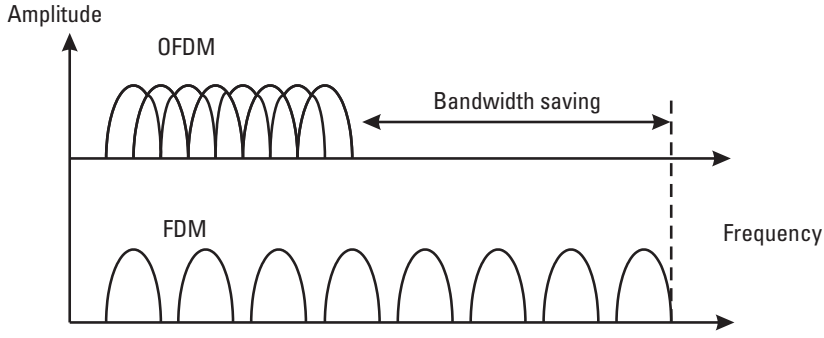


Figure 2.15 Difference between FDM and OFDM.

Europe, integrated services digital broadcasting (ISDB) in Japan, IEEE 802.11a/g, 802.16a, and power line networking (HomePlug). Because OFDM is suitable for high-data-rate systems, it is also being considered for the fourth generation (4G) wireless services, IEEE 802.11n (high-speed WLAN), and IEEE 802.20 (MAN).

2.7 Review Questions and Problems

1. List some of the most unusual applications of the wireless technology you have encountered. Describe the principle of operation, frequencies, and potential issues that users might face right now and in the future.
2. Calculate the wavelength of the signals that have frequency of 125 kHz, 13.56 MHz, 915 MHz, and 2.4 GHz. (*Answers:* 2.4 km, 2.2m, 0.328m, 0.125m.)
3. Although widely accepted as the inventor of radio, Guglielmo Marconi, an Italian engineer working in England, based his work on the patents and inventions of Nikola Tesla. Tesla was an inventor of (among other things) alternate currant, used today in every household (Figure 2.16). In 1911, Tesla refused to share a Nobel Prize in Physics for the invention of radio transmission with Marconi. Tesla's patents from 1900 were reversed in favor of Marconi in 1904 after large private investments (including Edison's) were made in Marconi's company. In 1943, the U.S. Supreme Court upheld Tesla's radio patent from 1900, and as it stands today, *Nikola Tesla is the official inventor of radio*. Although he is often nearly forgotten, Tesla holds more than 700 patents—and those are just the ones he remembered to patent.

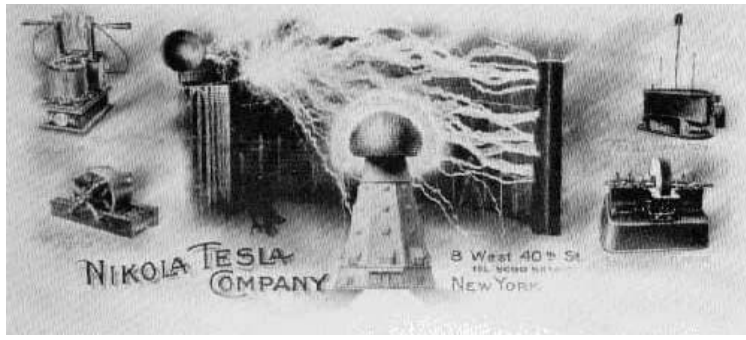


Figure 2.16 Nikola Tesla Company.

Write an essay describing the productive and very dynamic life of the genius and eccentric, Nikola Tesla.

4. Although it is possible to continually increase the EIRP of a system in order to achieve a path of any length, it is not necessarily desirable to do so. As the transmitting power of a system increases, the potential of the system to cause interference to other services also increases, which limits the use of spectrum in geographically adjacent areas. It is necessary to strike a balance between the need of one user for increased power for his or her system, and the need of another user for access to a channel to establish a service. During the 1920s, radio communication was a veritable free-for-all; anyone possessing radio equipment was allowed to broadcast signals over the air, resulting in chaos. Because interference resulted any time several transmitters operated in proximity, no one could be assured of reliable communications. By the early 1930s, radio sales and usage plummeted, and the market failure created by this chaos predestined today's regulatory environment. Accordingly, with the passage of the Communications Act of 1934, Congress created the Federal Communications Commission to regulate radio communications in the United States, the District of Columbia, and all U.S. possessions [10]. The FCC has historically controlled access to the radio spectrum by allocating specific frequency bands for use by licensed service providers.

   Today, many applications take advantage of the so-called license-exempt frequency bands. Research different applications and advantages and disadvantages of wide use of these (unlicensed) frequency bands. Show your results in a form of a table and briefly summarize your findings.

5. Verify whether the following two statements are correct and justify the answer:

a. A horizontally positioned, linearly polarized, receiver antenna is unlikely to capture any of the energy emitted by a vertically positioned, linearly polarized, transmitter antenna.

b. A circularly polarized transmitter antenna will be able to communicate with any receiver antenna, regardless of its orientation.

6. List some of the ethical issues in the modern wireless (E911, for example) communications systems. Discuss the topic from different points of view (for example, from that of an engineer, philosopher, doctor, clergyman, politician, or businessperson).
7. UWB, a technology that was just recently approved by the FCC for a number of communications and sensing applications, is a signaling method that relies on short pulses that create extremely wide bandwidths. In addition to their potential for communications systems, UWB technology can also support the operation of new low-power radar products that can provide precise measurement of distances or detection of objects underground or behind walls or other structures.

   Find at least three suppliers of commercial systems based on the UWB technology. Describe their products and systems.
8. If you were a person in charge of the decision about whether wireless human body implants should continue to be developed and used on people, what would you decide? Can you justify your decision in a way that is acceptable to all of the people all of the time?
9. What factors determine the range of a wireless link? Discuss the influence of each individual factor.
10. How would you shape a wire of fixed length to obtain the greatest and the smallest inductances?
11. There are many types of antenna radiation patterns, but the most common are omnidirectional, pencil beam, fan beam, and shaped beam. Discuss them in more detail. More information can be found in [11].

References

[1] Best, S., "Antenna Polarization Considerations in Wireless Communications Systems," Cushcraft Corporation, Manchester, NH, 2003.

[2] Ergen, S. C., *ZigBee/IEEE 802.15.4 Summary,* Internal Report to Advanced Technology Lab of National Semiconductor, Berkeley, CA, 2004.

[3] Stankovic, J. A., "Wireless Sensor Networks," Charlottesville, VA: Department of Computer Science, University of Virginia, June 19, 2006.

[4] Yang, G. -Z., *Body Sensor Networks,* New York: Springer-Verlag, 2006.

[5] Sauer, C., et al., "Power Harvesting and Telemetry in CMOS for Implanted Devices," *IEEE Trans. on Circuits and Systems,* Vol. 52, No. 12, December 2005.

[6] Miura, N., et al., "Analysis and Design of Inductive Coupling and Transceiver Circuit for Inductive Inter-Chip Wireless Superconnect," *IEEE J. of Solid-State Circuits,* Vol. 40, No. 4, April 2005.

[7] Nivi, B., et al., "Passive Wearable Electrostatic Tags: The Bodytag," Cambridge, MA: Physics and Media, MIT Media Lab, September 12, 1997.

[8] Bennett, C. L., and Ross, G. F., "Time-Domain Electromagnetics and Its Applications," *Proc. IEEE,* Vol. 66, No. 3, March 1978.

[9] Reed, J., *An Introduction to Ultra Wideband Communication Systems,* Upper Saddle River, NJ: Prentice-Hall, 2005.

[10] Carter, K. R., A. Lahjouji, and N. McNeil, "Unlicensed and Unshackled: A Joint OSP-OET White Paper on Unlicensed Devices and Their Regulatory Issues," Washington, D.C., Federal Communications Commission, Office of Strategic Planning and Policy Analysis, OSP Working Paper Series, May 2003, p. 61.

[11] Volakis, J. L., *Antenna Engineering Handbook,* 4th ed., New York: McGraw-Hill, 2007.

3

Automatic Identification Systems

The technologies used in the world of automatic identification and data capture (AIDC) are varied and often used in combinations to provide a broader base of information flow. This chapter attempts to summarize the technologies in common use today, and give the reader a basic understanding of the technology and its uses and limitations.

3.1 Barcodes

Perhaps the oldest of the AIDC technologies, barcode technology, can be looked on as the best known and probably most successful to date. We are all familiar with the basic barcode on our box of cereal or the jar of honey that we buy in the supermarket. This barcode is referred to as the UPC/EAN and is but one variation of more than 250 barcodes that have been designed over time. Barcodes like this are referred to as *linear* barcodes because they are made up off a collection of bars and spaces side by side (Figure 3.1). Fortunately, many of these barcodes have never gained broad acceptance, so only about 10 or 12 linear barcodes are in common use. The typical data content capacity varies from 8 to 30 characters, with some barcodes restricted to numerals only and others using full alphanumeric information.

Linear barcodes are used in applications where the use of a simple numeric or alphanumeric code can provide the key to a database of products. The most obvious limitation is the amount of data that can be stored in a linear barcode, though other problems can exist with the substrate on which the barcode is printed. The substrate might provide insufficient contrast or poor ink receptivity, which can cause the quality of the barcode to be less than ideal.

Figure 3.1 Linear and two-dimensional barcodes.

A new growth area in the world of barcodes is the two-dimensional versions. Several variations of two-dimensional barcodes are available, but because these do not all comprise bars and spaces, the more accurate name of two-dimensional *symbologies* is used. Two-dimensional symbologies provide a means of storing large amounts of data in a very small space. Various types include stacked symbologies (linear barcodes stacked on top of each other), matrix symbologies (comprised of a matrix of light and dark elements, circles, squares, or hexagons), and packet symbologies (a collection of linear symbols "randomly" arranged on a page). Examples of the three types include PDF417, Code 49 Code 16K (stacked), Code One, MaxiCode, Data Matrix, Aztec Code, QR Code (matrix), and Super Code (packet).

Two-dimensional symbologies have a major advantage over linear barcodes: They can store vast amounts of data. Individual symbols can store as many as 7,000 numeric-only or 4,200 alphanumeric characters. Many of the symbologies also have the ability to use a device called *structured append* that allows messages to be split over multiple symbols, providing almost infinite storage space. The disadvantage of the two-dimensional symbologies is that a special scanner is needed. Matrix symbologies need a vision-based scanner to read the data, although some of the stacked symbologies can be read with a special laser scanner.

3.2 Card Technologies

3.2.1 Magnetic Cards

The first *magnetic stripe cards* were used in the early 1960s on transit tickets and in the 1970s for bank cards. Since then, the use of magnetic stripes has continued to grow. Credit cards were first issued in 1951, but it wasn't until the establishment of standards in 1970 that the magnetic stripe became a factor in the use of the cards. Whether the card is a credit card–sized plastic card, a thin paper ticket, or an airline boarding card, the uses for magnetic-stripe technology have grown considerably. Today, with an infrastructure that provides every store in the mall with the ability to read the information on the magnetic stripe, the technology is everywhere. Although some limitations exist on the amount of

information that can be stored on the stripe and the security of the data, solutions from various vendor exist to solve these problems.

With the advent of new technologies many people have predicted the demise of the magnetic stripe. However, with the investment in the current infrastructure, this is not likely to happen any time soon. Magnetic stripe technology provides the ideal solution to many aspects of our lives. It is very inexpensive and readily adaptable to many functions, meaning that many applications can expect to be using magnetic stripe technology for quite a few years to come.

3.2.2 Smart Cards

Smart cards are not new. Invented in 1974 by French journalist Roland Moreno (http://www.rolandmoreno.com) and first used as prepaid phone cards in 1984, smart cards are credit card–sized pieces of plastic containing a data storage system. The technology was rapidly accepted in Europe because the high cost of telecommunications made online verification of transactions very expensive. The smart card provided the mechanism to move that verification offline, reducing the cost without sacrificing any of the security. The first plastic cards were used in the United States for club membership.

Several terms are used to identify cards with integrated circuits embedded in them. The terms *chip card, integrated circuit card,* and *smart card* really all refer to the same thing. There are two types of smart cards. The first type only contains memory and is used to store information; examples of this might include stored value cards, in which the memory stores a dollar value that the user can spend in a variety of transactions, such as in pay phones or vending machines. The second type of card is a true "smart" card, in which a microprocessor is embedded in the card along with memory. Now the card actually has the ability to make decisions about the data stored on the card. The card is not dependent on the unit to which it is attached to make the application work. Because the card has a microprocessor, various methods can be used to prevent access to the information on the card, providing a secure environment. This security has been touted as the main reason that smart cards will replace other card technologies.

The microprocessor-type smart card comes in two flavors: the *contact version* and the *contactless version.* Both types of cards have the microprocessor embedded in the card; however, in the contactless version, the gold-plated contacts are not visible on the card. The contactless card uses a technology to transfer data between the card and the reader without any physical contact being made. The advantage to this contactless system is that that the components are completely embedded in the plastic with no external connections, meaning there are no contacts to wear out and no chance of an electric shock coming

through the contacts and destroying the integrated circuit. The disadvantage to this system is that the card and reader are more sophisticated and, hence, more expensive. The biggest disadvantage today with smart cards is the cost to create a smart-card system. Individual card prices have fallen during the past few years, but they are still high when compared with a magnetic stripe card. The biggest advantage is, of course, the amount of data that can be stored and the security that can be built into the card.

3.2.3 Optical Cards

Optical memory cards use a technology similar to the one used for music CDs or CD-ROMs. A panel of the gold-colored laser-sensitive material is laminated on the card and used to store information. The material is comprised of several layers that react when a laser light is directed at them. The laser burns a tiny hole (2.25 μm in diameter) in the material, which can then be sensed by a low-power laser during the read cycle. The presence or absence of the burn spot indicates a 1 or a 0. Because the material is actually burned during the write cycle, the media is write-once, read-many (WORM) media, and the data is nonvolatile (i.e., not lost when power is removed). The optical card can currently store between 4 and 6.6 MB of data, which gives the ability to store graphical images such as photographs, logos, fingerprints, X-rays, and so forth.

The major disadvantage of the optical card is the fact that it is a write-once technology, so the amount of data storage available is used up with every piece of new data written. In some applications, this can be considered an advantage because it maintains the complete history of changes made to the card.

3.3 Radio-Frequency Identification

3.3.1 RFID Historical Background

The birth of radio-frequency identification technology was in October 1948 after the publication of a paper by Harry Stockman titled "Communications by Means of Reflected Power." The popular system Identification, Friend or Foe (IFF), for aircraft, was one of the first applications of RFID technology [1]. In early 1940, the British Royal Air Force outfitted airplanes with radio transponders that would respond when interrogated. This allowed pilots and ground crews to distinguish the RAF airplanes from the Luftwaffe's, which proved to be a decisive advantage in the Battle of Britain.

In the 1960s, the electromagnetic theory related to RFID applications was developed, and this was the prelude to the RFID explosion. Commercial activities exploiting the RFID also began during the 1960s, and the electronic article

surveillance (EAS) application is one example. The EAS is a simple 1-bit tag, because only the presence or the absence of a tag can be detected. During the rapid development of microelectronic technology during the 1970s, companies, universities, and government laboratories were actively engaged in the development of practical applications of RFID, such as animal tracking, vehicle tracking, and factory automation. The 1980s was the decade of mass deployment of RFID technology. Interest in the United States was mainly for transportation and access control, whereas in Europe the greatest interests were for animal tagging, industrial applications, and toll roads. Since the 1990s, many technological developments have dramatically expanded the functionality of RFID technology. Advances in microelectronics, embedded software, and microwave-circuit integration are opening the door to a new wireless system and expanding the application field for RFID. (*Note:* A search of the U.S. Patent Office alone will reveal more than 350 patents related to RFID and its use.)

Benefits of RFID technology are that it allows manufacturers, retailers, and suppliers to efficiently collect, manage, distribute, and store information on inventory, business processes, and security controls. RFID will allow retailers to identify potential delays and shortages; grocery stores to eliminate or reduce item spoilage; toll systems to identify and collect auto tolls on roadways; suppliers to track shipments; and in the case of critical materials, RFID will allow receiving authorities to verify the security and authentication of shipped items. These uses are seen as only the beginning, and as RFID is deployed across different sectors and services, increasing its efficiency and visibility, several other applications and benefits may arise.

3.3.2 RFID System Overview

In general terms, RFID represents a way of identifying objects or people using radio waves. Identification is possible by means of unique numbers that identify objects, people, and information, stored on microchips, which can be read automatically, unlike barcodes that need to be scanned manually. With the recent advancements in RFID technology, the automatic identification data capture industry is accelerating its efforts to identify new applications to take advantage of RFID. RFID is fundamentally based on wireless communication, making use of radio waves, which form part of the electromagnetic spectrum, and it is not unlike two other wireless technologies, WiFi and Bluetooth. The three technologies are all designed for very different uses and therefore have different functionalities, but there is shared ground between the three, with some hybrids starting to appear. RFID systems can utilize both WiFi and Bluetooth and need not see them as competitors.

All RFID systems are comprised of three main components:

- *RFID tag*, or transponder, which is located on the object to be identified and is the data carrier in the RFID system;
- *RFID reader*, or transceiver, which may be able to both read data from and write data to a transponder;
- *Data processing subsystem*, which utilizes the data obtained from the transceiver in some useful manner.

The essential requirement in an RFID system is to transfer data stored in a tag to a reader across a wireless air interface (the region between the tag and the reader). A two-way communication process is required to do this and requires a radio carrier signal suitably modified (modulated) to carry the data. The concept is depicted in Figure 3.2. The RFID concept works as follows: A reader transmits a signal that is received by an antenna integrated with a small RF chip. In general, the chip is activated only when an RFID reader scans it. When the chip wakes up, it sends the unique identifier number, which the reader passes along to applications such as inventory control and shipping. It is this level of application logic that provides the selection of particular tag and data manipulation criteria for those identified tag(s). Using a host computer application program, specific tags can be selected to be identified and the data contained within those tags to be acted on while those tags are active within the RF portal (interrogation area).

The *RF portal* is defined as the area where RFID tags can be read or written to; it can be stationary or mobile. *Stationary portals* are used mainly in

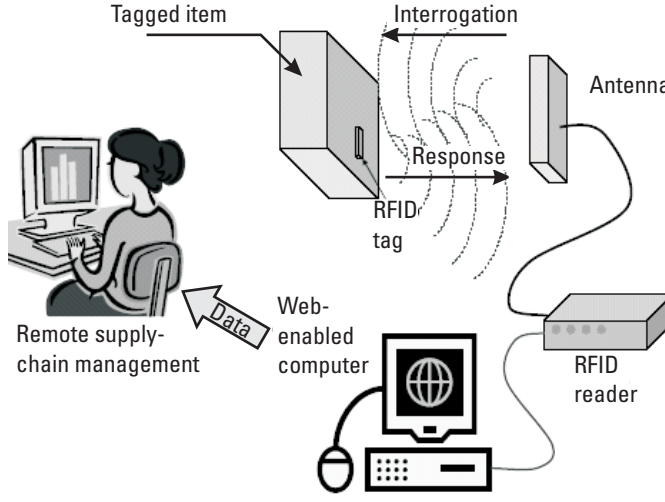


Figure 3.2 Basic components of an RFID system.

applications where the item containing the RFID tag has to follow some prescribed physical path or flow. An example of this would be where warehouse goods flow through a dock door, or with items that travel down a conveyor or assembly line. *Portable* or *mobile interrogator portals* are used in applications where the tagged items do not follow a predefined path. This type of RFID interrogator can be used in conjunction with a portable computing device, such as a portable data terminal or mobile pen computer. Typical applications include asset tracking, picking or moving inventory, inspection, and quality control. In portable or mobile applications, the interrogator is aimed into a certain physical area where RFID tags need to be scanned. It is more flexible in application scope because the user can define the RF portal by orienting the longitudinal and latitudinal axis of the interrogator. However, because this device operates on batteries, it is limited in its effective range of scanning, or RF portal depth of field. The energized period of operation of this type of interrogator must also be controlled so that battery life can be maximized.

There are many potential applications for RFID; the most obvious one is as a more robust replacement to barcodes. However, innovative companies are regularly finding new applications for the enhanced range, capacity, and read/write capability. The RFID readers can be big enough for forklifts to pass through or small enough to fit on retail shelves.

RFID operates in unlicensed spectrum space, sometimes referred to as ISM, but the exact frequencies that constitute ISM may vary, depending on the regulations in different countries. Typical carrier frequencies (the reader's transmitting frequency) in today's applications range from 125 kHz to 2.45 GHz (with 5.8 GHz also being considered). The frequency bands must be selected carefully for applications because each one has its own advantages and disadvantages (in the case of RFID we can truly say that one size does not fit all). The RFID systems themselves can achieve high levels of complexity, having incorporated memory, data processing capabilities that include communication encryption, and protocols (Figure 3.3). As we will see later, the air interface (communication between the reader and the tag), physical interrogator, data protocol processor, and application commands and interfaces are all covered by the different, but related, standards.

A high-level description of the sequence of communication is as follows:

- Host manages reader(s) and issues commands.
- Reader and tag communicate via RF signal:
  - The reader continuously generates an RF carrier sine wave, watching always for modulation to occur. Detected modulation of the field would indicate the presence of a tag.
  - Carrier signal sent out through the antennas.

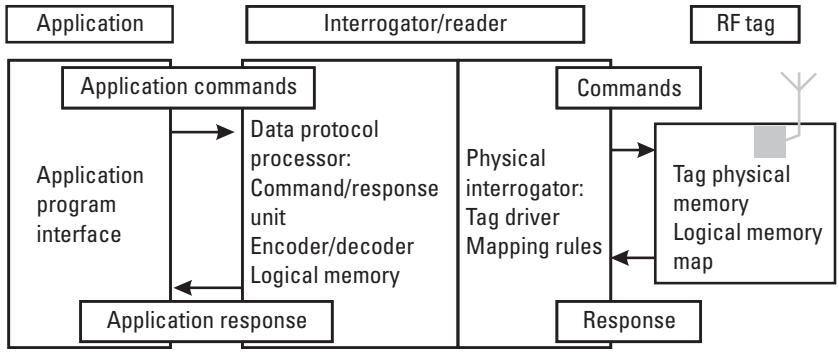


Figure 3.3 RFID system overview.

- Carrier signal hits tag(s) (Figure 3.4).
- Once the tag has received sufficient energy to operate correctly, it divides down the carrier and begins clocking its data to an output transistor, which is normally connected across the coil inputs.
- Tag receives and modifies the carrier signal and sends back a modulated signal (passive backscatter, which the FCC and ITU refer to as a *field disturbance device*). The tag's output transistor shunts the coil, sequentially corresponding to the data that is being clocked out of the memory array. Inductive coupling and load modulation are used at lower frequencies, as opposed to systems operating at 2.45 GHz and higher bands, where true RF communication links and backscatter principles are used.
- Shunting the coil causes a momentary fluctuation (dampening) of the carrier wave, which is seen as a slight change in amplitude of the carrier.

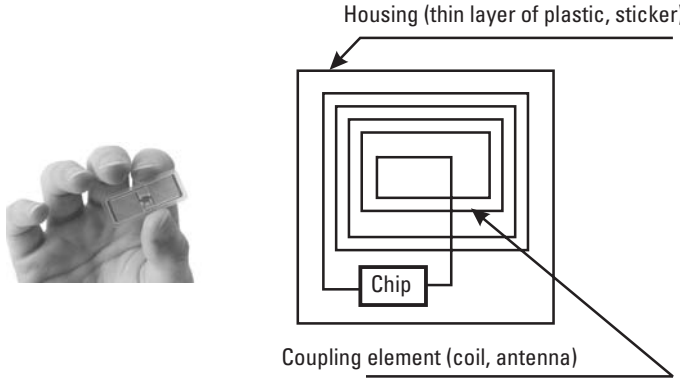


Figure 3.4 RFID tag.

- Antennas receive the modulated signal and send them to the reader.
- Reader decodes the data. The reader peak-detects the amplitude-modulated data and processes the resulting bitstream according to the encoding and data modulation methods used.

- Results are returned to the host application.

The general requirements for the tag antenna are small size, high efficiency, sufficient antenna beamwidth for reduction of orientation sensitivity, simplicity, and low manufacturing cost.

3.3.3 Principles of RFID Operation

The coupling between tag and reader is achieved in one of two ways, depending on the carrier frequency used and the system and antenna design (Figure 3.5). Low- (<135-kHz) and high-frequency (typically 13.56-MHz) systems invariably use reactive (typically inductive) coupling, wherein the predominantly magnetic field component carries the data in the communication between tag

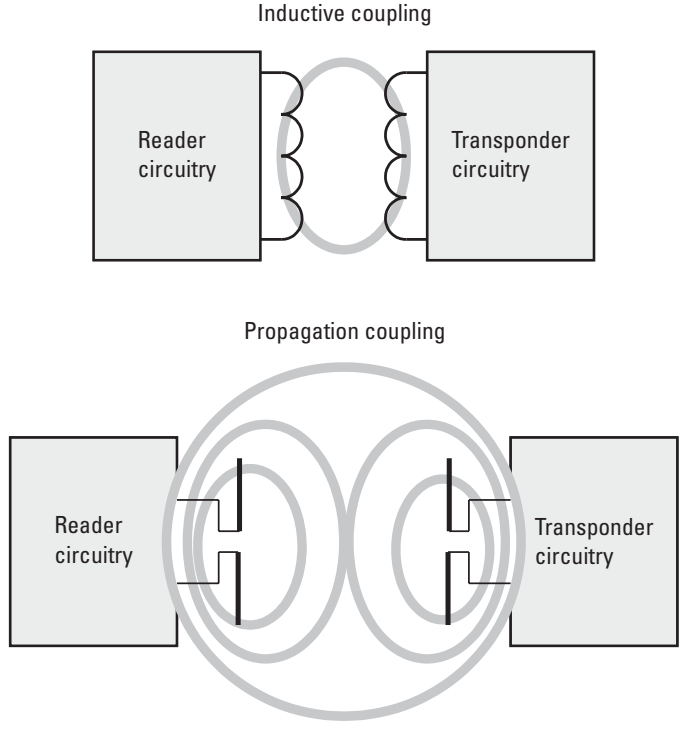


Figure 3.5 Inductive- and propagation-coupling RFID systems.

and reader, in much the same way as coupling occurs between primary and secondary coils in an air-cored transformer. In these systems the field is effectively tied to its source, and the field that couples with the tag is modulated by means of the tag circuitry, such that the data-related changes can be sensed by the reader.

The second form of coupling is by propagation of the electromagnetic field used to read or interrogate the tag. In these systems field components dissociate from their source in the reader and propagate into free space. The components of an RFID system that largely determine whether an RFID system couples by inductive or propagation means are the antenna and the manner in which it is driven in electrical terms. Loosely speaking, where the dimensions of the antenna are small in relation to the carrier wavelength, the associated electromagnetic field typically exhibits both reactive and propagation components. For low- and high-frequency RFID systems, the antennas are structured and driven in such a way that the propagation component is small or even nonexistent, while the reactive component is predominant. At ultrahigh frequencies and above the systems are essentially structured and driven to operate in the propagation mode, the antenna dimensions being matched or appropriately related to the carrier wavelength to achieve the desired result. In these systems, the reactive component is designed to be suitably small.

The *far field* begins where the *near field* ends, although not abruptly, at a certain distance from the transmitting antenna. In the near field, tag-to-reader communication is achieved via load modulation. Load modulation is achieved by modulating the impedance of the tag as seen by the reader. In the far field, tag-to-reader communication is achieved via backscatter. Backscatter is achieved by modulating the radar cross section of the tag antenna.

3.3.3.1 Inductive Coupling and Load Modulation

The data transfer from the transponder to the reader can be operated in three different ways: load modulation, load modulation by using a subcarrier (frequencies below 30 MHz), or by a subharmonic procedure (above 100 MHz). When a tag is placed within the alternating magnetic field created by the reader, it draws energy from the magnetic field. This additional power consumption can be measured remotely as a voltage perturbation at the internal impedance of the reader antenna. The periodic switching on/off of a load resistance at the tag therefore affects voltage changes at the reader's antenna and thus has the effect of an amplitude modulation of the antenna voltage by the remote tag. If the switching on and off of the load resistance is controlled by the tag's stored data stream, then this data is transferred from the tag to the reader. This type of data transfer is called *load modulation*. In load modulation the carrier signal is modulated by switching impedance from a matched condition to an unmatched condition to alter the reflection coefficient.

The incident radio electromagnetic (EM) wave is received by the tag through the antenna. The radiated energy is then converted to electrical current and travels down a transmission line configuration with intrinsic impedance Z_0. At the end of the transmission line, the electric waveform is met with a PIN diode, represented as a switch in Figure 3.6, and used to toggle between two types of load impedances. When the diode is forward biased, the current is allowed to flow through the matched load making $Z_L = Z_0$, and thus causes the reflection coefficient to equal zero. When the diode is reverse biased, the load impedance essentially become infinite and makes the reflection coefficient equal to 1. In the first case, all of the forward-traveling current is absorbed by the load and no power is sent backward through the transmission line. In the latter, no power is absorbed, and all of it is reflected back down the line. In either case, the reflected power is propagated down the transmission line and radiated out through the antenna. When all of the incident power is reflected, the tag transmits a logical bit 1. When no power is transmitted, the tag transmits a logical bit 0. These bits are propagated through the modulated backscatter, and ride on top of the reflected wave to the receiver. Several methods are used to encode the data onto the carrier wave.

The process of load modulation creates amplitude-modulated sidebands symmetrically placed around the 13.56-MHz interrogation carrier frequency. Because the coupling between reader antenna and tag is relatively weak and the voltage change created by the tag leads to relatively poor signal-to-noise ratios, reply code modulation with a subcarrier is utilized in most RFID chips. In this improved signaling method, the tag's data reply information is contained in a

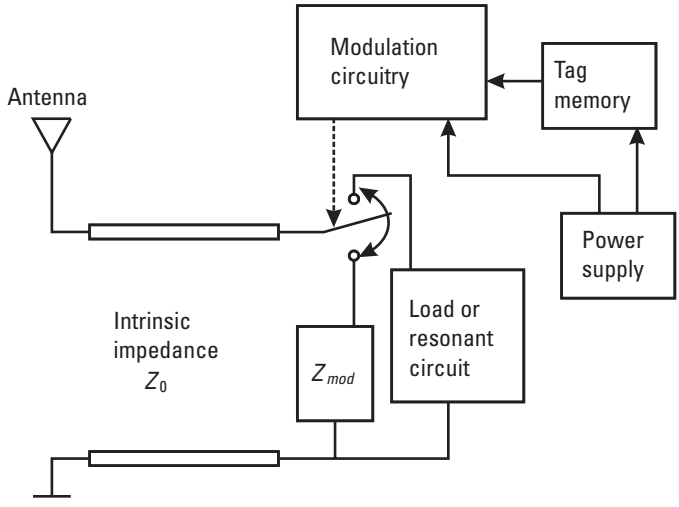


Figure 3.6 Load modulation circuitry.

pair of backscattered sidebands, which are subsequently demodulated in the RF and baseband signal processing sections of the reader to recover the tag's data stream.

In the *subharmonic procedure*, a second frequency (which is usually lower by a factor of 2) is derived by digital division by 2 of the reader's transmission frequency. The output signal of a binary divider can be modulated with the data stream from the transponder. One popular operating frequency for subharmonic systems is 128 kHz. This gives rise to a transponder response frequency of 64 kHz.

3.3.3.2 Propagation Coupling and Backscatter Modulation

The term *backscatter modulation* refers to the communication method used by a passive RFID tag to send data back to the reader. By repeatedly shunting the tag coil through a transistor, the tag can cause slight fluctuations in the reader's RF carrier amplitude. The RF link behaves essentially as a transformer; as the secondary winding (tag coil) is momentarily shunted, the primary winding (reader coil) experiences a momentary voltage drop. The reader must peak-detect this data at about 60 dB down (about 100 mV riding on a 100-V sine wave). This amplitude modulation loading of the reader's transmitted field provides a communication path back to the reader. The data bits can then be encoded or further modulated in a number of ways.

We know from the field of radar technology that electromagnetic waves are reflected by objects with dimensions greater than around half the wavelength of the wave. The efficiency with which an object reflects electromagnetic waves is described by its reflection cross section. Objects that are in resonance with the wavefront that hits them, as is the case for an antenna at the appropriate frequency, for example, have a particularly large reflection cross section.

Figure 3.7 shows a block diagram of a typical passive RFID reader/tag configuration. The reader radiates an unmodulated signal (called the incident wave), which impinges on the tag. The tag antenna intercepts this signal, absorbs part of it, and reradiates the rest. A switch (field effect transistor) is connected across the antenna, which when closed, causes a mismatch in the tag antenna. The mismatch causes a small percentage (say, 10%) of the signal to be absorbed, which in turn results in the other 90% being reradiated. The switch, also known as a modulator, is controlled by the tag output data stream and causes the tag data to be modulated onto the incident wave and be reradiated (or backscattered) as a modulated signal. This technique is known as *impedance-modulated backscatter* or *backscatter modulation* for short. The amount of energy intercepted and reradiated is determined by the *differential radar cross section* of the tag, which is, in turn, a function of the tag antenna aperture and modulation depth.

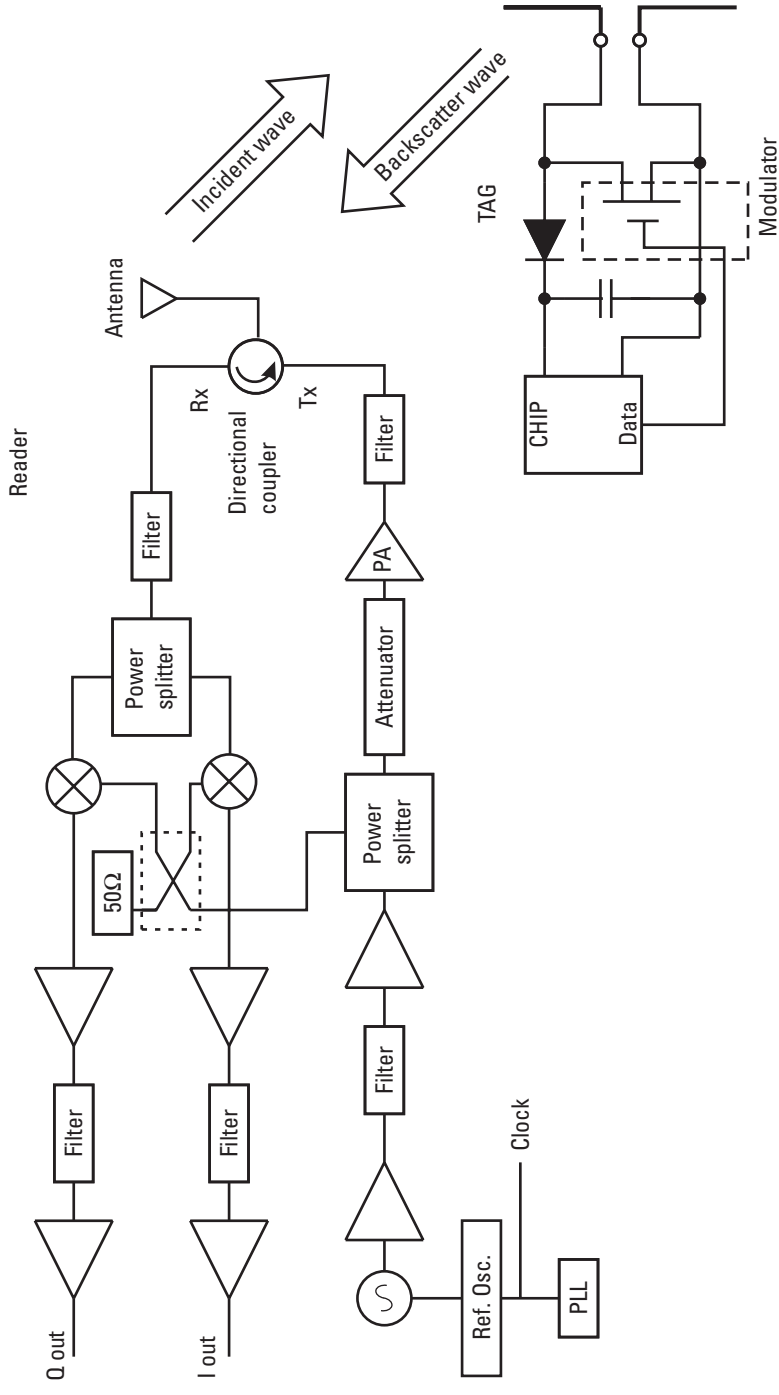


Figure 3.7 Backscatter modulation circuitry.

The operating range of the tag is determined by the reader's transmitted power, the reader's receiver sensitivity, the path loss in both directions between the reader and tag, and the differential radar cross section of the tag. The backscattered signal received at the reader receiver decreases as the fourth power of the distance between the reader and tag (the square of the power in each direction.) For example, a typical 900-MHz tag will have a radar cross section of at least 0.005 m^2 and typically 0.024 m^2. Present-day receiver technology allows a reader to reliably decode a signal having a strength of –80 dBm or higher. This means that the signal arriving back at the reader from a tag must be stronger than –80 dBm.

The ratio of power transmitted by the reader (incident wave) and power returning from the transponder (backscatter wave) can be estimated using the radar equation, a detailed explanation of which can be found in Chapter 5.

The reader's RF transceiver block diagram shown in Figure 3.7 uses the homodyne topology and represents the classic approach for backscatter radar where received signals are close in frequency to the transmitted carrier. A homodyne receiver performs a direct conversion of the received RF signals to a zero intermediate frequency (IF) baseband. Elimination of the IF reduces the need to perform image rejection filtering as is typical with other receiver approaches, such as the superheterodyne approach. This architecture uses two high-compression point MMIC mixers in quadrature along with lowpass filtering. Reducing continuous wave (CW) carrier leakage into the RF port of the mixers is critical to receiver performance since the resulting phase discriminator generates unwanted dc signal offsets. Leakage of the high-power transmitting signal into the receiving path can saturate the receiver and significantly degrade receiving signal sensitivity.

The power is supplied to the tag's antenna connections as HF voltage and after rectification by the diodes this can be used as turn-on voltage for the deactivation or activation of the power saving power-down mode. The diodes used here are low-barrier Schottky diodes, which have a particularly low threshold voltage. The voltage obtained may also be sufficient to serve as a power supply for short ranges.

Functions performed by the reader may include quite sophisticated signal conditioning and parity error checking and correction. Once the signal from a transponder has been correctly received and decoded, algorithms may be applied to decide whether the signal is a repeat transmission, and may then instruct the transponder to cease transmitting. This is known as the Command Response Protocol and is used to circumvent the problem of reading multiple tags, in a short period of time. Using interrogators in this way is sometimes referred to as *hands-down polling*. An alternative, more secure, but slower tag polling technique is called *hands-up polling*,, which involves the interrogator looking for tags with specific identities, and interrogating them in turn. This is contention

management, and a variety of techniques have been developed to improve the process of batch (multiple tags) reading.

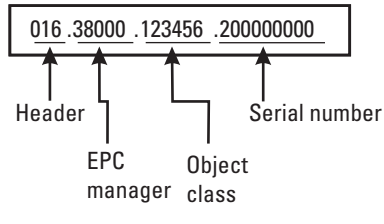
3.3.4 The Electronic Product Code System

Electronic product code (EPC) is an item numbering and networking concept that emerged from research undertaken at the Massachusetts Institute of Technology (MIT) Auto-ID Center, and associated centers, in which RFID data carriers were identified as the method of choice for carrying EPC numbers [2].

3.3.4.1 Electronic Product Code

The RFID technology itself offers several improvements over its predecessor technologies, the barcode and magnetic stripe cards. The central data feature of RFID technology is the electronic product code, which is viewed by many in the industry as the next generation barcode or Universal Product Code (UPC). The EPC is a 96-bit code created by the Auto-ID Center that would one day replace barcodes (Figure 3.8). The EPC has digits to identify the manufacturer, product category, and the individual item. It is backed by the United Code Council and EAN International, the two main bodies that oversee barcode standards. This EPC can carry more data than a UPC and can be reprogrammed with new information if necessary. Like the UPC, the EPC consists of a series of numbers that identify the manufacturer and product type. The EPC also includes an extra set of digits to identify unique items.

The EPC numbering scheme, initially comprising a 96-bit code structured as an 8-bit header, is followed by three data partitions for the EPC Manager (28 bits to facilitate identification of the item manufacturer or source provider for example), the Object Class identifier (24 bits to facilitate the identification of the type of product, such as a specific stock keeping unit), and serial number (36 bits to facilitate the unique identification of an individual item). The header,



Header: identifies the length, type, structure, version, and generation of EPC
Manager number: identifies the company
Object Class: similar to a stockkeeping unit or SKU
Serial number: specific instance of the Object Class being tagged

Figure 3.8 Electronic product code.

also known as the EPC version number, is used to distinguish multiple EPC formats, thus allowing designation of differing bit length tags as the technology matures (a 256-bit version has been proposed). The header can also be used to distinguish bit-length field variations to those indicated above with respect to manufacturer, product, and serial number support, thus allowing longer and more manufacturer identifiers for organizations with a small number of product types and serial number requirements. While the initial bit-length designation for EPC was 96 bits, a shorter 64-bit version has been introduced on an interim basis to help facilitate the realization of lower cost RFID data carrier devices. This seemed reasonable on the basis that the full identification capability of the 96-bit version would not be required for some time.

3.3.4.2 Object Naming Service

The EPC object naming service (ONS) was introduced to provide a directory service capable of supporting the linking of EPC numbers with additional data or information concerning the item to which the EPC tag is attached. This additional, item-associated data or information may be stored on a server connected to a local network or the Internet. The ONS is analogous to the domain name service used for location of information on the Internet.

3.3.4.3 Physical Markup Language

Physical Markup Language (PML) is structured to allow the information about an item or object to be appropriately specified. The PML is based on the popular XML metadata language, its syntax, and semantics to be administered and developed by a governing body (EPCglobal) in conjunction with the user community. Product definitions within this language markup facility, which began with food items, require the ongoing efforts of the governing body to build a sufficiently inclusive directory. Product descriptions already undertaken by standards bodies, such as the International Bureau of Weights and Measures and the National Institute of Standards and Technology (NIST), are being seen as valuable sources of information in this respect. In addition to fixed product information, the PML will also accommodate dynamic quantities, such as temperature, humidity, or vibration, which may change as a result of some local, environmental, or intrinsic effect, including changes over time (temporal effects). This adds a further dimension to the data gathering and handling processes and, when presented in a PML file, may offer innovative opportunities for process enhancement. For example, condition status information derived dynamically could be used to automatically determine product pricing.

3.3.4.4 Savant Software

Savant is the specification for standard RFID middleware, that is, software that bridges RFID hardware and enterprise applications. It defines an EPC event

handling framework and is thus the primary means of data gathering for any RFID deployment. It acts as the central nervous system of the EPCglobal Network [3]. Its most basic function is to receive the EPC number and direct a query over the Internet or other established network to the ONS, which then returns an address at which the item information is stored. The information is available to, and can be augmented by, Savant systems within the network, ostensibly around the world. The very high data handling envisaged within the EPC infrastructure indicates the potential need for companies to maintain ONS servers locally to support rapid retrieval of information. The Savant software is being developed to use a distributed architecture with a hierarchical structure to manage data flow. A highly extensive network of Savants is envisaged to support EPC data management, with Savant platforms running, for example, in factories, stores, distribution centers, regional support facilities, and even on mobile platforms such as container lorries and cargo planes. Creating such an infrastructure is one of the biggest challenges to realizing the EPC support objectives.

3.3.5 UWB and RFID

In recent years, many proposals have been made to address the privacy and security issues of the RFID systems. One of the proposals is to implement the link from RFID tag to the reader using UWB communications, since the use of an advanced modulation scheme offers a new approach to the RFID security [4]. By using the modulation spreading code as a secret parameter of the communications link, it is possible to make eavesdropping extremely difficult and therefore increase communication reliability. UWB is a license-free, low-power radio transmission scheme with an enormous potential bandwidth of 7.5 GHz. The use of UWB radio in RFID systems will bring significant benefits including reduced risk of interfering with sensitive medical equipment in hospital/health care applications, very precise positioning for the logistics and retail industries, and the ability to locate people through smoke and obstacles in hazardous search-and-rescue missions. These proposals and ideas are still under development.

3.3.6 RFID and Biometrics

Biometric technology is the use of human bodily characteristics or "physiological autographs" in an attempt to uniquely and absolutely identify individuals. The earlier forms of unique body characteristics were recognized in the science of fingerprints in the 1970s. In the 1980s, the Automated Fingerprint Identification System (AFIS), developed by NEC Technologies, completely changed the role of fingerprints. It combined computer graphics with special software programs and parallel processing to create forensic results. Today, biometric

technologies include retina prints, iris prints, signature and handwriting analysis, palm prints and hand geometry, voiceprints, face recognition, facial thermograms, silhouette identification and gait prints, and even specific task performance and writing styles. Of all the mentioned biometric identification systems, iris prints appear to be the most accurate. The iris patterns of each person's eyes are fixed before birth and remain unchanged throughout one's life, unless trauma interferes. Biometrics is widely used in fields as varied as e-commerce, network access, time and attendance, ATMs, corrections, banking, and medical record access. Due to the apparent ease of use and other factors, biometric technology applications are being used increasingly throughout private businesses and governmental sectors.

Although phenomenal growth in both smart card and biometric technologies has been witnessed, another area of more recent and rapid growth is the merging of these and many other technical elements into the field of RFID. Major initiatives by the United States and other governments aim to fuse RFID and biometric technologies in a new generation of identity cards. Together, RFID and biometric technologies promise to reduce fraud, ease identity checks, and enhance security.

As part of its US-VISIT program, the U.S. government has mandated the adoption of biometrically enabled passports by the 27 nations in its Visa-Waiver Program (VWP), among them Japan, most of the nations of Western Europe, and a handful of others [5]. Soon, all passports produced in the United States will carry biometric information. These passports will be based on guidelines issued by the International Civil Aviation Organization (ICAO), a body run by the United Nations with a mandate for setting international passport standards. The ICAO guidelines, detailed in ICAO Document 9303, call for incorporation of RFID chips—microchips capable of storing data and transmitting them in a wireless manner—into passports. Such chips are present in initial deployments of biometrically enabled U.S. passports and in the biometrically enabled passports of other nations as well. Next generation passports, sometimes called *e-passports*, will be a prominent and widespread form of identification within a couple of years.

The ICAO standard specifies face recognition as the globally interoperable biometric for identity verification in travel documents. Thus, e-passports will contain digitized photographic images of the faces of their bearers. The standard additionally specifies fingerprints and iris data as optional biometrics, and the goal is strong authentication through documents that unequivocally identify their bearers. Data integrity and physical integrity are vital to the security of ID cards as authenticators. For authorities to establish someone's identity with certainty, for example, the passport must carry a photograph of irrefutable pedigree, with a guarantee that no substitution or tampering has taken place. Without this guarantee, passports can be forged, enabling unauthorized persons

to enter a country. Strong authentication requires more than resistance to tampering. Data confidentiality, that is, the secrecy of data stored on ID cards, is also critical. Protecting biometric and biographical data is essential to the value and integrity of an authentication system. In particular, data secrecy affords an important form of protection against forgery and spoofing attacks. Therefore protecting e-passport data against unauthorized access is a crucial part of the security of the entire system. For a full review on the work leading to these decisions, see http://www.icao.int/mrtd.

Confidentiality protection for stored data is important for other reasons as well. Both RFID and biometrics are highly privacy-sensitive technologies. Sensitive data, such as birth date or nationality, are carried on passports. The privacy, physical safety, and psychological comfort of the users of next generation passports and ID cards will depend on the quality of data-protection mechanisms and supporting architecture.

3.3.7 Challenges of RFID Implementation

It has to be emphasized that the implementation of the technology itself has not been a difficult exercise; however, to gain the full benefits of implementing RFID within an enterprise, a more holistic view needs to be taken. A large volume of data is created when implementing RFID, data that has to be turned into information and intelligence. At the present time, enterprise resource planning technology suppliers are developing extensions to their products to work with RFID systems [6]. The following list represents the potential challenges to be considered when implementing an RFID solution. Some of those challenges are discussed in more detail later in this chapter.

Business Issues

Some of the business issues surrounding RFID implementation today include the following:

- No proven return on investment;
- Cost of initial implementation;
- Data sharing between supply chain partners;
- Intellectual property issues;
- Environmental (disposal) issues;
- Consumer privacy objections;
- Lack of organizational expertise;
- Lack of historic data.

Technology Issues

Technology issues include the following:

- Technology standards and interoperability;
- Reliability and maturity of technology;
- Data integration and evolving middleware;
- Environmental issues (heat, moisture);
- Spectrum congestion and frequency availability;
- Security of data on tags and readers;
- Accuracy of tag reading;
- Volume of data produced.

Large Volumes of Data

Readers can scan each RFID tag several times per second, which generates a high volume of raw data. Although the data is redundant and discarded at the reader level, processing large volumes of data can be difficult.

Operational Speed

The RFID system must provide accurate reads at all levels, item, case, and pallet, without requiring any reduction in throughput.

Product Information Maintenance

When the reader processes high volumes of RFID tags, the attributes of each tagged product must be continually retrieved from a central product catalog database, a process that results in challenges for large-scale implementations.

Configuration and Management of Readers and Devices

When a large number of readers and related hardware devices are deployed across multiple facilities, configuration and management can be challenging. The implementation of automated devices for these processes is essential.

Data Integration across Multiple Facilities

In an enterprise with multiple facilities that are geographically distributed, it is increasingly difficult to manage data in real time and instantaneously aggregate it into the central IT facility; such a process can place a significant burden on the network infrastructure.

Data Ownership and Partner Data Integration

When different companies are involved in business processes, such as the retail supply chain, it can create issues pertaining to the ownership and integration of the data, thereby compromising the integrity of the solution architecture.

Data Security and Privacy

Depending on the nature of the business application and the solution scenario, security and privacy challenges could have a significant impact on the architecture.

Cost

At an average cost of around 20 to 30 cents each, RFID tags are still too costly, especially for retail applications and certainly for use on inexpensive and low-margin products, such as a 50-cent candy bar or a $1 bar of soap. This is a key reason why mass-market consumer retail businesses operating on very thin profit margins have been slow to adopt RFID-based smart shelf and smart checkout technology. RFID tag developers are working to lower the cost of tags to 10 cents, or even 5 cents, during the next few years.

Materials

RFID signals are easily blocked. Over short ranges, these signals can be attenuated by certain materials (the most common is packing made from metallic substances). Over longer ranges, the signals, which are much weaker than commercial radio broadcast signals, can be blocked by common objects, including the human body. Researchers are working to solve this problem by using novel designs for tag antennas and more sensitive reader arrays.

Tag Form Factor/Size

Full flexibility in tag sizing is required in order to accommodate the smallest items, as well as cases and pallets, with the same high level of reliability and performance for all tags. In addition, tags must be able to be rigid as well as flexible, for example, to accommodate the curve of a pill bottle. And regardless of how tiny or flexible the tag must be, read-range requirements must still be met.

Tag Proximity and Orientation

RFID tags and readers are orientation dependent. Tags must be positioned properly relative to readers so that the antenna coils can exchange signals. The solution to this problem will come with the development of multiple-reader systems that use an array of readers positioned to cover all the possible orientations for tagged items that might be found, for example, in a display bin in a store. Part of this solution will involve protocols to coordinate the operation of these reader arrays.

Environmental Noise

Because other equipment may be operating in the environment that generates electromagnetic energy (for example, cordless phones, mobile radios, fluorescent lighting, electrical equipment, and other RFID readers), the RFID system selected must be able to reject the interference these products produce in order to ensure predictable and reliable system performance.

Accuracy of Tag Reading

RFID readers often experience false negatives and false positives. A false negative occurs when a valid tag passes within the prescribed range of an RFID reader, but the reader does not read the tag. This can happen for many reasons, for instance, when a case tag is buried deep inside a pallet, when reader signals are blocked or absorbed by substances such as metal or water, or when a case tag is not oriented properly (tag reads are more successful when tags are perpendicular to reader signals). A false positive occurs when a tag accidentally passes within range of an RFID reader, but was not intended to be read. False negatives and false positives often occur with closely packed items, where multiple tags in proximity shadow each other.

Competing Technical Standards

Competing standards prevent the universal adoption of RFID readers and tags. Different manufacturers are developing tag protocols that operate at different frequencies, with a variety of packets. Ideally, a single standard should be adopted to make all tags compatible with all readers. Both the cost and standardization challenges are being addressed by individual companies and by the Auto-ID Center and the International Organization for Standardization (ISO), industry consortia working to set standards for RFID tags.

3.4 Wireless Sensor Networks

3.4.1 About Wireless Sensor Networks

Wireless sensor network (WSN) is the generic name under which a broad range of devices and systems hide. Basically, any collection of devices equipped with a processor and having sensing and communication capabilities that are able to organize them into a network created in an ad hoc manner falls into this category. Research in the field of WSNs has increased tremendously during the past few years, and the initial projects have shown that building cheap smart sensors that can be networked if possible and the addition of wireless communication capabilities to sensors increase their functionality dramatically. Wireless-sensor networks bring monitoring capabilities that will forever change the way in which data is collected from the ambient environment [7].

The field of sensor networks is a relatively new one. Scientists from various communities approached this research area with enthusiasm and brought together knowledge from the various domains of computer science, electrical engineering, telecommunications, radiocommunications, and so forth. The initial directions of research were specific to each of these fields, with everyone trying to adapt their knowledge to make WSNs a reality. Let us take, for example, the traditional monitoring approach of a remote location for a given phenomenon, such as recording geological activity, monitoring the chemical or biological properties of a region, or even monitoring the weather at a certain place. The old approach was to build rather big and robust devices. Besides the sensor pack itself, these devices contained a big power supply and local data-storage capabilities. A team of technicians traveled together to the destination being monitored to place these expensive devices at predefined positions and calibrate all the sensors. Then, they returned after a certain amount of time in order to collect the sensed data. If, by misfortune, some hardware failed, then nothing could be done about it, and the information about the phenomenon itself would be lost.

The new approach is to construct inexpensive, small-sized, energy-efficient sensing devices. Because hundreds, thousands, or even more of these devices will be deployed, their reliability constraints will diminish. No local data storage is needed anymore because they will process locally and then transmit by wireless means the observed characteristic of the phenomenon to one or more access points connected to a computer network. Individual calibration of each sensor node is no longer needed because it can be performed by localized algorithms. The deployment will also be easier, by randomly placing the nodes (e.g., simply throwing them from a plane) onto the monitored region.

Wireless sensor networks are one of the most important tools of the third era of computing. They are the simplest intelligent devices around, having as their main purpose monitoring the environment surrounding us and alerting us about the main events happening. Based on the observation reported by these instruments, humans and machines can make decisions and act on them.

3.4.2 Applications of Wireless Sensor Networks

At this moment a large variety of sensors exists, and sensors have been developed to monitor almost every aspect of the ambient world: lighting conditions, temperature, humidity, pressure, the presence or absence of various chemical or biological products, detection of presence and movement, and so forth. The sensor networks field is still rapidly evolving. Although a large number of sensor network prototypes exist at this moment, the possible application areas are still being explored. The typical application one can think of has as the main goal some sort of monitoring with the most common one being the environmental monitoring. Some of the potential applications are listed next.

Intelligent Warehouses

Each item contained inside the warehouse will have a sensor tag attached that will be monitored by the sensor nodes embedded into the walls and shelves. Based on the read data, knowledge of the spatial positioning of the sensors, and time information, the sensor network will offer information about the traffic of goods inside the building, create automatic inventories, and even perform long-term correlations between the read data. The need for manual product scanning thus disappears.

Environmental Monitoring

This is the widest area of applications envisioned up to now; a particular application in this category is disaster monitoring. The sensor nodes deployed in the affected areas can help humans estimate the effects of the disaster, build maps of the safe areas, and direct needed human actions toward the affected regions. A large number of applications in this category address the monitoring of wildlife. This scenario has an increased complexity. The area of deployment is no longer accessible in an easy manner and no longer safe for the sensor nodes. There is hardly any infrastructure present and a large number of nodes have to be scattered around in a random manner; the network might also contain moving nodes.

Intelligent Highways

Cars have integrated sensors and these sensor nodes will communicate with each other to collect information about the traffic, routes, and special traffic conditions. On one hand, new information will be available to the driver of each car. On the other hand, a global view of the whole picture will also be available. The two main constraints that characterize this scenario are the large number of nodes and their high mobility. The algorithms employed will have to scale well and deal with a network that has a continuously changing topology.

Military Applications

Factors such as rapid deployment, self-organization, and increased fault tolerance make wireless sensor networks a very good candidate for usage in the military field. They are suited to deployment in battlefield scenarios due to the large size of the network and the automatic self-reconfiguration at the moment of the destruction/unavailability of some sensor nodes. Typical applications are for monitoring of friendly forces, equipment, and ammunition; battlefield surveillance; reconnaissance of opposing forces and terrain, targeting, battle damage assessment; and nuclear, biological, and chemical attack detection and reconnaissance.

Health Care Applications

Increasing interest is being shown in caring for the elderly population. Sensor networks can help in several areas of the health-care field. Monitoring can take place both at home and in hospitals. At home, patients can be under permanent monitoring, and the sensor networks will trigger alerts whenever there is a change in the state of the patient. Systems that can detect their movement behavior at home, detect a fall, or remind them to take their prescriptions are being studied. Also, inside hospitals, sensor networks can be used to track the positions of doctors and patients (their status or even errors in medication), expensive hardware, and so forth.

Home Applications

The home is the perfect application domain for the pervasive computing field—we can imagine a future in which all of the electronic appliances form a network and cooperate to fulfill the needs of the inhabitants. They will have to identify each user correctly, remember their preferences and their habits, and, at the same time, monitor the entire house for unexpected events. Sensor networks have another important role here: being the eyes and ears that will trigger actuator systems.

3.4.3 Sensor Network Design Considerations

The basic issue in communication networks is the transmission of messages to achieve a prescribed message throughput and quality of service (QoS). QoS can be specified in terms of message delay, message due dates, bit error rates, packet loss, economic cost of transmission, transmission power, and so forth. Depending on QoS, the installation environment, economic considerations, and the application, one of several basic network topologies may be used. A communication network is composed of nodes, each of which has computing power and can transmit and receive messages over communication links, wireless or wireline. The basic network topologies are shown in Figure 3.9 and include star, bus, ring, tree, mesh, and fully connected approaches. A single network may consist of several interconnected subnets of different topologies. Networks are further classified as LANs (e.g., inside one building) or WANs (e.g., between buildings).

Mesh networks are regularly distributed networks that generally allow transmission only to a node's nearest neighbors. The nodes in these networks are generally identical, so that mesh nets are also referred to as peer-to-peer networks. Mesh networks can be good models for large-scale networks of wireless sensors that are distributed over a geographic region, for example, personnel or vehicle security surveillance systems. Note that the regular structure reflects the communications topology; the actual geographic distribution of the nodes

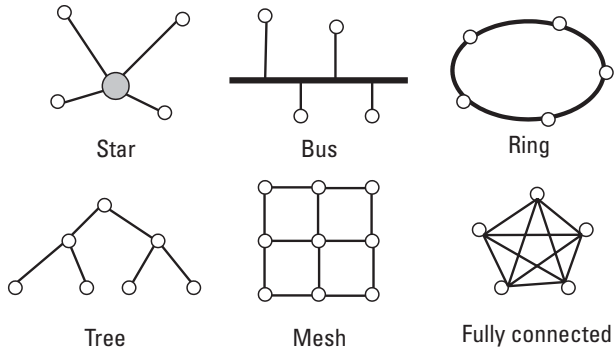


Figure 3.9 Basic network topologies.

need not be a regular mesh [8]. Because there are generally multiple routing paths between nodes, these nets are robust to the failure of individual nodes or links. An advantage of mesh nets is that, although all nodes may be identical and have the same computing and transmission capabilities, certain nodes can be designated as "group leaders" that take on additional functions. If a group leader is disabled, another node can then take over these duties.

The required transmission power in a wireless link increases as the square of the distance between source and destination. Therefore, multiple short-message transmission hops require less power than one long hop. In fact, if the distance between source and destination is R, the power required for single-hop transmission is proportional to R^2. If nodes between source and destination are taken advantage of to transmit n short hops instead, the power required by each node is proportional to R^2/n^2. This is a strong argument in favor of distributed networks with multiple nodes, that is, mesh networks.

If we take a look at the number of sensors deployed with respect to the area covered, we can make the following categorizations:

- *Coarse-grained sensor networks.* In this category usually fall the sensor networks made up of devices, each covering a large area. These devices are usually large and expensive, because they are equipped with high-quality sensors. The network topology is usually a star topology. The sensor nodes themselves are fixed.
- *Fine-grained sensor networks.* This category comprises the networks made up of large number of cheap devices, equipped with low-quality sensors having small amounts of resources available. The network topology is usually a multihop network. The large number of sensors and the dense deployment compensate for the low quality of the sensors, the network as a whole producing high-quality results.

Desirable functions for sensor nodes include ease of installation, self-identification, self-diagnosis, reliability, time awareness for coordination with other nodes, some software functions and DSP, and standard control protocols and network interfaces. There are many sensor manufacturers and many networks on the market today. It is too costly for manufacturers to make special transducers for every network on the market, so different components made by different manufacturers should be compatible. Therefore, in 1993 the IEEE and NIST began work on a standard that resulted in IEEE 1451, "Standard for Smart Sensor Networks." The objective of this standard is to make it easier for different manufacturers to develop smart sensors and to interface those devices to networks. Under the original concept of IEEE 1451, a sensor is divided into two parts. The first, called a smart transducer interface module (STIM), contains the sensing element (strain gauge, thermocouple, vibration sensor, and so on), the appropriate signal-conditioning circuits and A/D converter, plus a transducer electronic data sheet (TEDS), a memory chip that identifies the type of sensor, its make and model, its calibration information, its scale factor, and more. IEEE 1451.2 defines the basic STIM, and IEEE 1451.1 defines the Network Capable Application Protocol (NCAP) (Figure 3.10).

Although 1451.1 and 1451.2 worked well, they didn't cover enough configurations, and another substandard was started, IEEE 1451.3, "Standard for a Smart Transducer Interface for Sensors and Actuators, Digital Communication

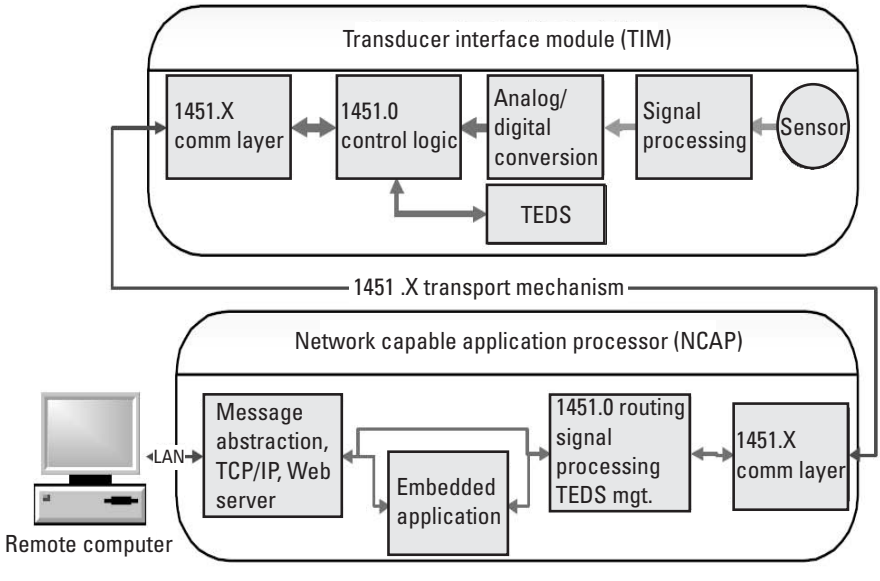


Figure 3.10 IEEE 1451 smart transducer concept.

and Transducer Electronic Data Sheet (TEDS) Formats for Distributed Multidrop Systems" (often called Dot3, and approved in October 2003), which allows multiple transducer modules (called transducer bus interface modules) of varying complexity and data rates to be multidropped to one NCAP via a local transducer bus.

One limitation of IEEE 1451 was lack of backward compatibility; that is, it did not address the large number of legacy sensor devices, with analog outputs and no digital communications, that are already in use. IEEE 1451.4 (accepted as a standard in August 2004), "Standard for a Smart Transducer Interface for Sensors and Actuators-Mixed-Mode Communication Protocols and Transducer Electronic Data Sheet (TEDS) Formats" (also called Dot4), was proposed as a way to do this. Unlike Dot2 and Dot3 sensors, a Dot4-compliant sensor has an analog output and no A/D converter, but it does contain a TEDS.

The first recently approved standard, IEEE 1451.0, "Standard for a Smart Transducer Interface for Sensors and Actuators—Common Functions, Communication Protocols and Transducer Electronic Data Sheet (TEDS) Formats," creates a common set of functions, protocols, and formats to facilitate interoperability among other standards in the IEEE 1451 series. It will also simplify the creation of future standards in the family for different physical layers.

The second approved standard, IEEE 1451.5, "Standard for a Smart Transducer Interface for Sensors and Actuators—Wireless Communication Protocols and Transducer Electronic Data Sheet (TEDS) Formats," establishes uniform wireless communication methods and data formats for transducers.

3.4.4 The Future of RFID Sensing

Industry is fast moving toward employing networked, digital, and wireless communications technologies for sensors. Using wireless connectivity for sensor networks increases the flexibility in deployment and reconfiguration and thus reduces the overall infrastructure cost. These advantages will enable sensor networks to monitor complex environments for applications ranging from industrial automation to battlefield surveillance to environmental monitoring to telemetry of first responder's health condition. RFID devices are going to play a key role in automated universal identification systems for accessing, securing, and tracking assets, personnel, equipment, and products throughout the supply chain. Combining RFID devices and sensors could expand the overall functionality and capability of the proposed applications.

In today's complex supply chain, even if technology standards are high, visibility still performs at very low efficiency levels. Asset tracking and management, anticounterfeit, and in-transit visibility are a few of the major applications of RFID technology to the real world. Several suppliers already have or are

preparing to meet their retailers' mandates, but what if you could not only find where your products are located, but also receive information about their conditions and status? Location is just one part of the equation; according to the U.S. Department of Agriculture, 10% of all perishable products spoil before they reach the consumer. The solution lies in enabling more parts of this equation, by adding several sensors to the RFID solution. So far RFID has enabled a number of end users to gain higher visibility of their supply chain, increase inventory efficiency, reduce out-of-stock situations, and gain higher anticounterfeit protection. RFID sensors will create and provide a new layer of protection and advance supply-chain visibility.

Efficient sensor networks require the sensor nodes to be cheap, to consume little energy, to be multifunctional, to be small, and to have the ability to communicate both among themselves and with other networks. Compared to mobile ad hoc networks, wireless sensor networks differ in various ways, including the larger number of nodes, the dense deployment, the attribution of fault proneness, the frequent topology changes, the main use of broadcast communication instead of point-to-point communication, and the limitations in power, storage, and processing units. Also, they often do not possess a global identity. Nevertheless, the demarcation of wireless sensor networks and mobile ad hoc networks is often vague in literature. In contrast to other areas, the field of wireless sensor networks is a new discipline.

The new IEEE P1451.7 proposed "Standard for a Smart Transducer Interface for Sensors and Actuators—Transducers to RFID System Communication Protocols and Transducer Electronic Data Sheet Formats" will address the integration of sensors in RFID infrastructures.

3.5 RFID Applications

RFID systems have been deployed in limited numbers for years; two of the most predominant have been in the form of toll-road collection transponders and security badges. Toll-road authorities around the country have equipped drivers with transponders that are connected to their credit cards. This allows them to pay their tolls at 40 mph rather than stopping to throw quarters into a basket and slowing the flow of traffic.

Security badges have been equipped with RFID chips to allow centralized control of access to facilities and specific rooms within buildings. These can also be used to track the locations of people in a facility by identifying the door they last passed through. Today, RFID has the potential for applications in virtually any area of industry, commerce, and services where items are handled and associated data collected and processed. These include:

- Supply chain logistics;
- Product authentication;
- Tracking and traceability;
- Security, ticketing, and access control;
- Lifetime item identification;
- Transient carrier labeling;
- Animal and specimen identification;
- Airline baggage handling.

The following sections will illustrate just a small number of examples—the possibilities are truly endless and limited only by our imagination.

3.5.1 Supply Chain Logistics

Global supply-chain logistics are expected to be the largest and fastest growing application for RFID. This will most likely be done through smart labeling of cases, cartons, and pallets. The key benefit is the ability to read the entire contents of mixed pallets all at once during material handling operations, such as truck loading or unloading. Managing pallets, totes, and other returnable transit containers with RFID represents one of the most dramatic cost-saving opportunities this technology can provide. Many returnable containers are never brought back from customer sites after shipment, forcing companies to carry excess inventory to ensure adequate supplies of shipping materials where they are needed. Identifying returnable containers with smart labels or fixed tags enables companies to augment their legacy barcode shipping applications by automatically recording materials shipped to customers. Companies can then find their own pallets in shipping yards or docks stacked with thousands of items belonging to dozens of companies. Incoming pallets or cartons with smart labels can be automatically routed for cross-docking or delivery directly to the manufacturing line. Fast-reading RFIDs enable instant identification of the shipping container, plus all of the individual items inside. For shipping, RFID readers can help packers quickly locate and aggregate all the items needed to complete an order.

The same principle is applied to improve warehouse picking. Workers scan shelves and bins with an RFID reader that automatically detects the storage location of the sought-after items. The system can also detect items that are stored in the wrong location and alert operators to the problem. Using RFID for these applications enables items to self-report their locations, rather than requiring human intervention to find them, thus reducing errors, saving labor, and lowering costs.

Wal-Mart Stores Inc., one of the originators of the RFID movement, continues to expand its RFID capability to additional facilities, enabling an additional 400 Wal-Mart stores by the end of the 2007 fiscal year. Wal-Mart claims that the current benefits include a 30% reduction of out-of-stocks, reduction of excess inventory in the supply chain, and sustainability impacts. Aside from the initial RFID implementation to track inventory, in the near future, customers will be able to enjoy advantages such as automatic warranty activation on electronics, freshness assurance on foods, thanks to cold-chain monitoring, and enhanced product safety as a result of faster, more accurate recalls and better freshness monitoring.

3.5.2 Product Authentication

The role of product authentication is to answer whether a given product is genuine or counterfeit (e.g., a product that infringes a trademark). An explicit way to authenticate products is needed in supply-chain applications because counterfeits can be very similar or even identical to authentic products. The starting point of automated nondestructive product authentication is to insert a special label or security feature into products, like a hologram or a watermark, and to authenticate this label. Product authentication can take place at the single-item level or in aggregated levels. Generally, multiple similar units are authenticated simultaneously, for example, when a shipment arrives at a retail store. The desired level of security, which can be defined as the effort an illicit actor has to undertake to break or bypass the security mechanism, has a major impact on the cost of a product authentication system. While minimizing the cost, the level of security should be high enough to protect the item over its entire life span. Because different products have varying security requirements, different levels of security and, thus, different solutions are needed.

The level of security of a product authentication system is defined by the level of security of a single security feature and by the granularity of the security features. By *granularity,* we mean how many products use an identical security feature; for example, applying weak but unique security features to all products can be more secure than using a strong but identical feature on the same products. One conceptual problem of automated product authentication is that it is only the security feature that is authenticated and not the product itself; therefore, a difference between label and product authentication should be made. The general requirements of a product authentication system in supply-chain applications are as follows:

- The system needs to be used by multiple parties from multiple locations.
- Authentication of products that are unknown to the system should be supported.

- The cost and effort to perform a check need to be low.
- The optimal solution should allow the customers to authenticate products as well.
- The product authentication system needs to have an appropriate level of security.

Among the requirements just listed, the level of security demands most attention in the system design. The level of security can be considered to be the resistance against attacks that are conducted against the authentication system. In supply-chain applications, product authentication is typically performed under the supervision of authorized personnel, thus restricting the possible attacks of counterfeit players. The general attack scenarios of illicit actors against product authentication systems can be divided into the following four categories:

- Omission of security features that are applied to the genuine objects refers to the counterfeiters' not taking any explicit actions to fool the authentication. These products form a considerable part of the counterfeit trade, due, for example, to consumer demand for counterfeits.
- The use of misleading security features means that the fake products are equipped with security features that make the products avoid closer inspection. Interviews with brand owners and customs reveal that this scenario, together with the aforementioned one, is dominant, especially for goods that are mass produced or where the consumers do not regularly check for the object's authenticity.
- The removal and reapplication of authentic security features remains a threat in all automated product authentication systems, if not explicitly addressed by binding the product and the label. However, because acquiring and reapplying authentic labels is costly, this attack does not threaten authentication systems on a large scale.
- The cloning and imitation of security features is the most obvious attack that a product authentication system has to resist. Because the underlying problem of counterfeits is that the products themselves can be cloned, the first line of defense is to integrate security features that are hard to replicate into products.

RFID has considerable potential in product authentication. The benefits of RFID, compared to those of old authentication technologies include nonline-of-sight reading, item-level identification, the nonstatic nature of security features, and cryptographic resistance against cloning. RFID systems, in general, comprise transponders, readers or interrogators, and an online database,

sometimes referred to as the back-end server. In many applications, RFID transponders are already being used for authentication, for example, in access control. Although RFID product authentication is very close to RFID access control when it comes to the authentication protocols, product authentication needs specific solutions because of the specific application requirements.

Resisting cloning and forgery is the most important security property of authentication tags. The simplest cloning attack against an RFID tag only requires reading the tag's serial number and programming the same number into an empty tag. There are two essential obstacles against this kind of replication. First, even the low-cost transponders (e.g., EPC Class 1, Generation 2) have a unique factory-programmed chip serial number (or transponder ID, TID) that is similar to the unique MAC address of PC network cards. To clone a transponder's TID would therefore also require access to hardware manufacturing.

The second obstacle against cloning is to place read-protected secrets on tags and to check if the tag knows these secrets, for example, by cryptographic challenge–response protocols. Even though this can provide significant improvements to a tag's ability to resist cloning, many ways remain in which to conduct a cloning attack against a single tag. These attacks include side-channel attacks, reverse-engineering and cryptanalysis, brute-force attacks, physical attacks, and different active attacks against the tag. In addition, shared secrets based on product authentication approaches are always vulnerable to data theft, where the secret PIN codes or encryption schemes of valid products are stolen or sold out by insiders, which would enable criminals to create phony tags. This scenario is especially interesting for adversaries because it would allow them to clone a large number of tags.

Other RFID security issues that have to be considered in product authentication comprise resistance against denial-of-service (DoS) attacks. In general, DoS attacks cause a loss of service to users. Even though it cannot be used to fool the product authentication, it can pose a threat to the overall process. In RFIDs, a DoS attack can be conducted, for example, by jamming the readers with hidden blocker tags or by desynchronizing the tag from its database entry.

We assume that product authentication is normally performed under the surveillance of authorized personnel or by the customer, which narrows down the possible attack scenarios. Therefore, active attacks, in which the adversary would need to participate in the authentication session and use special devices in the proximity of the reader (e.g., replay, relay, and man-in-the-middle attacks), are not considered as realistic threats against RFID product authentication.

3.5.3 Agriculture and Animals

RFID is already used to manage commercial livestock by farmers, improving farming efficiency and allowing livestock to be traced back to their origins (in

case of disease). The technology can also be used to effectively develop a track-and-trace system for meat, to keep track of pets and their vaccination records, for easy retrieval of pets by owners, to identify animals, and to reduce the potential spread of diseases when crossing borders (Figure 3.11). The technology is also being used for fish and wild game to track migrations, breeding patterns, population, and so forth, and to prevent poaching and illegal exporting of endangered species and ivory tusks.

Researchers have noticed recent cases of cannibalism in polar bears in the Arctic—something that has previously gone undocumented—and further investigation will be required. A study on polar bears in Alaska, by members of the U.S. Geological Survey (USGS) shows that some types of RFID tags can be read from as far away as 1,500 feet, even while the reader is in motion (in this case, a helicopter). In terms of savings, the RFID ear tags cost $35 and the battery lasts 5 years. In comparison, the older satellite-radio collars cost $4,000, with batteries lasting only 2 years. Obviously, the radio collars can be tracked at a greater distance.

3.5.4 Intelligent Transportation Systems

Traffic congestion in the largest cities of the world is a growing problem that has to be taken seriously, not only by governments, but also by the private sector. Different alternatives are being analyzed to solve this dilemma. The concept of an intelligent vehicle/highway system (IVHS) has been proposed as the best solution [9]. The IVHS has been called by different names, depending on the developing area and the application purposes: intelligent transportation system (ITS), intelligent cars and automated highways systems (AHS), automated vehicle/highway system (AVHS), and smart cars/smart highways are just some of the different names basically describing the same concept.

IVHS is an intelligent transportation system in which vehicles and highways exchange information through a two-way communication system (Figure 3.12). The automated highways will have a set of lanes in which vehicles with

Figure 3.11 Glass transponder for the identification of animals (and humans).

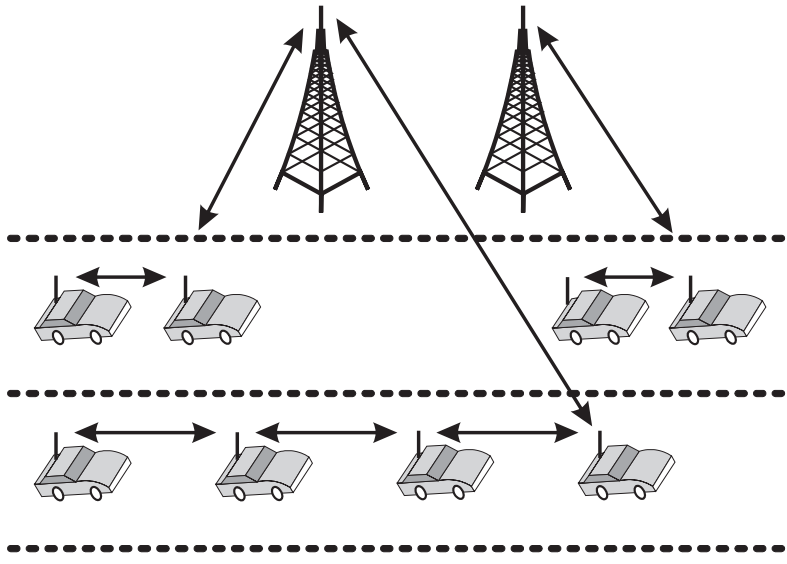


Figure 3.12 Intelligent transportation system.

specialized sensors and wireless communications systems could travel under computer control at closely spaced intervals. This type of arrangement is called a *platoon*. The vehicles could continuously exchange information with other vehicles and traffic control centers about speed, acceleration, braking, obstacles, road conditions, and other vehicle data. Sensor data can be processed and sent back to each vehicle, guaranteeing a continuous exchange of information.

The highway system will know the destinations and planned routes of individual vehicles. In that way the system can coordinate traffic flow more efficiently, reduce speed fluctuations, monitor unsafe vehicle operation and traffic shock waves, maximize highway capacity, and minimize avoidable traffic congestion. In addition, the system will respond rapidly to changing highway conditions. The vehicles might use several types of devices to sense its environment, such as magnetometers, visual sensors, infrared sensors, laser sensors, or accelerometers. Each vehicle has to have a powerful computer to process sensory data and the information that comes from the traffic control centers.

IVHS America has become the coordinating and planning entity in which the individual activities of state and local authorities, companies, and universities have a central orientation for constructing a national IVHS program. VERTIS (Vehicle, Road, and Traffic Intelligence Society) in Japan, ERTICO (European Road Transport Implementation Coordination Organization) in Europe, and IVHS America in the United States perform similar activities as major coordinators of the individual IVHS programs.

Some proposals have suggested using RFID technology for positioning [10]. This technique, however, would not replace GPS; rather, it is a complementary technique. RFID tags need to be installed on roads in a manner that maximizes the coverage and the accuracy of positioning. On installation, necessary information, such as the coordinates of the location where the tag is installed, needs to be written on each tag. The accuracy of this position information is very critical for this technique to be successful. The position information can be acquired by using DGPS or some other methods, which would take a long time to compute the location. Contrary to GPS in navigation systems where real-time positioning is necessary, the time for getting the accurate information would be tolerated since this computation would take place once. Vehicles, then, need to be equipped with an RFID reader that can communicate with the tags on a road. No matter how accurate the RFID positioning is, it only gives the position where the tags are. Therefore, the vehicles also need to be equipped with a GPS receiver and inertial sensors such as a gyroscope for positioning when there are no tags around. While driving, the vehicles constantly monitor the presence of a tag. On detection, the reader retrieves the information from the tag including a lane marker. The deployment should be done step by step; places such as tunnels where GPS signals are not available should be the first, intersections next, urban areas, and then nationwide. Due to the nationwide scale of the project, governmental participation would be necessary.

3.5.5 Document Management

Books and other materials are identified with smart labels that carry a unique, tamperproof ID code. Librarians at the circulation desk and patrons read the tags with RFID readers to check items in and out. The process is faster and more accurate than with traditional optical barcode labels. Some economic facts that help justify installing this system are as follows:

- A lost book typically costs a library around $45.
- An average library can have as many as 22 million items circulating each year.
- With RFID smart labels on items, check-in and check-out saves at least 1.5 minutes per transaction.

Besides the unique identification number, these labels can be programmed with additional information, such as type of media and storage location. In the retail RFID space, the EPCglobal suite of RFID specifications mandates that tags support an irrevocable kill command. In the library setting, however, tags must be reused to check in loaned items. Irrevocably killing a tag is not an

option. The tag has to be rewriteable, so libraries do not have to replace a book's digital identification tag when updating a book's status or flagging a book for reservation. Libraries are finding new ways to take advantage of tagged items, such as gathering statistics on what items are most often used. The main goal of libraries is to improve service to their patrons, particularly by having circulating items available when they are needed. RFID tracking greatly improves inventory management and optimizes resources.

RFID can also be used to improve the management of important individual document files in places such as libraries, insurance companies, and law offices, where the loss of such files can cause severe problems. RFID improves the tracking of documents so that files can be quickly located and document workflow more easily tracked. Each file is tagged with a smart label that contains a unique ID and human-readable information. The file description is entered into a database along with its tracking number. The file can be assigned certain parameters, such as expiration date, permitted movement, and persons authorized to see it. Over time, the database can build up an audit trail of the handling and workflow history of each document file.

3.5.6 Pharmaceutical and Health Care Industry

A number of pilot programs are already under way in the pharmaceutical supply chain and health-care markets for item-level management. Although much remains to be learned about the efficiencies and safeguards that can result from the use of RFID solutions in these markets, companies implementing RFID pilots are experiencing process improvements and safety benefits today. Suppliers to the medical industry, from garment to surgical instrument providers, as well as health-care institutions managing blood and tissue sample processing, are investigating the viability and reliability of HF technology solutions, and they are seeing significant returns in the field [11].

State and federal laws have been enacted to address the increasing threat to public health posed by counterfeit drugs. At the federal level, the U.S. FDA enacted the Prescription Drug Marketing Act (PDMA), which went into effect December 2006, with a mandated requirement that every drug must have a full pedigree (information required to ensure the security and authenticity of a drug as it travels through each step in the pharmaceutical supply chain). In addition, many states are enacting their own pedigree laws. Florida's pedigree law, which enables companies to select either paper or electronic record keeping, requires full documentation of the drug from the manufacturer to the store, requiring each shipment to be accompanied by the amount of the drug, dosage form and strength, lot numbers, name and address of each owner with owners' signatures, and complete shipping information. California's pedigree law, which specifies

the requirement for an electronic pedigree, went into effect on January 1, 2007. Many other states are considering similar laws.

The need for and value of RFID technology in the pharmaceutical industry are well recognized and documented. With the need to meet U.S. federal and state regulatory compliance in the very near future, many companies in the pharmaceutical industry are aggressively seeking to implement RFID now (Figure 3.13). Leading companies have already defined critical strategic and business needs, as well as the system specifications that will meet those needs. These requirements are some of the most demanding and stringent of any RFID applications to date. For pharmaceutical manufacturers and distributors, RFID will provide a solution for three critical issues: improvements in counterfeit protection; regulatory compliance by providing the ability to create a complete electronic pedigree, automatically, without human intervention; and rapid and cost-effective recalls by providing the ability to instantly and automatically identify the location of a product that must be recalled.

Locating Tissue Samples

Tissue-sample processing labs are using miniature HF tags to create efficiencies in locating single test tube samples among the hundreds in the lab at any given time. The RFID tag contains a unique serial number as well as memory that can be read, modified, and protected. The serial number is then linked to a database containing critical information on each tissue sample, including patient data and tissue treatments. Using a fixed desktop or lightweight handheld reader at a distance of a few inches, researchers and lab technicians searching for a specific sample on a tray of 100 tubes can quickly and easily read all of the tags in less than a few seconds. What was previously a painstaking and time-consuming task of locating and identifying samples can now be completed quickly with the simple pass of an RFID reader over the existing inventory.

Matching Blood Samples to Patients

Trials have shown significant reductions in administration time, both during sampling and laboratory processing. In the trial, the doctor or nurse taking the

Figure 3.13 RFID tag on a pill bottle.

blood sample enters the patient information into a handheld RFID device at the start of the blood-sampling process. This data is stored on an RFID label (HF type) on the patient's blood-sample tube and can be read by fixed readers and automatically transferred to the facility's database, enabling a fully automated process and replacing an entirely manual one. In addition, because the patient data is entered in electronic format at the beginning of the process, the integration of results into the patients' records is quick and simple.

Patient Identification and Care

The U.S. Navy is using HF technology to more efficiently track the status and location of hundreds of wounded soldiers and airmen, prisoners of war, refugees, and others arriving for treatment at Fleet Hospital Three, a 9-acre, 116-bed facility in Southern Iraq. The RFID-based system allows medical professionals to use RFID-enabled wristbands to identify patients and to update their status, location, and medical information in the system's electronic whiteboard automatically. The Navy implemented a new system to replace a labor-intensive, entirely manual system consisting of pen and paper, cardboard tags, and a centrally located whiteboard to show patient movement throughout the hospital. With the new electronic system, each patient receives an HF-enabled wristband, on which basic identification information is stored. Medical professionals use a handheld RFID device to read the unique identification number and to add or change data to create a digital treatment record that travels with the patient as he or she is moved throughout the facility. Using a wireless LAN, patient information is transferred to an electronic patient management system, further eliminating manual reentry of data at a central computer terminal.

3.5.7 Indoor Localization for First Responders

Most of the research and development for indoor localization includes that of a wireless network that integrates communications, precise tracking, and data telemetry, based on UWB technology, for use in hospital and manufacturing environments. In contrast, the system used by first responders is intended for an environment that is potentially much less friendly to RF propagation—the in-building environment may contain smoke, dust, or flames. The system is intended to leverage advances in ubiquitous RFID tag technology in combination with recent advances in miniaturized inertial sensors [12]. This research is a joint effort by components of three NIST laboratories: the Wireless Communication Technologies Group of the Information Technology Laboratory (ITL), the Fire Fighting Technology Group of the Building and Fire Research Laboratory (BFRL) in Gaithersburg, Maryland, and the Radio-Frequency Fields

Group of the Electronics and Electrical Engineering Laboratory in Boulder, Colorado.

The most widely used navigation system today is the global positioning system, which enables position determination through the measurement of time delays of signals from multiple satellites in known (moving) positions; the time-delay measurements are based on cross-correlating received satellite signals with local replicas to identify the signals' digital code position in time relative to the common reference. The difficulty in using GPS indoors and in urban outdoor areas is that the line-of-sight to the GPS satellites is obscured or severely attenuated. Without four good satellite signals, the GPS position solution is inaccurate. In addition, with weak signals, the GPS receiver continually loses its lock and must spend an inordinate amount of time attempting to reacquire the signals. Research is aimed at finding a way to implement a low-cost, reliable means for tracking firefighters and other first responders inside buildings, where navigation using GPS is not reliable or where the GPS signal may have been disabled temporarily to prevent exploitation by terrorists. Even if the GPS signals are not blocked or obscured for tactical advantage, the reception of GPS signals inside most buildings is not reliable.

Prior to the establishment of GPS, many techniques and devices for navigation were used. Today's navigation devices implement some very old navigation techniques, such as dead reckoning and waypoint navigation. Dead reckoning (DR) is the process of estimating position by advancing a known position using course, speed, time, and distance to be traveled; in other words, figuring out where one will be at a certain time if one holds the planned speed, time, and course. The usefulness of the technique depends on how accurately speed and course can be maintained in the air and on the sea, and the uncertainty of the DR position grows with time, so that it is necessary to check the position regularly with a fix of some kind, perhaps an RFID tag. It has been noted that 0.6 to 2 GHz is the best frequency band for propagation in buildings.

3.5.8 Passive Keyless Entry

When in 1993, a major German insurance company increased the pressure to either protect vehicles against massively increasing theft or else renounce full insurance coverage, nobody believed that RFID as a then-emerging technology would see one of its first major successes, still unbeaten in numbers by any other application today. Objection was raised as to the reliability of potential systems and their suitability in an automotive environment; and, of course, the lack of standards in what then seemed to be a mystic collection of proprietary systems was seen to be a major obstacle. An extremely fast but substantial development effort was undertaken by a few powerful automotive industry suppliers, resulting in a miniaturized ignition key system and its car lock transceiver

counterpart, which were fitted to new cars as early as 1994. This first generation system, still fitted to certain models of most major car makers, consisted of a 64-bit read-only, rod- or brick-type transponder, embedded in the ignition key of the vehicle and a transceiver antenna with its electronics package on a printed circuit board (PCB), integrated around and behind the lock. When the presented key does not match one of the prestored codes, the functionality of the system consists of safely cutting the power supply to the starter, the fuel pump, the system ignition, and other system elements that are required for the vehicle's operation during driving. Specifications for the miniaturized transponders were very stringent, requiring a maximum read distance in a metal-loaded environment over industry-practice thermal ranges with only a few-parts-per-million failure rate allowed. Three mainstream systems were adopted by the majority of the automotive industry and have since been fitted to millions of new cars with great success. Immobilizers today in Europe are part of the normal equipment of cars, and although no standard rules have yet been devised for the application nor the products, several systems have established a de facto standard and wide public acceptance.

Hands-free passive keyless entry (PKE) applications require bidirectional communication between the base station and transponder units. The base unit inside the vehicle transmits a low-frequency (LF) command that searches for a transponder in the field. Once located, the transponder in the vehicle owner's possession then automatically responds to the base unit. The base unit then unlocks the car doors if a valid authentication response is received.

3.5.9 Military Applications

The U.S. Department of Defense (DoD) became involved in RFID during the 1990s due to the identification of supply-chain challenges. During Operation Desert Storm in 1991, logistics and materiel distribution were major problems. The Defense Logistics Agency (DLA) became known for having mountains of unopened shipping containers in the middle of the Saudi Arabian desert. The lack of supply-chain visibility required 25,000 of the 40,000 containers to be opened in order to identify their contents. A Defense Research Projects Agency (DARPA) grant was awarded to identify whether RFID could help prevent similar supply chain problems in the future. This resulted in several initiatives during the next few years.

Evaluations of the DLA's effectiveness during Operation Desert Storm focused on the high cost of the DLA's supply chain. In 1995, the Joint Total Asset Visibility Office was formed with a charter to provide asset visibility in storage, in process, and in transit to optimize the DoD's operational capability. This had several results. First, it organized all RFID supply chain initiatives under one office, instead of being managed by individual armed forces or

distribution depots. Second, it provided a source for funding future RFID initiatives. By 2004, the DLA had spent more than $100 million on RFID initiatives; this level of funding would not have been available under the previous organizational structure. Finally, the implementation plan tied RFID usage to the overall strategic goals mandated by the department's charter. This forced a necessary pragmatism around RFID's relative advantages compared to other automatic identification technologies. These realistic expectations were a key contributor to the success of the DLA's RFID initiatives. By 2004, the DoD had joined EPCglobal.

When it comes to RFID, the DoD and the armed services have learned a thing or two from private industry. The DoD, for instance, is adding passive RFID tags to cartons and pallets, and rather than reinvent the wheel, the DoD is working with established industry technologies. Similarly, the armed services are implementing active RFID solutions to track mobile assets, like containers in the field, and work in process. The Defense Appropriations Act for fiscal year 2007 had a total of $17 million added for projects either directly or indirectly related to RFID [13].

According to the DoD's policy, which was finalized and released in July 2004, by January 1, 2005, suppliers had to put passive RFID tags on all individual cases, all cases packaged within a pallet, and all pallets of packaged troop rations, clothing, individual equipment and tools, personal items, and weapons systems repair parts and components shipped to the two DLA distribution centers. Beginning January 1, 2006, suppliers were required to tag cases and pallets of subsistence and comfort items, packaged petroleum, lubricants, oils, preservatives, chemicals, construction and barrier material, ammunition of all types, and pharmaceutical and medical material shipped to 32 depots throughout the United States and the two DLA distribution centers. Beginning January 1, 2007, all pallets of all commodities shipped to all DoD locations had to be tagged.

The Tobyhanna Army Depot, for instance, is using a real-time locating system (RTLS) to streamline the repair and overhaul of defense electronic systems. The Army maintains, repairs, and overhauls command, control, communications, intelligence, and reconnaissance systems at the 1.9-million-square-foot refurbishment center in Pennsylvania, including the army's radar system, which detects and tracks enemy mortar and artillery shells in Afghanistan and Iraq. These systems are shipped from the field to Tobyhanna. There, each system is disassembled, overhauled, and tested before being shipped out into the field again. The RTLS allows Tobyhanna to track hundreds of components during the process that might otherwise get misplaced in the massive facility. Being able to find the right part when it is needed will result in an estimated annual savings of nearly $8 million and the elimination of 837 repair cycle-days per year. The Army expects to obtain a complete return on its initial investment in

just more than 11 months. More importantly, the system will expedite the refurbishment process, enabling the Army to return systems to the field up to 35 days sooner than in the past [14].

The U.S. Army is developing an RFID system to track weapons usage. In an effort to make sure ground vehicles' weapons are properly maintained, Benét Laboratories is working on a system that uses sensors and passive RFID tags to record the number of rounds fired. The first vehicle to demonstrate the sensor system will be the M1 Abrams Main Battle Tank. The sensor systems will include microelectromechanical sensors (MEMS) integrated with RFID tags to help the vehicle's operator track how many times a weapon has been fired. This allows the gunner to determine whether it can be depended on to function properly. To date, vehicle and weapon operators have been required to log manually, on paper, how many times a weapon has been fired and the types of munitions used, and then bring that information back to the vehicle maintenance depot.

The first phase of the deployment, which will measure applied force, will focus only on using the MEMS sensors to count the rounds fired. However, the MEMS sensors will eventually be used to measure a fired round's physical effects on the gun barrel as well, including the intensity of the applied force, heat, and vibration. This will assist not only in counting the number of rounds fired, but also in recording the physical effects of the type of munitions fired in each instance, resulting in an accurate indication of the health and maintenance requirements of the barrel.

An operator can access the weapon-firing count on a tablet PC inside the vehicle. The tablet includes an RFID reader, and its data can be downloaded by Army personnel to keep a record of the weapon's firing history. When a vehicle operator prepares the weapon for firing, the reader automatically sends an RF signal energizing the passive RFID tag, which transmits its unique RFID number and a count of each time the weapon has been fired. Several frequencies may be tested with this system; the air-interface protocol has not yet been determined.

The tablet PC will include integrated data storage, communications, and display technologies. The MEMS will incorporate an RFID tag and a low-power microprocessor with limited memory, which would continue to record rounds fired even if the rest of the system were to fail. The prototype system was scheduled for demonstration by the end of 2007, with initial deployments planned for 2008 [15].

3.5.10 Other RFID Applications

Gillette (which announced in early 2003 that it would purchase 500 million RFID tags) has worked with retailers to test "smart shelves," as an adjunct to item-level tagging, for inventory control. With a reader on each shelf and a tag

on each package of razor blades, the data proprietor would always know how many packages are on the shelves, without having to count them.

Michelin has begun fleet-testing the use of RFID technology for passenger and light truck tires (Figure 3.14). Each tire's unique identification number will be associated in an external database with the vehicle identification number (VIN) of the car on which it is mounted and with information describing when and where the tire was made, its maximum inflation pressure, its size, and so forth. The transponder consists of a UHF/SHF RFID chip, circuit board, and two antennas. The antenna leads must be tuned to resonate at 868 and 960 MHz when cured into the tire. The current design has a nominal recommended tuned length of around 61 mm end-to-end (effective August 2005 and beyond). Exact tuned length will always depend on tire type.

Casinos are putting RFID tags in chips to block counterfeiting, identify stolen chips, and track gamblers' play. An Italian manufacturer has introduced a washing machine equipped to read RFID washing instruction tags in clothing.

A German supermarket, for a brief time, inserted RFID tags in supermarket loyalty cards, which gave the store the capability, while someone carrying the loyalty card was in the store, to pull up his entire buying history without his being aware that the query was taking place and without any other basis for the store's knowing who he was.

It is possible to imagine a whole lot of uses, indeed, for a technology in which objects can be uniquely identified without direct contact. This is the ideal technology if you want the milk in your refrigerator to notify you (or your supermarket) if you have failed to drink it by its pull date. Indeed, you could tie a slightly more elaborate tag to a nanosensor that checks for spoilage directly.

Students in an Osaka, Japan, elementary school will be getting RFID chips in their schoolbags, name tags, or clothing, to be read by readers installed in school entrances and exits. More than 50 million pets have RFID tags.

Federal Express (FedEx), headquartered in Memphis, Tennessee, is one of the world's largest express parcel delivery companies. FedEx delivers approximately 3.2 million parcels daily and operates a fleet of more than 42,500 vehicles worldwide. FedEx is constantly looking at ways to streamline their delivery processes [16]. The company's couriers drive millions of miles daily in the United States alone. Each time a courier makes a delivery, the driver must spend time searching for keys or using them to lock/unlock multiple doors to their

Figure 3.14 RFID tire transponder.

vehicle; if a courier misplaces his keys, he must wait for someone from a FedEx station to bring out spare keys, and the vehicle must be rekeyed at a cost of more than $200 per incident.

The FedEx system uses RFID readers mounted at each of the four doors to the delivery vehicle and a reader mounted on the right side of the steering column near the ignition switch. The company's couriers use an automatic keyless entry and ignition system that has RFID transponders embedded within a Velcro wristband. When the courier places his transponder wristband within 6 inches of the readers, the transponder's code is compared to those in the system's memory. If it is a match, the door unlocks for 5 seconds. The courier simply pulls on the door handle to enter the vehicle, while the three remaining doors stay securely locked to prevent unauthorized entry. To start the vehicle, the courier pushes a button on the right side of the steering column. The courier pushes another button near the start button to turn off the vehicle.

The University of Washington's Brockman Memorial Tree Tour uses RFID and PDA technologies [17]. The combined use of these new technologies increases the ease of locating and positively identifying each tree along the lengthy, spacious tour around the campus. Each tree is individually profiled including details such as origin, date planted, type of flowers, and other pertinent information. A map, located at the rear, visually locates each tree by its unique number. Originally, a nameplate was attached to each tree along the tree tour for quick identification from a distance. Over time, however, many nameplates were removed, damaged, or destroyed due to weather, growth, or vandalism. The optimum solution proposed was to embed RFID tags within the trees and retrieve the identification numbers via a tree-penetrating RF signal. The updated tour information is professionally read and recorded into PC-formatted sound files. The files are then embedded within the electronic tree tour version. At the click of a button, the text for each tree is played back through the onboard speaker on the PDA device. The advantage is that the user can simultaneously observe the tree and receive audio information. Also, the high-volume capability of the chosen portable information device enables groups to share a single system with one person designated as a chief navigator and tree scanner. In the electronic version of the tree tour, a map with significantly higher detail is displayed on a PDA device with a crisp, color screen at a magnification roughly four times that of the paper version.

The industry's biggest success to date, however, could be in the works in Europe, where rumor has it that the European Central Bank is working with vendors on weaving RFID technology into the fabric of its bank notes. The technology, most probably for incorporation into larger bills, would enable money to carry its own history. Hence, it would become more difficult for kidnapers to ask for unmarked bills. It would also enable law enforcement agencies

to "follow the money" in illegal transactions. The U.S. government is said to have expressed interest as well.

A 2.45-GHz, high-frequency, ultrasmall antenna, embedded in an ultrasmall IC chip (a μ-chip or mu-chip), for RF identification is shown in Figure 3.15. This ultrasmall ($0.3 \times 0.3 \times 0.06$ mm) μ-chip has been developed for use in a wide range of individual recognition applications. The chip is designed to be thin enough to be applied to the paper and paper-like media widely used in retailing to create certificates that have monetary value, as well as tokens. It was designed and fabricated using 0.18-μm standard CMOS technology. Conventional general-purpose μ-chips need an external antenna to transmit the 128-bit identification number inside the chip. In this new type of μ-chip, only a small IC chip is required for battery-less data transmission operations, because the embedded antenna on the chip is able to receive electromagnetic power from the reader [18].

This development enables RFID devices to be easily inserted inside a bank note or gift card. Furthermore, this feature also enables the identification devices to be attached to a narrow surface or inside thin materials. The embedded antenna is formed by using gold plating technology, widely used in the fabrication process of packaging connection terminals on semiconductor wafers.

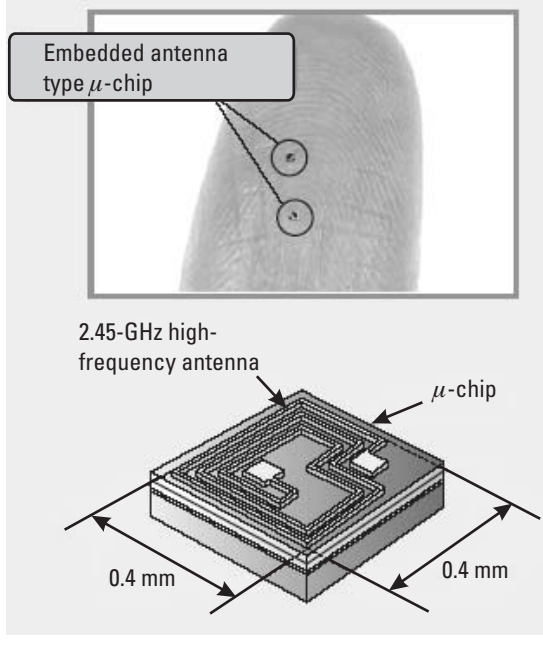


Figure 3.15 2.45-GHz antenna embedded in a chip.

Therefore, the antenna is not only ultrasmall, but also very easily made on the wafer by the fine and batch antenna forming process.

Chinese jails have used RFID to upgrade security by fitting its 6,000 prisoners with RFID wristbands. The new prisoner identification program at Jiangsu Longton Jail in the Nanjing Jiangsu province uses 13.56-MHz ISO 15693 tags and readers. Each wristband contains a smart label with encrypted data including the inmate's name, identification number, and security level. As inmates move around the prison they present their wristbands to a reader that records their identity and the time they entered and left a particular area. Guards are equipped with handheld readers for real-time spot checks and roll calls. Prisoners are required to check into the system for roll call at various times during the day, and if an inmate is not recorded by the system at the appropriate time, an alarm indicates which prisoner has not been identified. Guards can view that prisoner's last check-in point and locate the person quickly. The system has provided the jail with more accurate record keeping and improved prisoner management and is likely to be extended to allow prisoners to make purchases and to record the receipt of mail and other personal items.

To deal with security threats posed by the large volume of container shipping in the United States, the U.S. Customs Service (http://www.customs.gov) is proposing the Container Security Initiative (CSI) to identify high-risk containers and secure them with tamper-detection systems. The initiative aims to expedite processing of containers prescreened at points of embarkation in overseas megaports participating in the initiative. The CSI's basic goal is to first engage the ports that send the highest volumes of container traffic into the United States, as well as the governments in these locations, in a way that will facilitate detection of potential problems at the earliest possible time. To meet this requirement, high-end RFID tags could periodically monitor electronic seals on the containers during transit. This class of application requires tags that can integrate sensor management electronics, such as analog-to-digital converters, and digital data interfaces. Tampering can also be detected in real time, and the tags, as the lowest level of a multitier architecture, can relay data to alert the shippers or customs authorities of tampering as it occurs.

Similarly, tags are used extensively to monitor transportation of high-value goods in the United States as well as worldwide. For example, Norway is a major exporter of salmon, so the RFID tags record the temperature in the containers so that the buyer can verify product freshness. This can be especially important when the shipments are bound for southern locations such as Italy, Spain, or North Africa. The biosensor will have sensitive biological film coatings that will undergo changes in material properties on contact with target pathogens such as *Salmonella* and *Escherichia coli*.

Ski resorts are using smart label technology in lift tickets and passes. They not only authenticate the ticket as genuine and valid for that date and time, but

also increase the traffic-flow rate on the lifts since passes can be read on the fly. The reader signals if there is a problem with a particular ticket. Skiers pass quickly through lift gates without having to remove their gloves to swipe a card, making their experience more pleasurable.

3.6 Review Questions and Problems

1. If RFID has been around so long and is so great, why are only a few companies using it? Discuss at least two main reasons for the slow deployment rate of RFID systems.
2. Is RFID better than using barcodes? Will RFID ever completely replace barcodes?

   Did you know? June 26, 1974, Ohio, USA: The first product using barcodes, a 10-pack of Juicy Fruit chewing gum, is scanned at the check-out counter.
3. Figure 3.16 shows a flexible substrate. What is the importance of materials like this one in the development of RFID technology?
4. Use of RFID in the pharmaceutical industry could prevent most of the 1.25 million adverse reactions and 7,000 patient deaths that occur annually in the United States as a result of drug errors, according to the Meta Group Consultancy. Discuss this issue.
5. Transponder systems equipped with one or more sensors are able to measure continuously or time discrete physical values such as temperature, pressure, or acceleration. Measurements and data storage outside the electromagnetic field of the reader is only possible if an energy source like a battery or solar cell is available.

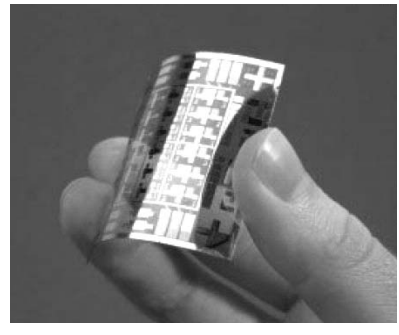


Figure 3.16 Flexible substrate.

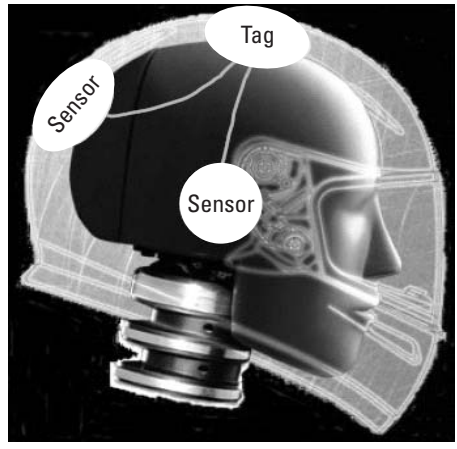


Figure 3.17 Motorbike helmet with sensor network.

A hard hat, like that shown in Figure 3.17, worn by building workers or motorcyclists has been supplied with an RF tag and a sensor network with three acceleration sensors [19]. The maximum acceleration values and their distribution can be measured in case of an accident. With a reader it is possible to read out these measurements quickly. Try to envision other applications of wireless sensors where human lives could be protected and/or saved.

6. RFID systems operate at four major frequency ranges. As a rule of thumb, low-frequency systems are distinguished by short reading ranges, slow read speeds, and lower cost. Higher frequency RFID systems are used where longer read ranges and fast reading speeds are required, such as for vehicle tracking and automated toll collection. Microwave requires the use of active RFID tags.

   Analyze Table 3.1 and answer the following questions:

   a. You are planning an RFID system for tracking the location of people within a building. What frequency band(s) would you consider for this application? Elaborate on your answer.
   b. You are planning an RFID system for counting a large number of fast moving and cheap products on a conveyer belt. What frequency band(s) would you consider for this application? Elaborate on your answer.
   c. You are planning to deploy an RFID system but the environment already contains a large WLAN network with many users. What RFID bands may not be suitable for your system? Elaborate on your answer.

Table 3.1
RFID Summary Table

Frequency	Range (ft)	Tag Cost	Applications
Low-frequency 125–148 kHz	3	$1+	Pet and ranch animal identification; car keylocks
High-frequency 13.56 MHz	3	$0.50	Library book identification; clothing identification; smart cards
Ultrahigh frequency 915 MHz	25	$0.50	Supply chain tracking: box, pallet, container, trailer tracking
Microwave 2.45 GHz	100	$25+	Highway toll collection; vehicle fleet identification

References

[1] Smail, T., et al, "Antennas for RFID Tags," *Joint SoC-EUSAI Conference,* Grenoble, France, 2005.

[2] *RFID Compendium & Buyer's Guide, 2004–2005,* Auto ID Service Providers Ltd., on behalf of AIM UK.

[3] Fuhrer, P., et al., "RFID: From Concepts to Concrete Implementation," Fribourg, Switzerland: Department of Informatics, University of Fribourg, 2005.

[4] Yu, P., et al., "Securing RFID with Ultra-Wideband Modulation," Blacksburg, VA: Electrical and Computer Engineering Department, Virginia Tech, 2006.

[5] Molnar, D. A., "Security and Privacy in Two RFID Deployments, with New Methods for Private Authentication and RFID Pseudonyms," Berkeley, CA: Department of Electrical Engineering and Computer Sciences, University of California–Berkeley, 2006.

[6] Griffin, S., and C. Williams, "RFID Futures in Western Europe," White Paper, Juniper Research, 2005.

[7] Dulman, S. O., "Data-Centric Architecture for Wireless Sensor Networks," Ph.D. Dissertation, University of Twente, Enschede, the Netherlands, 2005.

[8] Lewis, F. L., "Wireless Sensor Networks," Advanced Controls, Sensors, and MEMS Group, Automation and Robotics Research Institute, University of Texas at Arlington, 2004.

[9] Martin, A., H. Marini, and S. Tosunoglu, "Intelligent Vehicle/Highway System: A Survey, Part 1," Miami, FL: Department of Mechanical Engineering, Florida International University, 2000.

[10] Chon, H. D., et al., "Using RFID for Accurate Positioning," *J. of Global Positioning Systems,* Vol. 3, Nos. 1–2, 2004, pp. 32–39.

[11] "Item-Level Visibility in the Pharmaceutical Supply Chain: A Comparison of HF and UHF RFID Technologies," White Paper, Philips Semiconductors, TAGSYS, Texas Instruments, Inc., July 2004.

[12] Miller, L. E., et al., "RFID-Assisted Indoor Localization and Communication for First Responders," National Institute of Standards and Technology, 2005.

[13] H. R. 5631, "Defense Appropriations Act for Fiscal Year 2007," June 16, 2006, http://www.congress.gov.

[14] Trebilcoch, B., "Active RFID Enables the Military Supply Chain," *Modern Materials Handling,* http://www.mmh.com, September 28, 2006.

[15] Swedberg, C., "US Army Developing RFID System to Track Weapons Usage," *RFID Journal,* November 9, 2006.

[16] d'Hont, S., "The Cutting Edge of RFID Technology and Applications for Manufacturing and Distribution," Texas Instruments TIRIS, 2003.

[17] Hoyt, S., et al., "A Tree Tour with Radio Frequency Identification (RFID) and a Personal Digital Assistant (PDA)," *29th Annual Conference of the IEEE Industrial Electronics Society,* Roanoke, VA, November 2–6, 2003.

[18] Usami, M., "An Ultra-Small RFID Chip: μ-Chip," Central Research Laboratory, Japan Hitachi Technology, Kokubunji City, Tokyo, 2004–2005.

[19] Fischer, W. J., et al., "Smart RF-Transponder Chips with On-Chip or External Sensors for Usage in Portable Systems," Dresden, Germany: Department of Electrical Engineering, Dresden University of Technology, 2003.

4

RFID Standards Development Challenges

4.1 Regional Regulations and Spectrum Allocations

The growing need for interoperable products requires the standardization of regulatory controls on spectrum usage and international standards to support compatibility or interoperability of RFID systems. Standardization factors enter into the business thinking when considering identification and data capture solutions that are required to operate in different countries and/or require compatibility for the purposes of data capture and transfer. RFID data carriers (tags) and associated systems are generally considered to be part of a general category of radio-based short-range devices designed to operate in regions of the electromagnetic spectrum that do not require operating licenses and do not incur operating fees. The use of the EM spectrum, particularly for radio usage, is carefully controlled. Spectrum allocations are specified with a view toward avoiding systems interfering with one another. The unlicensed regions of the spectrum are generally allocated for ISM applications. Unfortunately, the regulations for spectrum usage vary from country to country. For the user of RFID systems, the expectation is that the necessary regulatory requirements will be satisfied. Although transparent to the user, the manufacturers of RFID systems will need to have a detailed understanding of the requisite regulations and associated standards and be able to demonstrate to users that their products are compliant.

Generally speaking, the manufacturers are required to identify and comply with the essential regulations and standards with respect to:

- Spectrum allocations and associated operating constraints;
- Health and safety;

- Electromagnetic compatibility;
- Avoidance of interference with other spectrum users;
- Compliance with national interface regulations;
- Other regulations and directives that may arise from time to time concerning system usage.

They may also be required to carry out essential testing for compliance purposes. The necessary requirements to be met are available from the spectrum management authorities in respective countries.

For low frequencies (<135 kHz) and high frequencies (13.56 MHz), a fair degree of allowed usage exists worldwide for RFID purposes. However, at the UHF carrier frequency, the situation is more complicated. Differences between the United Kingdom and the United States in the allocation of mobile phone usage, for example, prevent the use of common carrier frequencies and associated bands. Country-specific regulatory controls specify the field strengths and power levels allowable for devices and systems operating at the different carrier frequencies. These levels naturally have a determining influence on the ranges that are achievable for the reactively coupled and propagation coupled systems. It is therefore important to establish and confirm what is allowable within the country in which the technology is to be used. The standardization process and regional efforts to agree on frequency usage are likely, in the fullness of time, to provide a foundation for wider open systems exploitation.

It must be stressed that considerations for the use of RFID in different countries should include up-to-date information on spectrum allocation and constraints on use obtained from the appropriate regulatory authority within the country concerned, as well as any encompassing regulations operating at union and international levels.

It is understood that the use of any modern technology requires some form of harmonization at the national and international levels. Standardization ensures compatibility and interoperability between different manufacturers and technical applications. This chapter discusses the standards landscape and presents areas of concern for future standardization initiatives. In addition to the spectrum-related initiatives, RFID standardization is needed in the following areas:

- *Air interface and protocols:* Communication between tags and readers, readers and readers, RFID systems and other wireless communication systems;
- *Data structures:* Organization of the data (e.g., on a tag);
- *Conformance:* Tests ensuring that products meet the standards;

- *Applications:* Use of the technology for a particular purpose.

The importance of standards cannot be overemphasized for technologies, such as RFID, that have universal relevance and significant potential for open systems usage. In addition, RFID systems generate radio signals that can interfere with other radio applications. Therefore, they must comply with regulations that state which frequencies may be used and the maximum power of transmission. In Europe, the European Radiocommunications Committee (ERC) regulates the use of frequencies. In the United States, radio devices have to comply with the FCC's licensing regulations. RFID systems are not separately regulated, but they do need to comply with regulations based on the frequency used. The regulations of other countries often correspond with either the U.S. or European regulations. One exception is Japan, which has a regulation system that differs from both the U.S. one and the European one. However, it is often enough just to consider the FCC rules of the United States and the European ETSI standards.

4.2 Key Players in RFID Standardization

Various players are involved in working with and standardizing RFID technologies. This section presents some major players in RFID standardization activities and spectrum allocation.

The Automotive Industry Action Group (AIAG), an association of 1,600 members involved in the automotive and truck manufacturing supply chain, has been developing RFID specifications for the automotive industry. A general standard exists (*ARF 156: Application Standard for RFID Devices in the Automotive Industry*) and is accompanied by several standards dealing with special sectors of the automobile supply chain, such as AIAG B-11, a standard to identify tires and wheels with RFID devices. AIAG B-11 has been developed together with EPCglobal. AIAG also holds regular conferences on the use of RFID in the automotive industry.

The European Article Numbering (EAN) and the Uniform Code Council (UCC), which administers and manages the EAN-UCC standards system in the United States and in Canada, launched the Global Tag Initiative (GTAG) in March 2000. It is a standard that covers UHF RFID technology and data formats. The air interface aspects of GTAG have now been merged with ISO 18000 Part 6. EAN International and UCC started EPCglobal as a joint venture.

At a European level, the European Radiocommunications Office (ERO) has been working on radiocommunications policies and frequency allocation that are important to RFID technologies. The ERO is the permanent office that

supports the Electronic Communications Committee (ECC). This committee, in turn, is the telecommunications regulation committee for the European Conference of Postal and Telecommunications Administrations (CEPT). Decision ERC/DEC (01)04 of 2001 addresses the use of nonspecific SRDs, such as RFID, in certain UHF frequency bands. Other important CEPT regulations related to RFID technologies include CEPT/ERC 70-03 (relating to the use of SRDs), CEPT T/R 60-01 (low-power radiolocation equipment for detecting movement and for alerts, EAS), CEPT T/R 22-04 ("Harmonisation of Frequency Bands for Road Transport Information Systems, RTI").

The European Telecommunications Standards Institute (ETSI) has also been very active in the field of RFID standardization. In 2004, ETSI TG34, the technical group for electromagnetic compatibility and radio spectrum matters, completed, in cooperation with EPCglobal, standard EN 302 208, which allows readers to use more power and operate in a wider UHF band. Prior to that, ETSI had already developed the ETSI EN 300-220 standard.

The International Air Transport Association (IATA) is studying RFID technologies for airline baggage management. A subgroup of the Baggage Working Group (BWG) is responsible for this. IATA has already adopted 13.56 MHz and ISO/IEC 15693 in its "Recommended Practice for Airline Baggage RF Identification" (RP1745).

Within the framework of the International Civil Aviation Organization (ICAO), the Technical Advisory Group on Machine Readable Travel Documents (TAG/MRTD) has been working on international travel documents (e.g., a passport or visa) containing eye- and machine-readable data. Specifications for the design of these travel documents are contained in ICAO document 930376. An annex to that document contains provisions on contactless integrated circuits, referring in particular to the ISO/IEC 14443 standard.

Within the International Committee for Information Technology Standards (INCITS), Technical Committee T6 deals with RFID standardization. T6 has developed NCITS 256, an RFID standard for use in item management. INCITS is accredited by the American National Standards Institute (ANSI).

The ISO and the International Electrotechnical Commission (IEC) have undertaken major activities to standardize RFID technologies in various areas. The work on RFID standardization for the ISO and IEC has been carried out mainly under ISO/IEC Joint Technical Committee 1 (JTC 1), subcommittees 17 (Identification Cards and Personal Identification) and 31 (Automatic Identification and Data Capture Techniques). Inside of SC31, WG481 is responsible for RFID devices for item management. Other groups within the ISO framework that work on RFID topics include ISO TC 23/SC 19 (Agricultural Electronics), ISO TC 104/SC (Identification and Communication), and ISO/TC204 (Transport Information and Control Systems).

The International Telecommunication Union (ITU) is the United Nation's specialized agency for telecommunications. Two of its three sectors are dealing with RFID-related issues: the Radiocommunication Sector (ITU-R), and the Telecommunication Standardization Sector (ITU-T). As global spectrum coordinator, ITU-R plays an essential role in the management of the RF spectrum. It is governing the use of the radio spectrum by some 40 different services around the world. ITU-R Study Group 185 is responsible for spectrum management. The Telecommunication Standardization Sector (ITU-T) creates globally agreed-on and globally accepted ICT standards.

The Universal Postal Union (UPU) is the United Nation's specialized agency for cooperation between postal services. UPU's Technical Standards Board has been developing standards for using RFID technologies in postal applications for several years. Among these standards are S25-1G ("Data Constructs for the Communication of Information on Postal Items, Batches and Receptacles"), S23-1 Parts A, B, C, and G ("RFID and Radio Data Capture Systems"), UPU RF 0001.2 ("Identification and Marking Using RFID Technology: Data Schemes"), UPU Snn-1 ("Identification and Marking Using RFID Technology: Reference Architecture and Terminology"), UPU Snn-3 ("Identification and Marking Using RFID: System Requirements and Test Procedures").

4.3 ISO and EPC Approaches

Standards and specifications may be set at the international, national, industry, or trade-association levels, and individual organizations may call their own specifications standards. Many industry standards and specifications set by individual organizations are based on international standards to make implementation and support easier and to provide a wider choice of available products. Standards can be applied to include the format and content of the codes placed on the tags, the protocols and frequencies that will be used by the tags and readers to transmit the data, the security and tamper resistance of tags on packaging and freight containers, and applications use. The ISO and EPCglobal have both been leading figures in this debate; the ISO has their 18000 standard and the EPCglobal Center has introduced the EPC standard. Wal-Mart has decided to use the EPC standard, whereas the DoD wants to use the EPC for general purposes, but use the ISO standard for the air interface. This is putting a lot of pressure on the ISO and EPC to come to some kind of an agreement.

What is EPC? The Auto-ID Center has proposed a new Electronic Product Code (EPC) as the next standard for identifying products. The goal is not to replace existing barcode standards, but rather to create a migration path for companies to move from established standards for barcodes to the new EPC. To encourage this evolution, we have adopted the basic structures of the Global

Trade Item Number (GTIN), an umbrella group under which all existing barcodes fall. There's no guarantee that the world will adopt the EPC, but the proposal already has the support of the UCC and EAN International, which are the two main bodies that oversee international barcode standards. Other national and international trade groups and standards bodies are working together.

The ISO is based in Geneva, and its standards carry the weight of law in some countries. All ISO standards are required to be available for use around the world, so users of ISO RFID standards will not have to worry if their systems comply with the different regulations on frequencies and power output for each country where they do business. The ISO is very active in developing RFID standards for supply chain operations and is nearing completion of multiple standards to identify items and different types of logistics containers. As appropriate, EPCglobal submits its standards to ISO for review and ratification because ISO considers the new EPC standards to be a subset of its standards.

Some in the industry anticipate a future where everyday objects and appliances are connected to the Internet and each other via RFID tags. This so-called Internet of things will enable people to interact more with their environment. RFID will most likely be one of the most important standards in the decades to come and the promise of an Internet of things has caused major countries and geopolitical entities to support several national and regional standards. EPCglobal, the not-for-profit standards organization, is driving the development of a universal EPC system and a global information network to enable automatic identification of items in the supply chain. The EPC standard is backed mainly by U.S. vendors, whereas the ISO standard is more widely supported and accepted in Europe. Japan has UID (Ubiquitous ID), and China has recently announced its intent to create its own RFID standards. Even though the EPC standard has the strongest commercial support, the true global standards are still under development and will be subjected to geopolitical forces.

Appropriate standards allowing numerous companies to create interoperable products are a key prerequisite for the widespread use of RFID tags. ISO 15693, accepted in 2000, is one such standard (see http://www.iso.org). It is titled "Identification Cards—Contactless Integrated Circuit(s) Cards—Vicinity Cards" and has three parts: physical characteristics, the air interface and initialization, and an anticollision and transmission protocol. It specifies a 13.56-MHz RFID protocol, originally proposed by Texas Instruments and Philips Semiconductors in 1998, defining data exchange between RF tags and readers, and collision mediation when multiple tags are in a reader's RF field. Compliance guarantees that RF tags and readers using the ISO 15693-2 protocol will be compatible across companies and geographies.

4.4 RFID Systems and Frequencies

4.4.1 Power Emissions Conversion

Before discussing the details of the regulations, note that there are several accepted means of describing one of the most important regulated parameters of an RF device—its radiated emissions. All unlicensed devices have some limitation on the amount of output power, or radiated energy, they can produce. What differs among the various regulatory agencies is the means of describing this limit. Radiated energy can be described in terms of:

- Electrical field strength (E) measured at some distance from the radiator;
- Effective isotropic radiated power (EIRP);
- Effective radiated power (ERP).

The electrical field strength (E) is perhaps the most precise way of describing the actual RF energy present at a point in space that a receiving antenna could use. Because RF energy decreases with increasing distance from the transmitting antenna, the regulatory limits based on electrical field strength are specified at a specific distance from the transmitting antenna. Although the electrical field strength may be a precise measure, it is often not as useful from a design perspective as the effective isotropic radiated power, or EIRP.

The EIRP is the power that would have to be supplied to an ideal antenna that radiates uniformly in all directions in order to get the same electrical field strength that the device under test produces at the same distance; such an antenna is called an *isotropic radiator*. Given a distance r from the transmitting antenna, the EIRP can be calculated from E using the following formula:

$$\text{EIRP} = 10\log\left(4\pi E^2 r^2 / 0.377\right) = 10\log\left(E^2 r^2 / 0.03\right) \quad [\text{dBm}] \qquad (4.1)$$

where the EIRP is in dBm, E is the field strength in volts per meter, r is distance in meters, and 0.377 is expressed in V^2 (unit of measurement.) Contrary to the European regulations, the U.S. regulations are in most cases specified in terms of field strength, not power. The field strength is measured at 3m from the device under test.

The ERP is similar to the EIRP. It is the power that would have to be supplied to a half-wave dipole to get the same electrical field strength that the device under test produces at the same distance. A half-wave dipole is more representative of realistic simple antennas than an isotropic radiator. Because a half-wave dipole radiates more energy in some directions and less in others, it is said to

have antenna gain in the direction of the most energy. The amount of antenna gain is usually expressed in decibels relative to an isotropic radiator, or dBi. The maximum gain of a half-wave dipole is 2.15 dBi. Thus, in the direction of maximum gain, the amount of power required to produce the same electric field is less with a half-wave dipole than it is with an isotropic radiator (see Table 4.1). If both the ERP and the EIRP are expressed using a logarithmic scale, such as dBm, it holds that:

$$\text{ERP} = \text{EIRP} - 2.15\,\text{dB} \tag{4.2}$$

or

$$\text{EIRP} = 1.64\ \text{ERP} \tag{4.3}$$

If one of the three parameters—EIRP, ERP, or E—at a distance r is given, the other two can be calculated. These reference parameters are used to define the permitted radiated power in RF systems, and they are often quoted in RFID reader and tag specifications. The United States tends to use EIRP, whereas Europe uses ERP.

4.4.2 North American and International Frequency Bands

The RF spectrum is a scarce and shared resource, used nationally and internationally, and subject to a wide range of regulatory oversight. In the United States, the FCC is the key regulatory body that allocates spectrum use and resolves spectrum conflicts. The ITU is a specialized agency of the United Nations that plays the same role internationally.

Development of RFID systems suitable for deployment internationally faces some considerable problems as a result of disparate frequency regulations and the transmitting power of the reader. It is not only that the frequency bands

Table 4.1
ERP/EIRP Conversion Example

	EIRP (W)	Gain	ERP Power Fed to the Antenna (W)
Isotropic antenna	4	1	4
Dipole antenna	4	1.64	2.44
Antenna	4	3	1.33

have been allocated differently in different countries, but the permitted transmitting power for the reader also varies, which means that identical models can differ considerably in their range. The frequency ranges of 125 kHz and 13.56 MHz are generally regarded as having been largely standardized. The standardization process for the remaining frequency ranges is being carried out on the basis of the ISO/IEC 18000 standards.

Both the U.S. and EU regulatory agencies place limitations on the operating frequencies, output power, spurious emissions, modulation methods, and transmitting duty cycles, among other things. RFID tags and readers fall under the category of short-range devices, which although they do not normally require a license, the products themselves are governed by laws and regulations that vary from country to country. Today, the only globally accepted frequency band is the HF 13.56 MHz.

For passive UHF RFID the problem is much more complicated because frequencies allocated in some countries are not allowed in others due to their proximity to already allocated bands for devices such as mobile phones and alarms. This discontinuity has resulted in the ITU dividing the world into three regulatory regions (Figure 4.1):

- *Region 1:* Europe, Middle East, Africa, and the states of the former Soviet Union including Siberia;

- *Region 2:* North and South America and the Pacific east of the international date line;

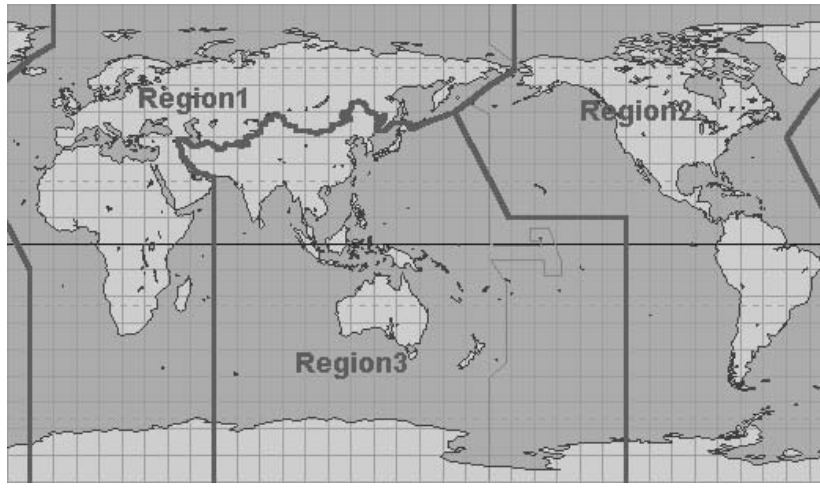


Figure 4.1 ITU regions.

- *Region 3:* Asia, Australia, and the Pacific Rim west of the international date line.

The main regulatory bodies in different regions are:

- In the United States, the FCC;
- In Europe, CEPT;
- In Japan, the Ministry of Public Management, Home Affairs, Posts and Telecommunication (MPHPT)

4.4.3 RFID Interoperability and Harmonization

When considering interoperability for global use of RFID devices, it is first necessary to consider spectrum allocation. Perhaps the most essential aspect of worldwide acceptance of RFID is to have spectrum available in all relevant markets so that RFID can be used where needed [1]. RFID systems in the United States, as with other ubiquitous RF devices used by the general public, are unlicensed. In the United States, RF radiation from intentional and unintentional unlicensed radiators is regulated by the FCC under Part 15 of Title 47 of the *Code of Federal Regulations*. These regulations delineate the technical specifications, such as allowable frequency, power limits, and other operational constraints, under which an intentional radiator may be operated without an individual license. Part 15 also allows operation of RFID systems over a vast range of frequencies but places limits on the allowable output power of the system.

Every available operating frequency has a specific power limit associated with it. The combination of frequency and allowable power level are the factors that dictate the functional range of the particular RFID application, whether over a range of centimeters or hundreds of meters. These allowable power levels are particularly relevant for the power output of the readers, which have a far higher power output than the tags.

Country-specific regulatory controls specify the field strengths and power levels allowable for devices and systems operating at the different carrier frequencies. These levels naturally have a determining influence on the ranges that are achievable for the reactively coupled and propagation coupled systems. It is therefore important to establish and confirm what is allowable within the country in which the technology is to be used. In addition to power limits, there are restrictions on the use of certain bands by unlicensed devices to prevent potentially harmful interference to systems used for services such as safety, search and rescue, aeronautical communications, and scientific research. In addition,

unlicensed systems operating in the United States under Part 15 must accept any interference from other systems in these bands, including interference from other unlicensed devices.

Many countries do not have rules like those in the United States for unlicensed systems. Instead, they allocate spectrum on a primary or secondary basis and require that all radio transmitters be licensed by the government. Also, some countries that allow RFID devices currently do not recognize active tags within their regulations. Similar to the United States, several other countries do not allocate spectrum specifically for RFID but instead allocate categories of service such as short-range devices.

The ITU-R recently released Recommendation SM.1538-1, "Technical and Operating Parameters for Short-Range Radio Communication Devices," outlining the spectrum requirements and regulatory approaches applicable to SRDs in Europe, the United States, the People's Republic of China, Japan, and South Korea. Table 4.2 contains a partial list of countries currently using RFID and the frequency bands allowed for RFID operations.

There are exceptions in terms of allowable frequency and functional use of the bandwidth for critical operations both in the United States and in other

Table 4.2
RFID Operational Frequencies

Band	Frequency	System	Regions/Countries
Low frequency (LF)	125–134 kHz	Inductive	United States, Canada, Japan, and Europe
High frequency (HF)	13.56 MHz	Inductive	United States, Canada, Japan, and Europe
Very-high frequency (VHF)	433.05–434.79 MHz	Propagation	In most of Europe, United States (active tags at certain locations must be registered with the FCC), and under consideration in Japan
Ultrahigh frequency (UHF)	865–868 MHz	Propagation	Europe, Middle East, Singapore, Northern Africa
Ultrahigh frequency (UHF)	866–869 and 923–925 MHz	Propagation	South Korea, Japan, New Zealand
Ultrahigh frequency (UHF)	902–928 MHz	Propagation	United States, Canada, South America, Mexico, Taiwan, China, Australia, Southern Africa
Ultrahigh frequency (UHF)	952–954 MHz	Propagation	Japan (for passive tags)
Microwave	2.4–2.5 and 5.725–5.875 GHz	Propagation	United States, Canada, Europe, Japan

countries. For example, the allowable bandwidth and power for RFID devices are not generally the same from ITU region to region. In ITU Region 1, which covers most of Europe, the ISM radio band (13.553 to 13.567 MHz) is much narrower than the bandwidth that is allowed in the United States for RFID. Also, use of the band at 433 MHz for active RFID in the United States is intended for container tracking and is allowed at higher power levels than generally permissible for unlicensed devices in this band. However, operations of such systems are limited primarily to industrial locations, such as railheads and shipyards, and the systems must be registered with the FCC. In Europe, however, the 433.05- to 434.79-MHz band is an ISM band and is allocated on a primary basis to short-range devices, such as RFID. In addition, many nations in East Asia (e.g., China, Japan, and South Korea) are currently developing their own regulations for RFID. For example, Japan is in the process of revising its regulations to allow the 950- to 956-MHz band to be used for unlicensed, low-power passive tag RFID systems. They are also establishing a licensing structure for high-power (power levels up to any level not hazardous to humans at specified distances) passive RFID systems that will be used in industrial areas.

In general, RFID chips can be used to track products grouped in various hierarchies:

- Individual items or single packages containing multiple items for consumer purchase;
- Cartons or cases of multiple items;
- Pallets of multiple cartons or cases;
- Loads (e.g., truckloads, shiploads, or railcar loads) of multiple pallets.

The products at each of these levels may be assigned an RFID label that is associated with information pertaining to at least one adjacent hierarchical level. For example, an RFID label on a pallet may be associated in a database with the RFID labels for each carton on the pallet, or may be associated with data pertaining to the RFID label from the truckload. Figure 4.2 illustrates needs for different requirements for different types (layers) of RFID tagging. The RF power, the frequency used, and therefore the reader-tag distance will be different for different applications.

Although some of the ISM frequency bands are internationally recognized (e.g., 13.56 MHz and 2.45 GHz), others are not. In the United States, 915 MHz is recognized as an ISM band, but 433 MHz is not; however, in Europe the 915 MHz is not recognized, but 433 MHz is. Differences like this add to the challenges of harmonization of RFID bands. In recognition of the complications associated with international harmonization of frequency bands, the U.S. DoD requires that passive RFID systems be both multimode and multiband to

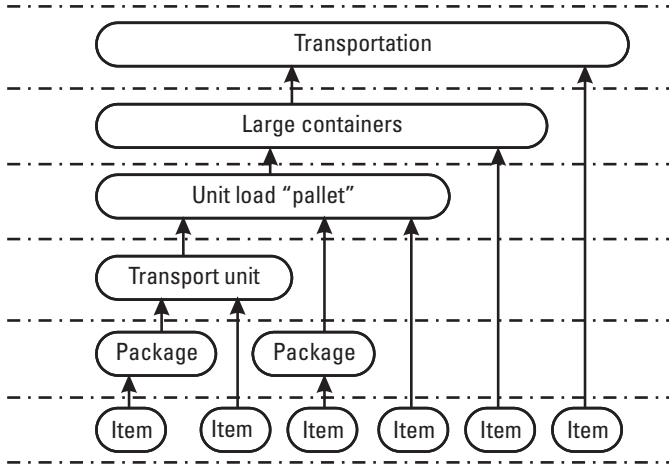


Figure 4.2 Different layers of RFID tagging.

meet global requirements. Accordingly, DoD has issued a set of policies mandating system performance and functionality for implementing RFID systems within the DoD supply chain. Additionally, it specifies operation in the 860- to 960-MHz frequency range and requires that passive RFID systems be capable of operating within this range.

4.4.4 Advantages and Disadvantages of Using the 13.56-MHz Frequency

Much current activity revolves around an RFID frequency of 13.56 MHz. The reason for this attention is that several RFID technology providers have designed their product offerings based on this particular frequency; 13.56 MHz is, after all, an ISO standard for smart card applications. Smart cards differ from RFID tags in that they are used for more than data storage. In some smart card applications, the card itself contains a microprocessor that can define parameters for data storage and byte allocation. Smart card applications require that certain RFID attributes be present for implementations of that technology to be successful. Since 13.56 MHz provides for RFID near-field read/write capability, it is ideally suited to smart card applications where secure financial transactions are being transmitted. Understandably, a smart card user would not want his or her bank account number to be able to be read at a distance of 3m to 5m. Why are these attributes important? If a technology or specific frequency in this case were targeted for a particular application (smart cards), then it may not have the necessary attributes that would be needed for another RFID application, such as item management. The 13.56-MHz inductive passive RFID systems are one of the mainstream RFID products and some of the main advantages and disadvantages of this band are listed next.

Advantages of 13.56-MHz RFID:

- Frequency band available worldwide as an ISM frequency;
- Well suited for applications requiring reading small amounts of data and minimal distances;
- Popular smart card frequency;
- ISO 15693, ISO 14443, and HF EPC standardization for the air interface;
- Robust reader-to-tag communication;
- Excellent immunity to environmental noise and electrical interference;
- Well-defined and localized label interrogation zones;
- Minimal shielding effects from adjacent objects and the human body;
- Penetrates water/tissue well;
- Freedom from the environmental reflections that can affect UHF and microwave systems;
- Good data transfer rate;
- High clock frequency and synchronous subcarrier;
- On-chip capacitors for tuning transponder coil can be easily realized;
- Uses normal CMOS processing, cheap ICs, and disposable tags;
- Cost-effective antenna coil manufacturing;
- Low RF power transmission so EM regulation compliance does not cause problems;
- No user licenses for reader systems required (ISM band);
- Possible to use the systems in industrial and in hazardous environments with potential for explosive substances.

Disadvantages of 13.56-MHz RFID:

- Government-regulated frequency (United States and Europe recently harmonized);
- Does not penetrate or transmit around metals;
- Large antennas (compared to higher frequencies);
- Larger tag size than higher frequencies;
- Tag construction requires more than one surface to complete a circuit;
- Reading range of less than 1.0m (3 feet).

4.4.5 Operation in the 900-MHz Band

Traditionally, passive transponders operate at 125 kHz or 13.56 MHz, using coils as antennas. These transponders operate in the magnetic near field of the base station's coil antenna, and their reading distance is typically limited to less than 1.0m (3 feet). A problem of these systems is the low efficiency of reasonably large antennas at such low frequencies. Due to great demand for higher data rates, longer reading distances, and small antenna sizes, there is a strong interest in UHF frequency band RFID transponders, especially for the 868- to 915-MHz bands. The frequency band of 902 to 928 MHz is one of the ISM bands in the United States (FCC Part 15.247 regulations), commonly abbreviated as the 915-MHz ISM band. In this band, there are no restrictions on the application or the duty cycle because they were intended for periodic applications only. Furthermore, the allowed power output is considerably higher. Because of the lack of restrictions and higher allowed power, this band is very popular for unlicensed short-range applications including audio and video transmission. FCC Section 15.249 allows 50 mV/m of electrical field strength at 3m distance in the frequency band of 902 to 928 MHz. This corresponds to an EIRP of –1.23 dBm. The harmonics limit is one-hundredth of the fundamental level, 500 μV/m, corresponding to an EIRP of –41.23 dBm [2].

The frequency of operation for the reader to tag communication in passive UHF RFID is not fixed. The reader does frequency hopping in the ISM UHF band for communicating with the tags. The frequency hopping avoids interference that might occur due to other devices using some part of the ISM band's spectrum. Also, the modulation schemes used in reader-to-tag communication depend on the type of the protocol being read.

The U.S. readers will have a maximum output power of +30 dBm (1W), the European readers will have a maximum output power of +27 dBm, and the Japanese readers will have a maximum output power of +30 dBm. A new requirement for *dense reader mode* is being introduced in the United States that will be used when there are a high number of readers in proximity. Most 1W readers to date in the United States are designed for low-density usage, and their emissions and out-of-band requirements are defined by the FCC. For the U.S. and Japanese readers, a saturated power amplifier can be used, which allows for higher power amplifier efficiency (PAE) on the order of 50%. European and dense reader mode readers will be required to operate the power amplifier in the linear range, which reduces PAE to approximately 30%.

Even higher output power can be used if the system employs some form of spread spectrum, such as frequency-hopping or direct-sequence spread spectrum. The reason such allowances are made is that spread-spectrum systems are less likely to interfere with other systems than are single-frequency transmitters and they are often more immune to interference from other systems.

For more details on UHF regulations in Europe, see publication ERC REC 70-73, which can be found at http://www.ero.dk/doc98/official/pdf/rec7003e.pdf.

Advantages of 915-MHz RFID:

- Unlicensed ISM band in United States;
- Read range of up to 20 feet (6m);
- Significantly higher data rates than LF RFID;
- Recent widespread use (2004–2006);
- Can cover dock door portals up to 9 feet wide; therefore, is suitable for inventory tracking applications including pallets and cases.

Disadvantages of 915-MHz RFID:

- Absorbed by liquids, although Gen 2 tags and antenna designs are less susceptible;
- Unpredictable performance near metal, although Gen2 tags and antenna designs less susceptible;
- Interferes with existing bands in some countries, but receiving certification;
- Common sources of interference include 900-MHz cordless phones, older 900-MHz wireless LANs, metal supports, equipment, and cabinets;
- As a consequence of sometimes unpredictable radio propagation due to reflections, some tags may be successfully read from a large distance away from the reader, while neighboring tags may receive little power from the reader.

Efficient antennas are commonly constructed with dimensions on the order of a half of the wavelength, in this case around 6 inches (150 mm), an inconveniently large size for many desirable applications. It is possible to compress the linear dimension by bending the conductive regions, or using large-area structures, though inevitably with some compromise in radiation resistance and thus performance. At these frequencies, tags are also sensitive to environmental effects, and optimal designs may change depending on the material to which the tag is to be attached.

4.4.6 Operation in the 2.45- and 5.8-GHz Bands

The basic operating principle of 2450-MHz RFID systems is energy and data transmission using propagating radio signals (E-field transmission). This is exactly the same principle as that used in long-range radio communication systems. The reader's antenna generates a propagating radio wave, which is received by the antenna in the tag. A passive power tag converts the signal to dc voltage to supply the tag with energy. Data transmission from the reader to the tag is achieved by changing one parameter of the transmitting field (amplitude, frequency or phase). The return transmission from the tag is accomplished by changing the load of the tag's antenna (amplitude and/or phase). In this context, the microwave, 13.56-MHz and 125-kHz systems use the same principle. For microwave RFID systems, this method is called *modulated backscatter*. Alternatively, a signal of different frequency can be generated, modulated, and transmitted to the reader. These latter types of systems are referred to as using active RF transmitter tags. The 2.4- to 2.4835-GHz band is another ISM band covered by FCC Sections 15.247 and 15.249.

Advantages of 2.4-GHz RFID:

- The 2.4-GHz band is a worldwide unlicensed band, and this is an important advantage compared to, for example, the 902- to 928-MHz band.
- The 2.4-GHz band also has a wider bandwidth than the 902- to 928-MHz band, which means more available channels.
- Good reflections off metal surfaces allow for relatively good propagation in cluttered environments.
- It has reasonable propagation through nonconductive materials, such as wood and wood-based products, natural and synthetic garments, and plastics.
- Because tags operating in the E-field do not require antennas with extremely low impedances, inexpensive flexible antennas that are able to withstand considerable bending are achievable.

Disadvantages of 2.4-GHz RFID:

- The active components are more expensive and have higher current consumption.
- It has a reduced propagation distance for the same power as other systems.

- Moisture and moisture-containing substances can exhibit energy-absorbing mechanisms at microwave frequencies.
- For ranges in excess of 1.0m (3 feet), multipath effects and fading need to be considered.
- It shares spectrum allocation with spread-spectrum radio LANs, WLANS, microwave ovens, Bluetooth, TV devices, and so forth.
- It is more susceptible to electronic noise than lower UHF bands (e.g., 433 MHz, 860 to 930 MHz).
- Regulatory approvals are still in process.

For single-frequency or other systems that do not qualify as spread spectrum, the same transmitting power limits as in the 902- to 928-MHz band apply. The only difference is that an averaging detector can be used in the 2.4-GHz band, allowing a higher peak output power, with the limitation that the peak electrical field strength must not be more than 20 dB above the average value. Thus, similar to the control and periodic applications described earlier, the transmitting strength can be up to 20 dB larger than the limits for continuous signals if the duty cycle is reduced accordingly.

As with the 902- to 928-MHz band, larger transmitting power levels are allowed if the spread spectrum is used. The criteria for a frequency-hopping system in the 2.4-GHz ISM band are as follows:

- The transmitter hops pseudo-randomly between at least 15 nonoverlapping frequency channels.
- The average time of occupancy at any frequency must not be larger than 0.4 second within a time period of 0.4 second multiplied by the number of channels.

The permitted peak transmitting power measured at the antenna input of a frequency-hopping system with at least 75 hopping frequencies is +30 dBm. For systems with less than 75 but at least 15 hopping frequencies, a peak transmitting power of +21 dBm is allowed. Similar to the 902- to 928-MHz band, the power has to be reduced if the isotropic antenna gain is larger than 6 dBi. This allows a maximum EIRP of +36 dBm in a system with at least 75 channels or +27 dBm in a system with less than 75 but at least 15 channels.

Systems operating at microwave frequencies may be considered to exhibit quasi-optical features with the facility to form well-defined beams. By exploiting these spatial directivity features, systems can be configured for interrogating defined areas. However, costs for transponders in the category are generally

higher, by a factor of 2 or more, than lower frequency devices. The antenna structures often used for microwave tags yield a directional beam, in contrast with the near spherical field patterns associated with lower frequency antenna structures. Each antenna structure will exhibit a primary forward lobe and, depending on antenna dimensions in relation to the wavelength, a number of sidelobes. Because of these features, the direction in which the antennas are deployed (directivity) in relation to a readable tag will influence the range and ability to detect the tag.

Today 2,450-MHz tags are available in many different shapes and with different functionality, influenced by applications and its requirements. Unlike inductive RFID tags, which require substantial surface area, many turns of wire, or magnetic core material to collect the magnetic field, UHF and microwave tags can be very small, requiring length in only one dimension. Thus, in addition to longer range over the inductive systems, the UHF and microwave tags are easier to package and come in a wider variety of configurations. Tag lengths of 20 to 100 mm (1 to 4 inches) are typical. The tag's thickness is limited only by the thickness of the chip, because the antenna can be fabricated on thin, flexible materials.

Today, in Europe, 5.8 GHz is used as the European road telematics frequency. It offers the advantages of being a less congested band, resulting in less interference, but it has a number of disadvantages: It is not available in the United States or many other countries, orientation of the antennas is very important, and chips are difficult and expensive to build. Microwave RFID has been effectively used in special applications where its directionality, range (particularly with active tags), and very fast data transfer features are required. These areas include asset tracking, factory automation (particularly in automotive manufacturing), and barrier-based access control.

4.5 ISO/IEC 18000 RFID Air Interface Standards

ISO/IEC 18000 is a series of standards being created by ISO/IEC/JTC for the RFID air interface for the item identification world. The ISO/IEC 18000 series of standards are currently being revised to update them and encompass new information. The update is in two main areas: an update (1) to 18000-6 to add the EPCglobal Generation 2 specification as Type C and (2) to fix any issues with Types A and B, published as an amendment to the standard. In addition, a revision to all parts of 18000 includes fixes to the standards based on actual issues discovered during the use of the standards along with the addition of the capabilities to use batteries and sensors with the existing technologies.

The standard currently contains the following parts:

- ISO/IEC 18000; Information Technology AIDC Techniques; RFID for Item Management; Air Interface;
- 18000-1 Part 1; Generic Parameters for the Air Interface for Globally Accepted Frequencies;
- 18000-2 Part 2; Parameters for Air Interface Communications below 135 kHz;
- 18000-3 Part 3; Parameters for Air Interface Communications at 13.56 MHz;
- 18000-4 Part 4; Parameters for Air Interface Communications at 2.45 GHz;
- 18000-5 Part 5; Parameters for Air Interface Communications at 5.8 GHz (withdrawn due to the lack of global interest);
- 18000-6 Part 6; Parameters for Air Interface Communications at 860 to 960 MHz;
- 18000-7 Part 7; Parameters for Air Interface Communications at 433 MHz.

As can be seen, each of these parts deals with a different aspect of RFID. The first part is the defining document that explains how the standard works, and the rest are divided by frequency.

ISO/IEC 18000-1:2004 defines the parameters to be determined in any standardized air interface definition in the ISO/IEC 18000 series. The subsequent parts of ISO/IEC 18000 provide the specific values for definition of the air interface parameters for a particular frequency or type of air interface, from which compliance (or noncompliance) with ISO/IEC 18000-1:2004 can be established. In addition, it provides descriptions of example conceptual architectures in which these air interfaces are often to be utilized. ISO/IEC 18000-1:2004 is an enabling standard that supports and promotes several RFID implementations without making conclusions about the relative technical merits of any available option for any possible application.

ISO/IEC 18000-2:2004 defines the air interface for RFID devices operating below 135 kHz used in item management applications. Its purpose is to provide a common technical specification for RFID devices to allow for compatibility and to encourage interoperability of products for the growing RFID market in the international marketplace. ISO/IEC 18000-2:2004 defines the forward and return link parameters for technical attributes including, but not limited to, operating frequency, operating channel accuracy, occupied channel bandwidth, spurious emissions, modulation, duty cycle, data coding, bit rate, bit rate accuracy, and bit transmission order. It further defines the communications protocol used in the air interface.

ISO/IEC 18000-2:2004 specifies the physical layer that is used for communication between the interrogator and the tag; the protocol and the commands; and the method to detect and communicate with one tag among several tags (anticollision). It specifies two types of tags: Type A (FDX) and Type B (HDX). These two types differ only by their physical layer. Both types support the same anticollision and protocol. FDX tags are permanently powered by the interrogator, including during the tag-to-interrogator transmission. They operate at 125 kHz. HDX tags are powered by the interrogator, except during the tag-to-interrogator transmission. They operate at 134.2 kHz. An alternative operating frequency is described, as well as an optional anticollision mechanism.

ISO/IEC 18000-3:2004 provides physical layer, collision management system, and protocol values for RFID systems for item identification in accordance with the requirements of ISO 18000-1. It relates solely to systems operating at 13.56 MHz. ISO/IEC 18000-3:2004 has two modes of operation, intended to address different applications. The modes, although not interoperable, are noninterfering.

ISO/IEC 18000-4:2004 defines the air interface for RFID devices operating in the 2.45-GHz ISM band used in item management applications. Its purpose is to provide a common technical specification for RFID devices that may be used by ISO committees developing RFID application standards. ISO/IEC 18000-4:2004 is intended to allow for compatibility and to encourage the interoperability of products for the growing international RFID market. ISO/IEC 18000-4:2004 defines the forward and return link parameters for technical attributes, including, but not limited to, operating frequency, operating channel accuracy, occupied channel bandwidth, maximum EIRP, spurious emissions, modulation, duty cycle, data coding, bit rate, bit rate accuracy, bit transmission order, and, where appropriate, operating channels, frequency hop rate, hop sequence, spreading sequence, and chip rate. It further defines the communications protocol used in the air interface. ISO/IEC 18000-4:2004 contains two modes. The first is a passive tag operating as an interrogator that talks first, whereas the second is a battery-assisted tag operating as a tag that talks first.

ISO/IEC 18000-6:2004 defines the air interface for RFID devices operating in the 860- to 960-MHz ISM band used in item management applications. Its purpose is to provide a common technical specification for RFID devices that may be used by ISO committees developing RFID application standards. ISO/IEC 18000-6:2004 is intended to allow for compatibility and to encourage interoperability of products for the growing RFID market in the international marketplace. ISO/IEC 18000-6:2004 defines the forward and return link parameters for technical attributes including, but not limited to, operating frequency, operating channel accuracy, occupied channel bandwidth, maximum EIRP, spurious emissions, modulation, duty cycle, data coding, bit rate, bit rate

accuracy, bit transmission order, and, where appropriate, operating channels, frequency hop rate, hop sequence, spreading sequence, and chip rate. It further defines the communications protocol used in the air interface. ISO/IEC 18000-6:2004 contains one mode with two types. Both types use a common return link and are reader-talks-first types. Type A uses pulse interval encoding (PIE) in the forward link and an adaptive ALOHA collision arbitration algorithm. Type B uses Manchester in the forward link and an adaptive binary tree collision arbitration algorithm.

ISO/IEC 18000-7:2004 defines the air interface for RFID devices operating as an active RF tag in the 433-MHz band used in item management applications. Its purpose is to provide a common technical specification for RFID devices that may be used by ISO committees developing RFID application standards. This standard is intended to allow for compatibility and to encourage interoperability of products for the growing RFID market in the international marketplace. ISO/IEC 18000-7:2004 defines the forward and return link parameters for technical attributes, including, but not limited to, operating frequency, operating channel accuracy, occupied channel bandwidth, maximum power, spurious emissions, modulation, duty cycle, data coding, bit rate, bit rate accuracy, bit transmission order, and, where appropriate, operating channels, frequency hop rate, hop sequence, spreading sequence, and chip rate. ISO/IEC 18000-7:2004 further defines the communications protocol used in the air interface.

There are a number of other closely related standards, such as ISO/IEC 19762-1 (-2 and -3):2005 that provide general terms and definitions in the area of automatic identification and data capture techniques.

4.6 UHF and EPCglobal Gen 2

4.6.1 The EPC Class Structure

The RFID class structure, depicted in Table 4.3, provides a framework to classify tags according to their primary functional characteristics. The RFID class structure classifies tags as belonging to one of five classes: Class 1 (identity tags), Class 2 (higher functionality tags), Class 3 (semipassive tags), Class 4 (active ad hoc tags), or Class 5 (reader tags). Each successive class within this framework builds up, that is, is a superset of, the functionality contained within the previous class, resulting in a layered functional classification structure. Class 1 forms the foundation of this framework [3]:

- The *Class 1 identity tag* is designed to be the lowest cost, minimal-usable-functionality tag classification. Identity tags are pure passive RFID tags that are expected to implement a resource discovery

Table 4.3
The EPC Class Structure

Class 1	Read-only passive identity tags, no battery
Class 2	Passive tags with additional functionality, such as memory or encryption
Class 3	Semipassive tags (battery assisted); may support broadband communication
Class 4	Active tags that may be capable of broadband, peer-to-peer communication with other active tags in the same frequency band and with readers
Class 5	Essentially reader tags; they can power other Class 1, 2, and 3 tags and also communicate with Class 4 tags and with each other wirelessly.

mechanism and store a unique object identifier only. The signaling and modulation defined for Class 1 tags are the foundation for all passive communication within this hierarchy.

- *Class 2 higher functionality tags* build on the identity tag by providing more functionality, such as a tag identifier and read/write memory, while still maintaining a pure passive power and communication scheme.
- *Class 3 semipassive tags* add an on-tag power source, such as a battery, to their higher functionality foundation. Semipassive tags combine passive communication with an on-tag power source that enables a tag to operate without the presence of a passive tag reader (i.e., a Class 5 reader tag).
- *Class 4 active ad hoc tags* encompass the Class 3 semipassive tags and, in addition, are ad hoc networking devices that are capable of communicating with other Class 4 tags, using active communication, and with Class 5 reader tags, using both passive and active communication. Because they may initiate communication, Class 4 tags are necessarily active. Functionally, these tags lie in the realm of ubiquitous computers or smart dust.
- *Class 5 reader tags* encompass the functionality of a Class 4 active ad hoc tag and, in addition, are able to power and communicate with pure passive Class 1 and Class 2 tags and communicate with Class 3 tags via passive communication.

The main technological difference between passive and semipassive (battery-assisted) labels has to do with the source from which the tags receive the appropriate amount of power in order for the IC to reach the excitation level (or to wake up) and send signals back to the reader. Passive tags gather energy from the reader's signal, whereas semipassive tags contain an integrated power source

and have no need to gather energy from the reader. This enables semipassive labels to outperform passive labels in two key areas:

- Higher reliability of up to 100% read/write rates (theoretically), even for liquids, metals, and foils;
- Increased ranges of up to 60 feet.

4.6.2 UHF Gen 2

Developed by the EPCglobal industry group, the EPC Generation 2 standard defines the physical and logical requirements for a passive-backscatter, interrogator-talks-first (ITF), RFID system operating in the 860- to 960-MHz frequency range. EPC Gen 2 is a new standard for RFID tags, specifying the operation of the tag and the communication protocol for interoperability with EPC readers worldwide (see Table 4.4). It was developed by a collaboration of leading RFID users and vendors, working through EPCglobal, a nonprofit trade group. EPCglobal is part of the UCC/EAN organization, which has long administered barcode and other standards around the world. The frequency used for UHF RFID systems varies between 860 and 960 MHz. UHF RFID systems operate at 915 MHz in the United States and 868 MHz in Europe, and they are being widely deployed due to RFID mandates from several large corporations, including international retailers, and the DoD. In addition to retail, UHF systems are employed in various supply-chain management applications.

Gen 2 (as well as its counterpart ISO 18000-6c) is a technical standard that specifies the air interface protocol, that is, how tags and readers

Table 4.4
UHF Systems Worldwide

	North America	Europe	Singapore	Japan	Korea	Australia	Argentina Brazil Peru	New Zealand
Band size [MHz]	902–928	866–868	866–869 923–925	950–956	908.5–914	918–926	902–928	864–929 (parts)
Power	4W EIRP	2W ERP	0.5W ERP (2W in upper band)	4W EIRP	2W ERP	4W EIRP	4W EIRP	0.5–4W EIRP
Number of channels	50	10	10	12	20	16	50	Varied

communicate. The leadership of EPCglobal, along with collaborative efforts throughout the industry, has now set the stage for real-world implementation of Gen 2 technologies that are compliant with the newest EPCglobal RFID specifications for the UHF band. While primarily intended for the supply chain market, the Gen 2 RFID systems may also be used in asset tracking, baggage tagging, manufacturing, and a wide assortment of other applications where a long reading range is required. The Gen 2 chip is intended for use in the manufacture of passive RFID tag products operating in the 860- to 960-MHz frequency band. So, Gen 2 addresses a very small part (although a very important part) of the total standards environment.

The Gen 2 protocol takes the best features of the Gen 1 Class 1, Gen 1 Class 2, and ISO protocols to make what promises to be a greatly improved standard. The Gen2 standard promises to be a global, open, interoperable standard with a number of very sophisticated features. In other words, Gen 2 incorporates the frequency and performance requirements for worldwide use. Some of the features of Gen 2 are as follows:

- Has a high read rate of 1,500 tags per second in North America, 600 tags per second in Europe. This is especially important in countries where the narrow bandwidth limits the data rates to 30% of that we can achieve in the United States.
- Has a proven air interface with forward link PIE ASK, backscatter link FM0 or Miller-modulated subcarrier.
- Provides an operating mode for dense reader environments.
- Provides better read algorithms (bit mask filtering) that eliminate duplicate reads, allow tags to enter the reader field late and still be read, and allow tags to stay quiet until asked to talk, making it faster to find a specific tag.
- Each tag is security enhanced with 32-bit password encryption and permanent kill capability.
- Write schemes enhance the write speed function.
- In addition to the required EPC data, there is optional memory to support user-specific data.
- Meets global regulatory compliance standards.

The Gen 2 protocol was very similar to, but not in full conformance with, the existing protocols described by ISO 18000 Part 6a and Part 6b specifications. To make Gen2 ISO compliant, EPCglobal was required to modify their Gen 2 document allowing for the optional use of an application family identifier (AFI). The AFI is a value in the data string that is used to identify a numbering

administration authority. Manufacturers of RFID tags and readers are shipping Gen 2–compliant products, and implementations have begun. Both Wal-Mart and the DoD have embraced the Gen 2 standard, and ISO has approved it for publication as ISO 18000-6c.

More information on Gen 2 can be found in the *TI UHF Gen2 Protocol, Reference Guide, 11-09-21-700,* June 2006.

In April 2007, EPCglobal announced a new industry standard providing the capability for unprecedented visibility into the movement, location, and disposition of assets, goods, and services throughout the world. Electronic Product Code Information Services (EPCIS) allows for the seamless, secure exchange of data at every point in the life cycle of goods and services. EPCIS, by providing a standard set of interfaces for EPC data, enables a single way to capture and share information, while still allowing the flexibility for industry- and organization-specific implementations. The specification supports powerful business cases and consumer benefits such as container tracking, product authentication, promotions management, baggage tracking, electronic proof of delivery, chain of custody, returns management, and operations management.

In October 2006, EPCglobal successfully completed interoperability testing of the platform along with 12 other large and small solution providers from Japan, Korea, and North America, including Auto-ID Labs, Avicon, BEA Systems, Bent Systems, IBM, Globe Ranger, IIJ, NEC, Oracle, Polaris Systems, Samsung, and T3Ci. The interoperability test marked a significant milestone in the development of EPCIS, which is the result of years of effort by more than 150 companies and organizations participating in the EPCIS working group. The positive results of this test and solution provider support have led to the ratification of this standard.

4.6.3 UHF RFID Tag Example

The EPC tag class structure is often misunderstood. *Class* is not the same as *Generation. Class* describes a tag's basic functionality—for example, whether it has memory or a battery—whereas *generation* refers to a tag specification's major release or version number. The full name for what is popularly called EPC Generation 2 is actually EPC Class 1 Generation 2, indicating that the specification refers to the second major release of a specification for a tag with write-once memory. The example shown in Figure 4.3 and description of the UHF RFID tag were taken from [4].

Antenna performance at UHF and microwave frequencies are dependent on the substrate, that is, thickness and electromagnetic properties (conductivity, permittivity, and permeability). It also depends on the conductive quality of electrodes. Low-cost constraints and the diversity of RFID applications have led researchers to consider and investigate nonstandard materials to be used for both

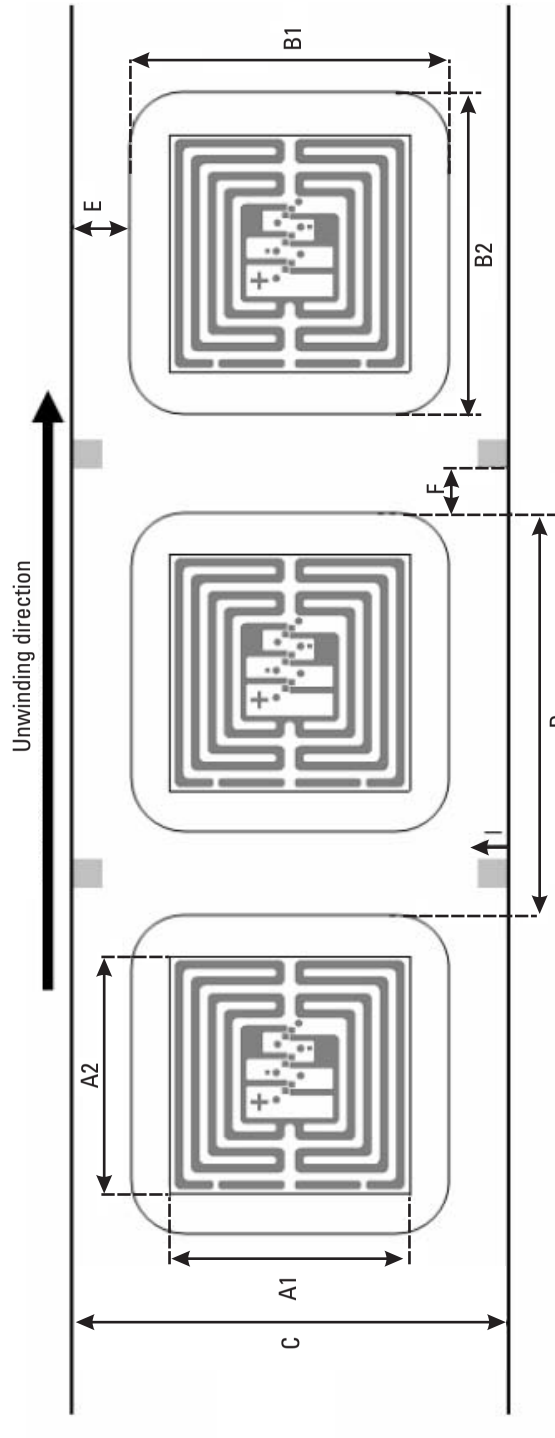


Figure 4.3 UHF RFID tag layout.

tags and antennas. These investigations concern low-cost substrate materials, such as paper, plastic, and polymers (Figure 4.4).

The use of conductive inks is an alternative to the usual electrodes made with standard conductors such as copper and aluminum. The performance of an antenna made with conductive ink is limited by the conductivity and the thickness of the deposited ink. To avoid excess loss and get good directivity, the thickness of the ink must be larger than skin depth (which is dependent on the conductivity and the frequency) [5].

The tag antenna is generally omnidirectional to ensure identification in all directions. The structure of the tag antenna should also be as small as possible in size. Because of its simplicity and omnidirectionality, the $\lambda/2$ dipole is one of the most preferred forms. At UHF frequencies the typical size is 150 mm (6 inches), which is big. Usually the dipole is folded in order to reduce its size. This usually requires full-wave electromagnetic simulation in order to take into account the capacitive and inductive coupling introduced by the folded form. Some typical specification parameters are shown in Table 4.5.

4.7 Review Questions and Problems

1. Why is standardization very important in the RFID world? What are the most important standardization organizations in this arena?
2. What is the current status of harmonization of the UHF RFID standards?
3. What is the most widely used RFID frequency/system today? Explain why.
4. Explore the latest utilization and standardization status of the 2.4- and 5.8-GHz bands RFID systems in the United States and the rest of the world. Why are microwave bands not more popular in RFID applications?
5. An antenna has a gain of 16 dBi, and the power delivered to the antenna is 100 mW. What is the effective isotropic radiated power in dBm and in watts? (*Answer:* 36 dBm/4W.)

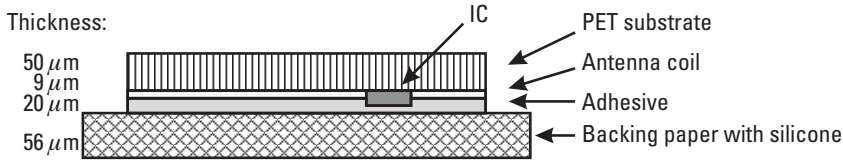


Figure 4.4 UHF RFID tag cross section.

Table 4.5
UHF RFID Specifications

Integrated circuit	96-bit EPC Class 1 Gen 2
Free-air frequency	915 ± 15 MHz, loaded mode
Read sensitivity	Minimum 7.8 V/m
Operating temperature (electronics parts)	–40°C/+65°C
Thermal-cycle resistance (electronics parts)	200 cycles –40°C/+80°C
Temperature-humidity resistance (electronics parts)	80°C, 85% RH, 168h
ESD voltage immunity	+/– 1-kV peak, HBM
Storage	(15°–25°C, 40%–60% RH, maximum 2 years
Bending diameter (*D*)	> 50 mm, Tension less than 10N
Static pressure (*P*)	<10 MPa (10 N/mm^2)

6. What frequency is used by UHF RFID devices in the United States?

   a. 902–928 MHz
   b. 121–124 kHz
   c. 2.4 GHz
   d. 915–928 MHz

7. The Spanish Post Office has implemented Europe's largest UHF RFID system in sorting centers in 16 cities across Spain. Reusable tags are inserted into an envelope and sent through the system to monitor the movement of letters, as well as the system's real-time performance. How would you conceptually (without going into details) design a system like that in the country in which you live?

8. You want an RFID tag that supports longer distance communications and does not rely on the reader to provide power to the tag. What kind of tag do you need?

   a. Passive tag
   b. Semipassive tag
   c. Active tag
   d. Powered tag

9. Discuss briefly the pros and cons of active, passive, and semipassive tags.

References

[1] "Radio Frequency Identification—Opportunities and Challenges in Implementation," Washington, D.C., U.S. Department of Commerce, 2005.

[2] Loy, M., et al., "ISM-Band and Short Range Device Regulatory Compliance Overview," Application Report SWRA048, Texas Instruments, May 2005.

[3] Engals, D. W., and S. E. Sarma, "Standardization Requirements within the RFID Class Structure Framework," Cambridge, MA: Auto-ID Labs, Massachusetts Institute of Technology, January 2005.

[4] http://www.upmraflatac.com, accessed October 13, 2006.

[5] Tedjini, S., et al., "Antennas for RFID Tags," *Joint sOc-EUSAI Conference,* Grenoble, France, October 2005.

5

Components of the RFID System

5.1 Engineering Challenges

An RFID system consists of an RFID reader, RFID tag, and information managing host computer. The reader contains an RF transceiver module (transmitter and receiver), a signal processor and controller unit, a coupling element (antenna), and a serial data interface (RS232, RS485) to a host system. The tag acts as a programmable data-carrying device and consists of a coupling element (resonant tuned circuit) and a low-power CMOS IC. The IC chip contains an analog RF interface, antenna tuning capacitor, RF-to-dc rectifier system, digital control and electrically erasable and programmable read-only memory (EEPROM), and data modulation circuits. RFID involves contactless reading and writing of data into an RFID tag's nonvolatile memory through an RF signal. The reader emits an RF signal and data is exchanged when the tag comes in proximity to the reader signal. Tags can be categorized as follows:

1. Active tag, which has a battery that supplies power to all functions;
2. Semipassive tag, which has a battery used only to power the tag IC, and not for communication;
3. Passive tag, which has no battery on it. The absence of a power supply makes passive tags much cheaper and more reliable than active tags.

Given the increase in RFID usage, many new challenges face design engineers. Currently, these challenges include multiple tag standards, 20% tag failure rate, installation and placement issues, the need for cost-effective

management and maintenance of readers, the need for reductions in the reader size that allows them to be imbedded into structures and handheld devices, and intellectual property protection and secure access control protocols. A number of different parameters will influence the quality and reliability of the RFID system: tag size, reader/writer antenna size, tag orientation, tag operating time, tag movement velocity, effect of metallic substances on operating range, multiple-tag operating characteristics, and the effect of the number of tags on operating success rate, tag overlapping, and so forth.

5.2 Near- and Far-Field Propagation

RFID systems on the market today fall into two main categories: *near-field* systems that employ inductive (magnetic) coupling of the transponder tag to the reactive energy circulating around the reader antenna, and *far-field* systems that couple to the real power contained in free space propagating electromagnetic plane waves [1]. Near-field coupling techniques are generally applied to RFID systems operating in the LF and HF bands with relatively short reading distances, whereas far-field coupling is applicable to the potentially longer reading ranges of UHF and microwave RFID systems. Whether or not a tag is in the near or far field depends on how close it is to the field creation system and the operating frequency or wavelength. There is a distance, commonly known as the *radian sphere,* inside which one is said to be in the near field and outside of which one is said to be in the far field. Because changes in electromagnetic fields occur gradually, the boundary is not exactly defined; the primary magnetic field begins at the antenna and induces electric field lines in space (the *near field*).

The zone where the electromagnetic field separates from the antenna and propagates into free space as a plane wave is called the *far field*. In the far field, the ratio of electric field E to magnetic field H has the constant value of 120π or 377Ω. The approximate distance where this transition zone happens is given as follows:

$$r = \frac{\lambda}{2\pi} \tag{5.1}$$

It is also important to notice that this expression is valid for small antennas where $D << \lambda$.

The reactive near-field region is a region where the E- and H-fields are not orthogonal; anything within this region will couple with the antenna and distort the pattern, so the antenna gain is not a meaningful parameter here. Using (5.1), at 13.56 MHz ($\lambda = 22$m), this places the near-field–far-field boundary at about 3.5m (10 feet).

It has been estimated that the far-field distance for the case in which $D > \lambda$ is given as follows:

$$r = \frac{2D^2}{\lambda} \tag{5.2}$$

where D is the maximum dimension of the radiating structure and r is the distance from the antenna. Note that this is only an estimate, and the transition from near field to far field is not abrupt. Typically D for reader antennas is 0.3m (1 foot.) The far-field distance in the UHF ISM band in the United States (915 MHz, $\lambda = 0.33$m) is estimated to be 0.56m.

Generally speaking, the radiating near-field or transition region is defined as a region between the reactive near field and a far field. In this region, the antenna pattern is taking shape but is not fully formed, and the antenna gain measurements will vary with distance:

$$\frac{\lambda}{2\pi} < r < \frac{2D^2}{\lambda} \tag{5.3}$$

The solution of Maxwell's equations for the fields around an antenna consists of three different powers of the range $1/r$, $1/r^2$, and $1/r^3$. At very short ranges, the higher powers dominate the solution, while the first power dominates at longer ranges. This can be interpreted as the electromagnetic wave breaking free from the antenna. The near field may be thought of as the transition point where the laws of optics must be replaced by Maxwell's equations of electromagnetism.

5.2.1 Far-Field Propagation and Backscatter Principle

RFID systems based on UHF and higher frequencies use *far-field communication* and the physical property of backscattering or "reflected" power. Far-field communication is based on electric radio waves where the reader sends a continuous base signal frequency that is reflected back by the tag's antenna. During the process, the tag encodes the signal to be reflected with the information from the tag (the ID) using a technique called *modulation* (i.e., shifting the amplitude or phase of the waves returned) [2].

The concept of the radian sphere, which has a value for its radius of $\lambda/2\pi$ helps in the visualization of whether the tag coupling is in the near or far field. If the tag is inside this sphere, the reactive energy storage fields (dipolar field terms) dominate and near-field coupling volume theory is used. If the tag falls outside the sphere, then propagating plane wave EM fields dominate and the familiar antenna engineering concepts of gain, effective area or aperture, and

EIRP are used. These often more familiar EM concepts whereby real power is radiated into free space are relevant to the cases of UHF and microwave tagging technologies.

Most theoretical analyses, at least in the first approximation, assume the so-called free-space propagation. *Free space* simply means that there is no material or other physical phenomenon present except the phenomenon under consideration. Free space is considered the baseline state of the electromagnetic field. Radiant energy propagates through free space in the form of electromagnetic waves, such as radio waves and visible light (among other electromagnetic spectrum frequencies). Of course, this model rarely describes the actual propagation situation accurately; phenomena such as reflection, diffraction, and scattering exist that disturb radio propagation. In the wireless industry, most models and formulas we use today are semiempirical, that is, based on the well-known radio propagation laws but modified with certain factors and coefficients derived from field experience. RFID is definitely an area where this practice is required; short distances cluttered with multiple tags and/or other objects are potential obstacles to radio propagation and will cause serious deviations, predictable or not, from the theoretical calculations.

A backscatter tag operates by modulating the electronics connected to the antenna in order to control the reflection of incident electromagnetic energy. For successful reading of a passive tag, two physical requirements must be met:

1. *Forward power transfer:* Sufficient power must be transferred into the tag to energize the circuitry inside. The power transferred will be proportional to the second power of the distance.
2. *The radar equation:* The reader must be able to detect and resolve the small fraction of energy returned to it. The power received will be reduced proportional to the fourth power of the distance.

5.2.1.1 Forward Power Transfer

A typical RFID tag consists of an antenna and an integrated circuit (chip), both with complex impedances. The chip obtains power from the RF signal transmitted by the base station, called the RFID reader. The RFID tag antenna is loaded with the chip whose impedance switches between two impedance states, usually high and low. At each impedance state, the RFID tag presents a certain *radar cross section* (RCS). The tag sends the information back by varying its input impedance and thus modulating the backscattered signal.

In Figure 5.1, $Z_A = R_a + jX_a$ is the complex antenna impedance and $Z_C = R_c + jX_c$ is the complex chip (load) impedance; chip impedance may vary with the frequency and the input power to the chip. The power scattered back from the loaded antenna can be divided into two parts. One part is called the *structural*

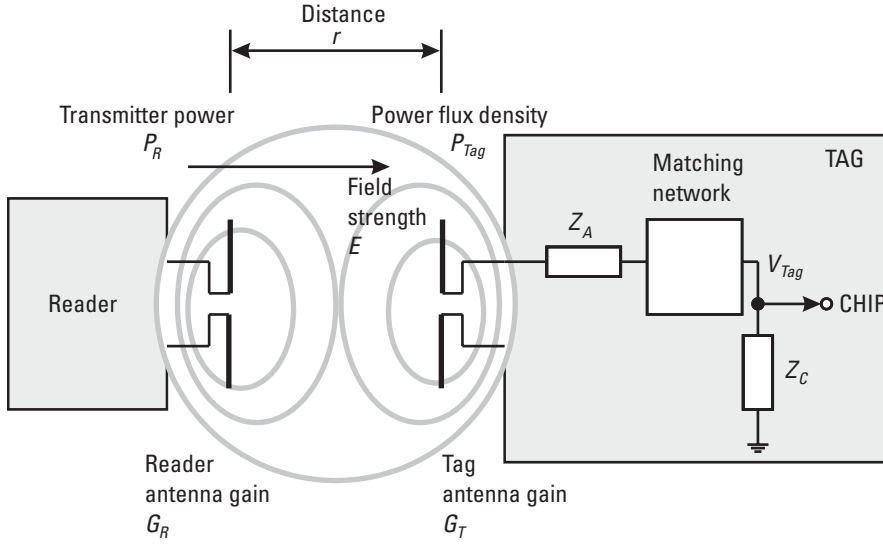


Figure 5.1 Forward power transfer.

mode and is due to currents induced on the antenna when it is terminated with complex conjugate impedance. The second part is called the *antenna mode* and results from the mismatch between antenna impedance and load impedance.

The separation between the antennas is r, which is assumed to be large enough for the tag to be in the far field of the reader. E is the electric field strength of the reader at the tag location. The efficiency of the matching network will be taken as unity and ignored (losses in the network may also be accounted for in the value of G_T). Antenna gains G_R and G_T are expressed relative to an isotropic antenna. From considerations of power flux density at the tag, with λ as the wavelength, we get:

$$P_{Tag} = \left(E^2/120\pi\right)\left(\lambda^2/4\pi\right)G_T = V_{tag}^2/R_c \tag{5.4}$$

and

$$E^2/120\pi = P_R G_R/4\pi r^2 \tag{5.5}$$

After some manipulation of these equations, we obtain:

$$P_{Tag} = \left(P_R G_R/4\pi r^2\right)\left(\lambda^2 G_T/4\pi\right) = P_R G_R G_T \lambda^2/(4\pi)^2 r^2 \tag{5.6}$$

The typical maximum reader output power is 500 mW, 2W (ERP, CEPT), and 4W (EIRP, FCC). Converted to dBm, the permitted maximum limits are about 29 dBm (500 mW ERP, 825 mW EIRP), 35 dBm (2W ERP, 3.3W EIRP), and 36 dBm (4W EIRP). The gain of the transmitter (reader) antenna (typical value) is assumed to be 6 dBi. Therefore, the maximum output power from the power amplifier should be 23, 29, and 30 dBm, respectively. The tag available power versus distance can be seen in Figure 5.2. From the industrial experience, the minimum RF input power of 10 μW (−20 dBm) to 50 μW (−13 dBm) is required to power on the tag. The power received by the tag is then divided in two parts: the reflected power and the available power used by the chip. The distribution of these two parts is very critical for a maximum distance. For dipole antennas presented in the best orientation, G_T may be taken as 2 dBi (gain over isotropic with allowance for losses, approximately 1.6).

We can also say that:

$$V_{Tag} = (\lambda/4\pi r)\sqrt{P_R G_R G_T R_c} \tag{5.7}$$

Note that $P_R G_R$ is the EIRP of the reader. The maximum practical value of R_c is 600Ω. The received voltage V_{Tag} must be large enough to be rectified and power the tag; a voltage in excess of 1.2 V_{rms} may be required. This is with the tag presented to the interrogating field in the ideal orientation and with no power margin.

Using (5.9), at 915 MHz (λ = 0.33m), for example, it can be seen that with 1.6 V_{rms} tag voltage (assuming both the gain of the reader and the tag to be

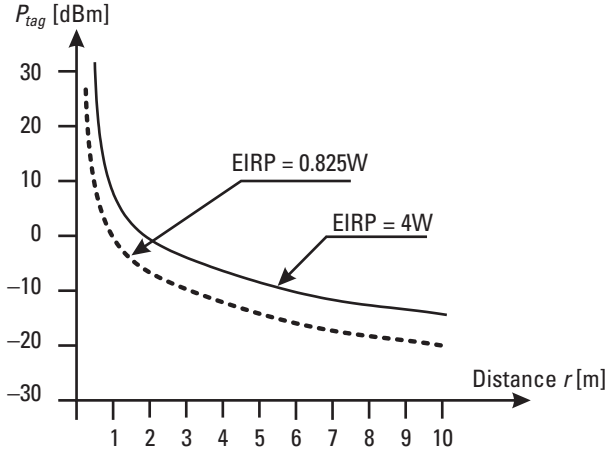


Figure 5.2 Tag-received power versus distance.

2 dBi), the required reader power at 1m distance is 2.416W or EIRP reader power of 3.96W.

$$P_{Tag} = V_{Tag}^2 / R_c = 1.6^2 / 600 = 0.0043\,[\mathrm{W}] \tag{5.8}$$

$$\begin{aligned} P_R &= \left(4\pi r V_{Tag}/\lambda\right)^2 \left(1/\left(G_R G_T G_c\right)\right) \\ P_R &= \left(4\pi \cdot 1 \cdot 1.6/0.33\right)^2 \left(1/(1.6 \cdot 1.6 \cdot 600)\right) \\ P_R &= 2.416\,[\mathrm{W}] \end{aligned} \tag{5.9}$$

$$P_{\mathrm{REIRP}} = P_R \cdot 1.64 = 2.416 \cdot 1.64 = 3.96\,[\mathrm{W}] \tag{5.10}$$

The relationship between the electrical field strength and the power flux density is the same as that between the voltage and the power in an electrical circuit; from (5.4) we can say that the electric field strength of the reader at the tag position ($P_R G_R$ is the EIRP of the reader) is equal to:

$$\begin{aligned} E &= \sqrt{30 \cdot P_R G_R}/r \\ E &= \sqrt{30 \cdot 3.96}/1 = 10.9\,[V/m] \end{aligned} \tag{5.11}$$

Note that the gain of 2 dBi is approximately equivalent to the gain of 1.6 and can be calculated as follows:

$$G[\mathrm{dBi}] = 10 \log G \rightarrow G = 10^{\frac{G[\mathrm{dBi}]}{10}} \tag{5.12}$$

5.2.1.2 The Radar Equation

Radar principles tell us that the amount of energy reflected by an object is dependent on the reflective area of the object—the larger the area, the greater the reflection. This property is referred to as the radar cross section (RCS). The RCS is an equivalent area from which energy is collected by the target and retransmitted (backscattered) back to the source. For an RFID system in which the tag changes its reflectivity in order to convey its stored identity and data to the reader, this is referred to as *differential radar cross section* or ΔRCS. Calculations of the complete return signal path are conveniently conducted in terms of the ΔRCS of the backscatter device.

For the antenna to transfer maximum energy to the chip, the impedance of the chip must be a conjugate of the antenna impedance. However, it is important to remember that the logic circuits of a chip used in a tag draw very little power relative to the amount of power consumed by the chip RF input

circuits. As the modulator switches between two states, the load impedance of the chip Z_C will switch between two states. The reflection due to a mismatch between antenna and load in a backscatter tag is analogous to the reflection found in transmission lines and may be expressed in terms of a coefficient of reflection. The coefficient of reflection ρ will therefore change as the modulator switches between two states. When the tag modulator is in the off state, the chip input impedance will be closely matched to the antenna impedance; therefore, the reflectivity will be low and hence the SWR will approach 1 (5.13). When the modulator is in the on state, the tag antenna impedance will be mismatched and so the reflectivity will be high, and the SWR will tend to infinity, causing the maximum amount of power to be reflected:

$$\rho = \frac{Z_C - Z_A^*}{Z_C + Z_A} \tag{5.13}$$

The tag varies its RCS by changing the impedance match of the tag antenna between two (or more) states. The ratio between the states is called the *differential coefficient of reflectivity,* represented by the symbol $\Delta\rho$, and it can be calculated using well-known transmission line theory, a summary of which is provided next. Signal propagation follows the well-known Friis transmission formula; analytical approaches such as the Friis equation assume undisturbed near-field conditions (i.e., no proximity of dielectric and metal objects), known antenna characteristics, and no diffraction and reflection effects.

An antenna of gain G has an effective aperture as calculated here:

$$A_e = \lambda^2 G_T / 4\pi \ \left[m^2\right] \tag{5.14}$$

The $\Delta\rho$ is the differential reflection coefficient of the tag modulating circuitry and can be calculated as shown:

$$\Delta\rho = p_1\left(1 - |\rho_1|^2\right) + p_2\left(1 - |\rho_2|^2\right) \tag{5.15}$$

where the IC in states 1 and 2 for a fraction of and of time, p_1 and p_2 of time, respectively [3].

It is worth mentioning that if the tag modulator switches from a perfectly matched state ($\rho_2 = 0$) to a short circuit state or to an open circuit state ($\rho_1 = 1$), $\Delta\rho$ will be approximately 0.5. Lower (and more realistic) modulation ratios will result in $\Delta\rho < 0.5$. However, in those cases where the modulator switches from a state where the chip has an impedance higher than the antenna impedance, to a condition where it is lower than the antenna impedance, $\Delta\rho$ will represent the

difference between the two states, in which case $\Delta\rho$ could be greater than 0.5 (but never higher than 1).

In modulation schemes, where one of the two states is active most of the time (e.g., $p_1 < p_2$), this is a good choice in terms of power efficiency, but these schemes require a much larger bandwidth (due to the short gaps), which is often prohibited by national authorities' regulations. Assuming that both states are active an equal amount of time (as it is in ASK with total mismatch in one state), that is, $p_1 = p_2 = 0.5$, and assuming there are no antenna losses, 50% of the available input power is actually available for rectification, 25% is used as backscattered modulated power, and the remaining 25% is wasted.

The σ is the ΔRCS of the tag and P_{Ret} is the power returned to the reader. The ΔRCS of the tag antenna is equivalent to the antenna effective aperture A_e, when the tag is matched. For the ΔRCS) when the tag antenna is mismatched, theoretically speaking, we can say that:

$$\sigma = \Delta_{\mathrm{RCS}} = A_e \cdot G_T \cdot (\Delta\rho)^2 = \frac{\lambda^2 G_T^2 (\Delta\rho)^2}{4\pi} \left[m^2\right] \tag{5.16}$$

where G is the gain of the tag antenna (G is squared because the signal is received and reradiated), λ is the wavelength, and ρ is the differential reflection coefficient of the tag modulator.

First we assume that electromagnetic waves propagate under ideal conditions, that is, without dispersion. If high-frequency energy is emitted by an isotropic radiator, then the energy propagates uniformly in all directions. Areas with the same power density therefore form spheres ($A = 4\pi r^2$) around the radiator (Figure 5.3). The same amount of energy spreads out on an incremented spherical surface at an incremented spherical radius. That means that the power density on the surface of a sphere is inversely proportional to the radius of the sphere.

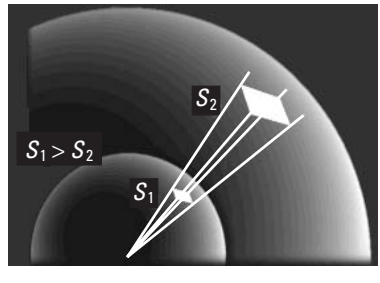


Figure 5.3 Illustration of the power-flux density.

So we obtain the formula for calculating the nondirectional power-flux density:

$$S = \frac{P_R}{4\pi r^2} \left[W/m^2\right] \tag{5.17}$$

where P_R is the power transmitted from the reader.

Because a spherical segment emits equal radiation in all directions (at constant transmitting power), if the power radiated is redistributed to provide more radiation in one direction, an increase of the power density in direction of the radiation results. This effect is called *antenna gain* and it is obtained by directional radiation of the power. So, from the definition, the directional power flux density is:

$$S_D = S \cdot G_R \tag{5.18}$$

The target (tag in our case) detection is not only dependent on the power density at the tag's position, but also on how much power is reflected in the direction of the radar (reader in our case). To determine the useful reflected power, it is necessary to know the radar cross section σ. This quantity depends on several factors, but it is true to say that a bigger area reflects more power than a smaller area. That means that a jumbo jet offers more RCS than a sporting aircraft in the same flight situation. Beyond this, the reflecting area depends on design, surface composition, and materials used. With this in mind, we can say that the returned (reflected) power P_{Ret} toward the RFID reader depends on the power density S_D, the reader's antenna gain G_R, and the variable RCS σ:

$$P_{Ret} = \frac{P_R}{4\pi r^2} \cdot G_R \cdot \sigma \left[W\right] \tag{5.19}$$

Because the reflected signal encounters the same conditions as the transmitted power, the power density yielded at the receiver of the reader is given by:

$$S_{REC} = \frac{P_{Ret}}{4\pi r^2} = \frac{P_R \cdot G_R}{(4\pi)^2 r^4} \cdot \sigma \tag{5.20}$$

The backscatter communication radio link budget, a modification of the monostatic radar equation, describes the amount of modulated power that is scattered from the RF tag to the reader (5.21):

$$P_{REC} = S_{REC} \cdot A_e = \frac{P_R \cdot G_R \cdot \sigma}{(4\pi)^2 r^4} \cdot \frac{\lambda^2 \cdot G_R}{4\pi} = \frac{P_R G_R^2 \lambda^2 \sigma}{(4\pi)^3 r^4} \tag{5.21}$$

For successful operation, we require both that the signal at the reader's receiver be above the noise floor and that the ratio of the power received and transmitted from the reader not be too small. The spreadsheet shown in Table 5.1 shows the return signal ratio for various situations. Ratios below 100 dB are both manageable in terms of signal processing and ensure that the return signal is significantly above the thermal noise floor.

Noise is the major limiting factor in communications system performance. Noise can be divided into four categories: thermal noise, intermodulation noise, crosstalk, and impulse noise. For this analysis, we consider only thermal noise and neglect other potential sources of noise. Now, we have to calculate the power of the reflected signal at the receiver of the reader and compare it with the thermal noise threshold.

Thermal noise results from thermal agitation of electrons; it is present in all electronic devices and transmission media and is a function of temperature and the channel bandwidth. Thermal noise is independent of any specific frequency. Thus, the thermal noise power in watts present in a bandwidth of B hertz can be expressed as shown here:

$$N = kTB \tag{5.22}$$

where Boltzmann's constant $k = 1.3803 \times 10^{-23}$ J/K and T is the temperature in kelvin ($T = 273.16 + t$ [°C]).

In dBW, (5.22) would look like this:

$$N = -228.6 + 10 \log T + 10 \log B \tag{5.23}$$

It is also possible to define rms noise voltage across some resistance R by applying Ohm's law to (5.22):

$$E_N = \sqrt{4RkTB} \tag{5.24}$$

The *noise figure* or *noise factor* (NF) is a contribution of the device itself to thermal noise. It is commonly defined as the signal-to-noise ratio at the input divided by the signal-to-noise ratio at the output and is usually expressed in decibels. Typical noise figures range from 0.5 dB for very low noise devices, to 4 to 8 dB.

Table 5.1
Received UHF RFID Power for Various Distances

f [MHz]	λ [m]	*r* [m]	*P* Reader [W]	*P* Reader [dBm]	Reader Antenna Gain	Tag Antenna Gain	$\Delta\rho$	σ [m^2]	*P* Received [μW]	*P* Received [dBm]	Power Ratio [dB]
915.00	0.33	1.00	2.00	33.01	1.60	1.60	0.50	0.0055	1.5185	–28.19	–61.20
915.00	0.33	2.00	2.00	33.01	1.60	1.60	0.50	0.0055	0.0949	–40.23	–73.24
915.00	0.33	4.00	2.00	33.01	1.60	1.60	0.50	0.0055	0.0059	–52.27	–85.28
915.00	0.33	8.00	2.00	33.01	1.60	1.60	0.50	0.0055	0.0004	–64.31	–97.32
433.00	0.69	1.00	2.00	33.01	1.60	1.60	0.50	0.0244	30.2793	–15.19	–48.20
433.00	0.69	2.00	2.00	33.01	1.60	1.60	0.50	0.0244	1.8925	–27.23	–60.24
433.00	0.69	4.00	2.00	33.01	1.60	1.60	0.50	0.0244	0.1183	–39.27	–72.28
433.00	0.69	8.00	2.00	33.01	1.60	1.60	0.50	0.0244	0.0074	–51.31	–84.32

For a 500-kHz bandwidth, at room temperature, we can use (5.22) to calculate a thermal noise of −117 dBm; with addition of the 3-dB receiver noise factor, total noise is −114 dBm, leaving enough reader signal margin to the noise threshold.

We can see that the return power ratio conditions are met at relatively longer ranges and that forward power transfer is the limiting factor in UHF backscatter tags. Battery-powered backscatter tags overcome this limitation and can be read at significantly greater ranges.

5.2.2 Near-Field Propagation Systems

5.2.2.1 Magnetic Field Calculations

At low to mid-RFID frequencies, RFID systems make use of near-field communication and the physical property of inductive coupling from a magnetic field. The reader creates a magnetic field between the reader and the tag and this induces an electric current in the tag's antenna, which is used to power the integrated circuit and obtain the ID. The ID is communicated back to the reader by varying the load on the antenna's coil, which changes the current drawn on the reader's communication coil. In the near field, it is possible to have an electric field with very little magnetic field, or magnetic field with very little electric field. The choice between these two alternatives is determined by the design of the interrogation antenna, and RFID systems are generally designed to minimize any incidental electric field generation.

In the near field, the magnetic field strength attenuates according to the relationship $1/r^3$, that is, the magnetic field intensity decays rapidly as the inverse cube of the distance between the reader antenna and the tag. In power terms, this equates to a drastic $1/r^6$ reduction with distance (60 dB/decade) of the available power to energize the tag. The magnetic field strength is thus high in the immediate vicinity of the transmitting coil, but a very low level exists in the distant far field; hence a spatially well-confined interrogation region or localized tag-reading zone is created. Note that magnetic loop reader antennas can also be designed that exhibit good electrical symmetry and balance to eliminate stray electric E-field pickup.

The tag's ability to efficiently draw energy from the reader field is based on the well-known electrical resonance effect. The coupling or antenna element of the tag is really an inductor coil and capacitor connected together and designed to resonate at the 13.56-MHz system operating frequency (Figure 5.4). The current passing through the inductor creates a surrounding magnetic field according to Ampere's law. The created magnetic field B is not a propagating wave [4], but rather an attenuating carrier wave, with its strength given as illustrated in Figure 5.5 and described by formula (5.25):

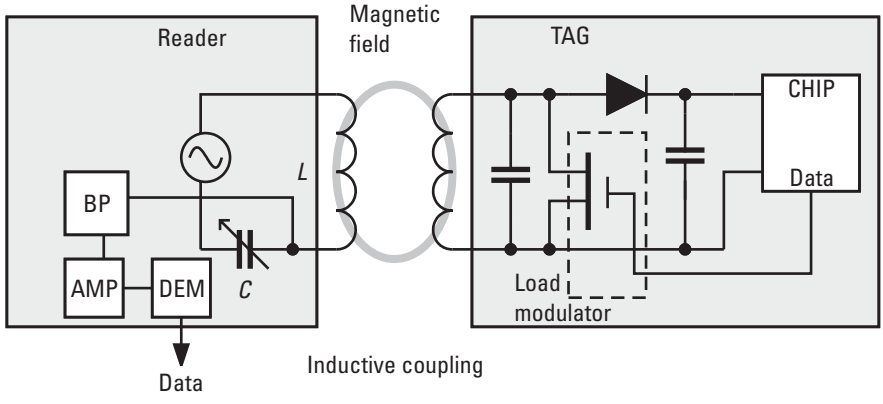


Figure 5.4 Principle of inductive (near-field) coupling.

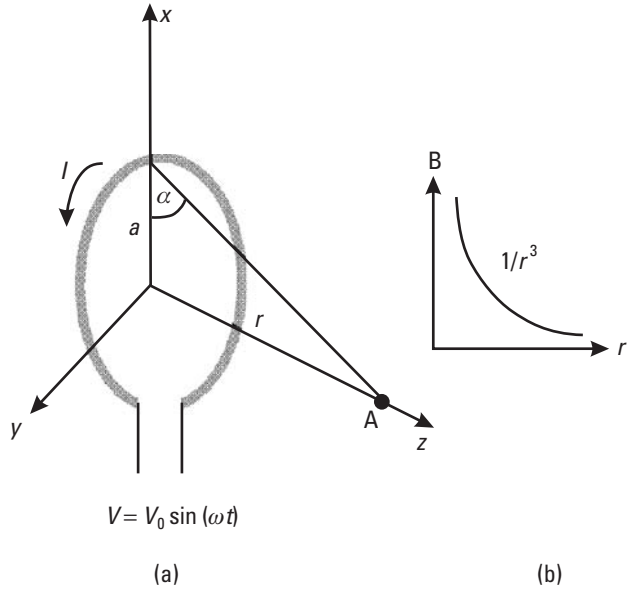


Figure 5.5 Calculation of the magnetic field away from the coil: (a) coil in 3D space, and (b) magnetic field decay with distance.

$$B = \frac{\mu_0 INa^2}{2r^3}\left[\text{Weber/m}^2 \text{ or tesla}\right] \tag{5.25}$$

where:

I = current through the coil;

N = number of windings in the coil;

a = radius of the coil;

μ_o = permeability of free space ($4\pi \times 10^{-7}$ H/m);

r = perpendicular distance from antenna to point A and $r >> a$.

As one moves away from the source with $r >> a$, the simplified (5.25) shows the characteristic $1/r^3$ attenuation. This near-field decaying behavior of the magnetic field is the main limiting factor in the read range of an RFID device.

We use Ohm's law for ac circuits:

$$I = \frac{V}{Z_L} = \frac{V}{\omega L} \tag{5.26}$$

and assume that L can be approximated as follows:

$$L \approx \mu_0 \pi a N^2 \tag{5.27}$$

We can then rewrite (5.25) as shown here:

$$B = \frac{Va}{2\omega N \pi r^3} \tag{5.28}$$

From (5.28) with a given coil voltage at some distance from the coil, we can now see that B is inversely proportional to N. This is due to the fact that the current increases at the rate of $1/N^2$ with a given coil voltage V. Only the case of an air-coiled inductor has been described, but a ferrite-cored inductor could be used as well. Adding a core has the effect of increasing the effective surface area, enabling one to reduce the physical size of the coil.

To maximize the magnetic field, given fixed antenna dimensions, (5.25) dictates that the current delivered to the antenna must be maximized. Additionally, to maximize current, the antenna must resonate at the excitation frequency provided by the reader circuit. Resonance frequency (f_0) of the reader is determined by the inductance (L) of the antenna (determined by the radius of the coil, the number of windings, the thickness of the windings, and the length of the coil) and a tuning capacitor (C) and is calculated as follows:

$$f_0 = 1/2\pi\sqrt{LC} \tag{5.29}$$

The same formula is used to calculate a tag's resonant frequency, which is determined by choosing the inductive and total capacitive values, so that the

value for the tag's resonant frequency f_0 is achieved. In the case of a tag's resonant frequency:

$L[\mathrm{H}]$ = inductance of tag antenna coil
$C[\mathrm{F}]$ = capacitance of a tag's tuning capacitor

$$Z(j\omega) = R + j(X_L - X_C) \tag{5.30}$$

In practice, when the tuned circuit is resonating, the sum of its capacitive and inductive reactance is zero ($X_L = X_C$), and the impedance shown in (5.30) becomes purely resistive.

Total resistance is thereby minimized and current through the antenna is maximized, yielding a maximized magnetic field strength. Passive tags utilize the energy provided by the carrier wave through an induced antenna-coil voltage. The voltage is proportional to the product of the number of turns in the tag antenna and the total magnetic flux through the antenna. The ASIC within the tag must receive a minimum voltage (threshold voltage or power) to operate.

5.2.2.2 Voltages Induced in Antenna Circuits

Faraday's law states that a time-varying magnetic field through a surface bounded by a closed path induces a voltage around the loop. Figure 5.6 shows a simple geometry for an RFID application. When the tag and reader antennas are in proximity, the time-varying magnetic field B that is produced by a reader antenna coil induces a voltage (called electromotive force or simply EMF) in the closed tag antenna coil. The induced voltage in the coil causes a flow of current on the coil. The induced voltage on the tag antenna coil is equal to the time rate of change of the magnetic flux Ψ:

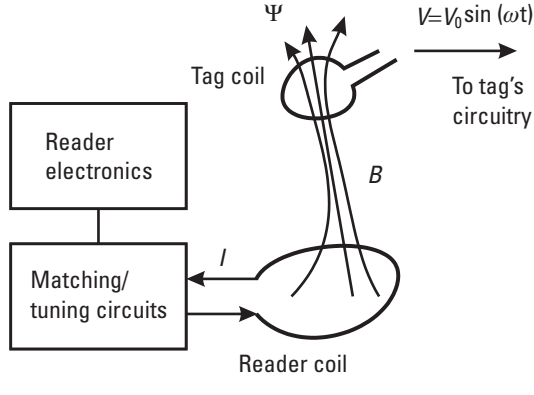


Figure 5.6 Basic reader and tag configuration.

$$V = -N\frac{d\Psi}{dt} \tag{5.31}$$

where N is the number of turns in the antenna coil and Ψ is the magnetic flux through each turn. The negative sign indicates that the induced voltage acts in such a way as to oppose the magnetic flux producing it. This is known as Lenz's law, and it emphasizes the fact that the direction of current flow in the circuit is such that the induced magnetic field produced by the induced current will oppose the original magnetic field. The magnetic flux Ψ in (5.31) is the total magnetic field B that is passing through the entire surface of the antenna coil, and it is found by:

$$\Psi = \int B \bullet dS \tag{5.32}$$

where:

B = magnetic field given in (5.21).

S = surface area of the coil.

• = inner product (cosine angle between two vectors) of vectors B and surface area S. Both magnetic field B and surface S are vector quantities.

The presentation of the inner product of two vectors suggests that the total magnetic flux Ψ that is passing through the antenna coil is affected by the orientation of the antenna coils. The inner product of two vectors becomes minimized when the cosine angle between the two is 90°, or the two (B field and the surface of coil) are perpendicular to each other and maximized when the cosine angle is 0°. The maximum magnetic flux that is passing through the tag coil is obtained when the two coils (reader coil and tag coil) are placed in parallel with respect to each other. This condition results in maximum induced voltage in the tag coil and also maximum read range. The inner product expression also can be expressed in terms of a mutual coupling between the reader and tag coils. The mutual coupling between the two coils is maximized in the preceding condition.

Combining expressions given so far, the voltage across the tag antenna, at the resonant frequency, can be calculated as in the following equation:

$$V_{Tag} = 2\pi fNQB(S\cos\alpha) \tag{5.33}$$

where:

f = frequency of the carrier signal;

S = area of the coil in square meters;

Q = quality factor of the resonant circuit;

B = strength of the magnetic field at the tag;

α = angle of the field normal to the tag area.

The ($S \cos \alpha$) term in (5.33) represents an effective surface area of the antenna that is defined as an exposed area of the loop to the incoming magnetic field. The effective antenna surface area is maximized when $\cos \alpha$ becomes unity ($\alpha = 0°$), which occurs when the antennas of the base station and the transponder units are positioned in a face-to-face arrangement. In practical applications, the user might notice the longest detection range when the two antennas are facing each other and the shortest range when they are facing orthogonally.

Voltage is built up in an onboard storage capacitor, and when sufficient charge has accumulated to reach or surpass the circuit operating threshold voltage, the electronics power up and begin transmitting data back to the reader. Both the reader and the tag must use the same transmission method in order to synchronize and successfully exchange data. Two main methods of communication occur between the reader and tag; full duplex and half-duplex. In a full-duplex configuration, the tag communicates its data by modulating the reader's carrier wave by applying a resistive load. A transistor (load modulator) within the tag shorts the antenna circuit in sequence to the data, removing the antenna from resonance at the excitation frequency, thereby removing its power draw from the reader's carrier wave. At the reader side, the loading and unloading are detected and the data can be reconstructed. In a half-duplex RFID system, the carrier wave transmits power and then pauses. Within the pause, the tag transmits the data back to the reader.

For a given tag, the operating voltage obtained at a distance r from the reader is directly proportional to the flux density at that distance. The magnetic field emitted by the reader antenna decreases in power proportional to $1/r^3$ in the near field. Therefore, it can be shown that for a circularly coiled antenna, the flux density is maximized at a distance r (in meters) when:

$$a = \sqrt{2} \cdot r \tag{5.34}$$

where a is the radius of the reader's antenna coil. Thus, by increasing a the communication range of the reader can be increased, and the optimum reader antenna radius a is 1.41 times the demanded read range r.

The *quality factor* or Q value of the coupling element defines how well the resonating circuit absorbs power over its relatively narrow resonance band. In smart-label RFID applications, the Q value demanded is reasonably high. Because most of the resonant circuit's tuning capacitance is located within the IC microchip where high capacitor Q can be realized, the effective circuit Q value is determined mainly by the antenna coil losses. The coil Q is usually

calculated (without taking into account additional parasitic capacitance losses) according to this equation:

$$Q = \frac{\omega L}{R_S} = \frac{1}{\omega CR} = \frac{\omega}{\omega_2 - \omega_1} \tag{5.35}$$

In general, the higher the Q, the higher the power output for a particular size of antenna. Unfortunately, too high a Q may conflict with the bandpass characteristics of the reader, and the increased ringing could create problems in the protocol bit timing. In (5.35), R_s is the coil's total effective series loss resistance, taking into account both the dc resistance and the ac resistance due to high-frequency current flow concentration caused by skin-effect phenomena in the conductor windings. Practical smart-label systems usually operate with a coupling element resonator Q, within the range of 20 to 80. The Q of the LC circuit is typically around 20 for an air-core inductor and about 40 for a ferrite-core inductor. Higher Q, values than this are generally not feasible because the information-bearing, amplitude-modulated reply sidebands are undesirably attenuated by the resonator's bandpass frequency response characteristic. At resonance, the induced RF voltage produced across the tuned tag and delivered to the microchip will be Q times greater than for frequencies outside of the resonant bandwidth.

Figure 5.7 shows the frequency response curve for a typical serial resonant tank circuit. A good rule of thumb is to stay within the –3-dB limits; the individual manufacturing tolerances for capacitance and inductance of 2%, a Q of 30, can be used. Lower tolerance components may be used at the expense of

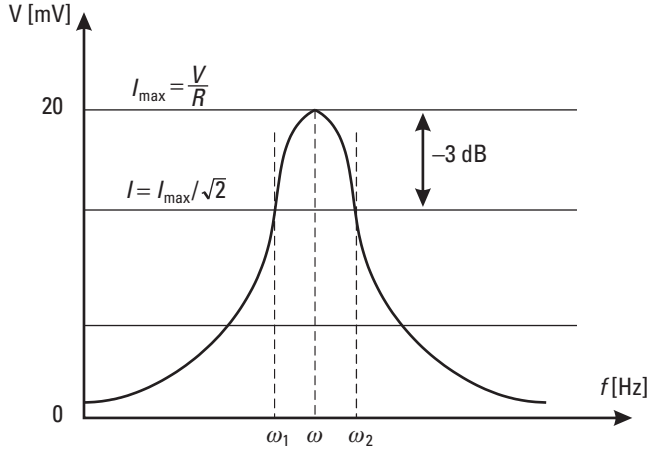


Figure 5.7 Frequency response curve for resonant tank circuit.

sensitivity and, thus, yield a lower range. The corresponding final design must accommodate a wider bandwidth and will, therefore, have a lower response.

As a resonant application, the smart-label tag can be vulnerable to environmental detuning effects that may cause a reduction in transponder sensitivity and reading distance. Undesirable changes in the tag's parasitic capacitance and effective inductance can happen easily. The presence of metal and different dielectric mediums can cause detuning and introduce damping resulting from dissipative energy losses. Such permeable materials can also distort the magnetic flux lines to weaken the energy coupling to the tag. However, these effects can largely be overcome when they are taken into account during the label and system design phase.

Clusters of tagged objects that sometimes come together in physical proximity to each other can also exhibit significant *detuning effects* caused by their mutual inductances. This shift in tuning is called *resonance splitting,* and it is an expected outcome when two or more tags are brought too close to one another. They become coupled tuned circuits, and the degree of coupling (called the *coupling coefficient, k*) determines the amount of frequency shift. The value of *k* depends on the coil geometry (size and shape) and spacing distance. Bigger area coils are inherently more susceptible to deleterious mutual coupling effects. When in proximity, the magnetic flux lines of the individual coils overlap, and the coils exhibit mutual inductance. *This mutual inductance generally adds to the coil's normal inductance and produces a downward shift in the effective resonant frequency.* This in turn results in the tag receiving less energy from the reader field and, hence, the reading distance decreases accordingly. The higher the tag *Q,* the more pronounced the effect. Closely coupled tags can also have problems with commands signaled from the reader being misinterpreted due to cross-coupling between tags.

This rapid attenuation of the energizing and data communication field with increasing distance is the fundamental reason why 13.56-MHz passive RFID systems have a maximum reading distance on the order of about 1m (3 feet). This is also the reason why well-designed near-field RFID systems have good immunity to environmental noise and electrical interference. All of these characteristics are particularly well suited to many smart-label applications.

The efficiency of power transfer between the antenna coil of the reader and the transponder is proportional to the operating frequency, the number of windings, the area enclosed by the transponder coil, the angle of the two coils relative to each other, and the distance between the two coils. As frequency increases, the required coil inductance of the transponder coil—and thus the number of windings—decreases. For 135 kHz it is typically 100 to 1,000 windings, and for 13.56 MHz, typically 3 to 10 windings. Because the voltage induced in the transponder is still proportional to frequency, the reduced number of windings barely affects the efficiency of power transfer at higher frequencies.

5.3 Tags

5.3.1 Tag Considerations

There really is no such thing as a typical RFID tag. The read range is a balancing act between a number of engineering trade-offs and ultimately depends on many factors: the frequency of RFID system operation, the power of the reader, and interference from other RF devices. Several general RFID tag design requirements whose relative importance depends on tag application are discussed here. These requirements largely determine the criteria for selecting an RFID tag antenna:

- *Frequency band:* Desired frequency band of operation depends on the regulations of the country where tag will be used.
- *Size and form:* Tag form and size must be such that it can be embedded or attached to the required objects (cardboard boxes, airline baggage strips, identification cards, and so on) or fit inside a printed label (Figure 5.8).
- *Read range:* Minimum required read range is usually specified.
- *EIRP:* EIRP is determined by local country regulations (active versus passive tags).
- *Objects:* Tag performance changes when it is placed on different objects (e.g., cardboard boxes with various content) or when other objects are present in the vicinity of the tagged object. A tag's antenna can be designed or tuned for optimum performance on a particular object or designed to be less sensitive to the content on which the tag is placed.

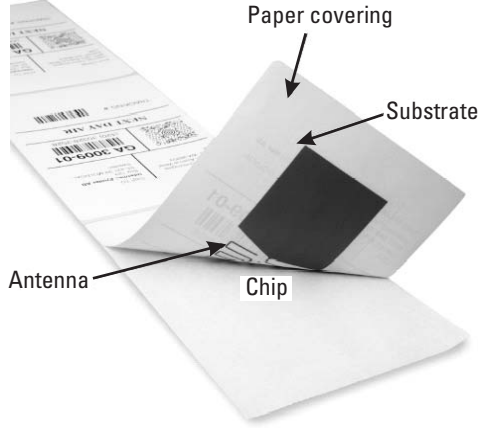


Figure 5.8 RFID label cross section.

- *Orientation (also called polarization):* The read range depends on antenna orientation. How tags are placed with respect to the polarization of the reader's field can have a significant effect on the communication distance for both HF and UHF tags, resulting in a reduced operating range of up to 50%, and in the case of the tag being displaced by 90° and not being able to read the tag at all. The optimal orientation for HF tags is for the two antenna coils (reader and tag) to be parallel to each other (Figure 5.9). UHF tags are even more sensitive to polarization due to the directional nature of the dipole fields. Some applications require a tag to have a specific directivity pattern such as omnidirectional or hemispherical coverage.

- *Applications with mobility:* RFID tags can be used in situations where tagged objects, such as pallets or boxes, travel on a conveyor belt at speeds up to either 600 feet/minute or 10 mph. The Doppler shift in this case is less than 30 Hz at 915 MHz and does not affect RFID operation. However, the tag spends less time in the read field of the RFID reader, demanding a high-read-rate capability. In such cases, the RFID system must be carefully planned to ensure reliable tag identification.

- *Cost:* The RFID tag must be a low-cost device, thus imposing restrictions both on antenna structure and on the choice of materials for its construction including the ASIC used. Typical conductors used in tags are copper, aluminum, and silver ink. The dielectrics include flexible polyester and rigid PCB substrates, such as FR4.

- *Reliability:* The RFID tag must be a reliable device that can sustain variations in temperature, humidity, and stress and survive such processes as label insertion, printing, and lamination.

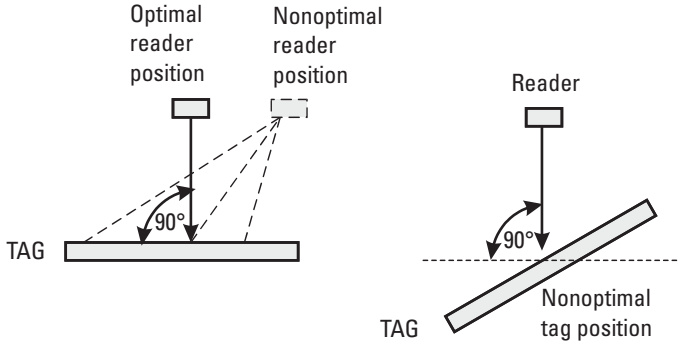


Figure 5.9 Optimal and nonoptimal tag and reader position.

- *Power for the tag:* An active tag has its own battery and does not rely on the reader for any functions. Its range is greater than that of passive tags. Passive tags rely on the reader for power to perform all functions, and semipassive tags rely on the reader for powering transmission but the battery for powering their own circuitry. For comparison, see Table 5.2.

5.3.2 Data Content of RFID Tags

5.3.2.1 Read-Only Systems

Read-only systems can be considered low end; these tags usually only contain an individual serial number that is transmitted when queried by a reader. These systems can be used to replace the functionality of barcodes. Due to the structural simplicity of read-only tags, costs and energy consumption can be kept down.

More advanced tags contain logic and memory, so they support writing and information can be updated or changed remotely. High-end tags have microprocessors enabling complex algorithms for encryption and security. More energy is needed for these than for less complex electronics.

One-bit tags can be detected, but they do not contain any other information. This is very useful for protecting items in a shop against shoplifters. A system like this is called electronic article surveillance (EAS) and has been in use since the 1970s. In practice, this system can be identified by the large gates of coils or antennas at the exits of shops.

Different principles of operation can be used for the one-bit tags. For example, the principle of the microwave tag is quite simple; it uses the generation of harmonics by diodes, that is, frequencies that are an integer times the original frequency. The tag is a small antenna that has a diode in the middle. Because the diode only lets current pass one way, the oscillations that get trapped behind the diode generate a frequency that is twice that of the original one. The system sends out a microwave signal, for example, 2.45 GHz, and listens for the first harmonics, that is, 4.90 GHz. If a tag is present, it generates harmonics that can be detected. However, false alarms may be caused by other

Table 5.2
Power for the Tag

Tag Type	Power Source	Memory	Communication Range
Active	Battery	Most	Greatest
Semipassive	Battery and reader	Moderate	Moderate
Passive	Reader	Least	Least

sources of this particular frequency. To avoid such false alarms, a modulation signal of, for example, 100 kHz is added to the interrogation signal. This means that the same modulating signal can also be found in any reflections from the tags. Microwave EAS tags are usually used to protect clothing. They are removed at the checkout and reused.

Read-only tags that contain more than 1 bit of data are simple ones that only contain a unique serial number that it transmits on request. The contents of the read-only chips are usually written during manufacturing. The serial number can, for example, be coded by cutting small bridges on the chip. Usually these simple chips also contain some logic for anticollision; that is, they allow multiple tags to be read simultaneously.

5.3.2.2 Read/Write Systems

Read/write systems specify what is possible in terms of read-only and read/write capability. It is worth remembering that many read-only tags are factory programmed and carry an identification number (tag ID). Other tags, including read/write devices, can also carry a tag identifier that is used to unambiguously identify a tag. This identifier is distinct from user-introduced identifiers for supporting other application needs.

Read-only devices are generally less costly and may be factory programmable as read only or one-time programmable (OTP). One-time programmability provides the opportunity to write once then read many times, thus supporting passport-type applications, in which data can be added at key points during the lifetime or usage of an item, and thus provide an incorruptible history or audit trail for the item data.

Some chips allow writing only once, and they are often referred to as write once/read many (WORM). These tags are versatile because they can be written with a serial number when applied to an item, instead of linking a predefined serial number to an item. More advanced chips allow both reading and writing multiple times, and the contents of a tag can be altered remotely by a scanner.

Read/write data carriers offer a facility for changing the content of the carrier as and when appropriate within a given application. Some devices will have both a read-only and a read/write component that can support both identification and other data carrier needs. The read/write capability can clearly support applications in which an item, such as a container or assembly support, is reusable and requires some means of carrying data about its contents or on what is being physically carried. It is also significant for lifetime applications such as maintenance histories, where a need is seen to add or modify data concerning an item over a period of time. The read/write capability may also be exploited within flexible manufacturing to carry and adjust manufacturing information and item-attendant details, such as component tolerances. A further important use of read/write is for local caching of data as a portable data file, using it as and

when required, and selectively modifying it as appropriate to meet process needs.

Additionally, the chip must be able to resolve who can access it and prevent the wrong people from altering its contents. For secure data transmission, some kind of encryption is added as well.

Time to read is the time it takes to read a tag, which of course is related to the data transfer rate. For example, a system operating at a 1-Kbps transfer rate will take approximately 0.1 second to read a 96-bit tag, with a bit of time for the communication management. Various factors can influence read time, including competing readers and tags (reader access and multiple tags).

5.3.3 Passive Tags

5.3.3.1 About Passive Tags

Passive RFID devices have no power supply built in, meaning that electrical current transmitted by the RFID reader inductively powers the device, which allows it to transmit its information back. Because the tag has a limited supply of power, its transmission is much more limited than an active tag, typically no more than simply an ID number. Similarly, passive devices have a limited range of broadcast, requiring the reader to be significantly closer than an active one would. Uses for passive devices tend to include things such as inventory, product shipping and tracking, use in hospitals and for other medical purposes, and antitheft, where it is practical to have a reader within a few meters or so of the RFID device. Passive devices are ideal in places that prevent the replacement of a battery, such as when implanted under a person's skin.

Tags consist of a silicon device (chip) and antenna circuit (Figure 5.10). The purpose of the antenna circuit is to induce an energizing signal and to send a modulated RF signal. The read range of a tag largely depends on the antenna circuit and size. The antenna circuit is made of an LC resonant circuit or E-field dipole antenna, depending on the carrier frequency. The LC resonant circuit is

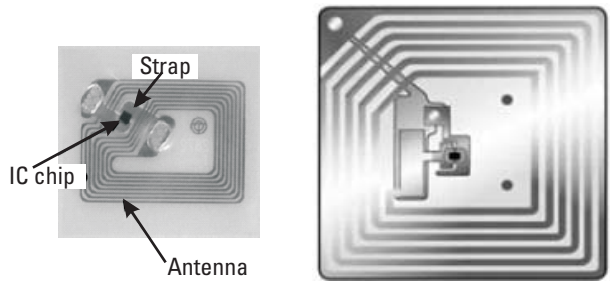


Figure 5.10 13.56-MHz RFID tags.

typically used for frequencies of less than 100 MHz. In this frequency band, the communication between the reader and tag takes place with magnetic coupling between the two antennas through the magnetic field. An antenna utilizing inductive coupling is often called a *magnetic dipole antenna*. The antenna circuits must be designed in such a way as to maximize the magnetic coupling between them. This can be achieved with the following parameters:

- The LC circuit must be tuned to the carrier frequency of the reader.
- The Q of the tuned circuit must be maximized.
- The antenna size must be maximized within the physical limits of application requirements.

The passive RFID tags sometimes use backscattering of the carrier frequency for sending data from the tag to the reader. The amplitude of the backscattering signal is modulated with modulation data from the tag device. The modulation data can be encoded in the form of ASK (NRZ or Manchester), FSK, or PSK. During *backscatter modulation*, the incoming RF carrier signal to the tag is loaded and unloaded, causing amplitude modulation of the carrier corresponding to the tag data bits. The RF voltage induced in the tag's antenna is amplitude modulated by the modulation signal (data) of the tag device. This amplitude modulation can be achieved by using a modulation transistor across the LC resonant circuit or partially across the resonant circuit. Changes in the voltage amplitude of the tag's antenna can affect the voltage of the reader antenna. By monitoring the changes in the reader antenna voltage (due to the tag's modulation data), the data in the tag can be reconstructed. (See Chapter 6 for more details on modulation.)

The RF voltage link between the reader and tag antennas is often compared to weakly coupled transformer coils; as the secondary winding (tag coil) is momentarily shunted, the primary winding (reader coil) experiences a momentary voltage change. Opening and shunting the secondary winding (tag coil) in sequence with the tag data is seen as amplitude modulation at the primary winding (reader coil).

5.3.3.2 RFID Chip Description

An RFID tag consists of an RFID chip, an antenna, and tag packaging. The RFID circuitry itself consists of an RF front end, some additional basic signal processing circuits, logic circuitry to implement the algorithms required, and EEPROM for storage. The RFID chip is an integrated circuit implemented in silicon [5]. The major blocks and their functions of the RFID front end are as follows:

- *Rectifier:* Generates the power supply voltage for front-end circuits and the whole chip, as well from the coupled EM field;
- *Power (voltage) regulator:* Maintains the power supply at a certain level and at the same time prevents the circuit from malfunctioning or breaking under large input RF power;
- *Demodulator:* Extracts the data symbols embedded in the carrier waveforms;
- *Clock extraction or generation:* Extracts the clock from the carrier (usually in HF systems) or generates the system clock by means of some kind of oscillator;
- *Backscattering:* Fulfills the return link by alternating the impedance of the chip;
- *Power on reset:* Generates the chip's power-on reset (POR) signal;
- *Voltage (current) reference:* Generates some voltage or current reference for the use of front-end and other circuit blocks, usually in terms of a bandgap reference;
- *Other circuits:* These include the persistent node or short-term memory (or ESD).

Figure 5.11 is a block diagram for RFID IC circuits and lists many of the circuit's associated function blocks. The RF front end is connected to the antenna, and typically, at UHF, an electric dipole antenna is used, while HF tags use a coil antenna. The front-end circuitry impacts the semiconductor process by requiring a process that allows for mixed-mode fabrication. Passive RF tags have no power source and rely on the signal from the reader to power up; thus, the RF front end implements modulators, voltage regulators, resets, and connections to an external antenna. RFID chips have control logic that typically consists of a few thousand gates. The lowest level chip uses very few gates, on the order of 1,500 gate equivalents. Functions in the logic include the error and parity/CRC checkers, data encoders, anticollision algorithms, controllers, and command decoders. More complex RFID chips may include security primitives and even tamperproofing hardware. The size of the circuit affects the number of mask, metal, and poly layers required in the semiconductor process, and RFID systems usually use CMOS.

A certain amount of information is stored on-chip in an EEPROM. The size of this EEPROM increases as more information is required to be on the RFID chip. The size of the required EEPROM is a factor in determining the number of mask, metal, and poly layers required in the semiconductor fabrication process. It is also a factor in determining the size of the final semiconductor die. Silicon cost is directly proportional to both the die size and the number of

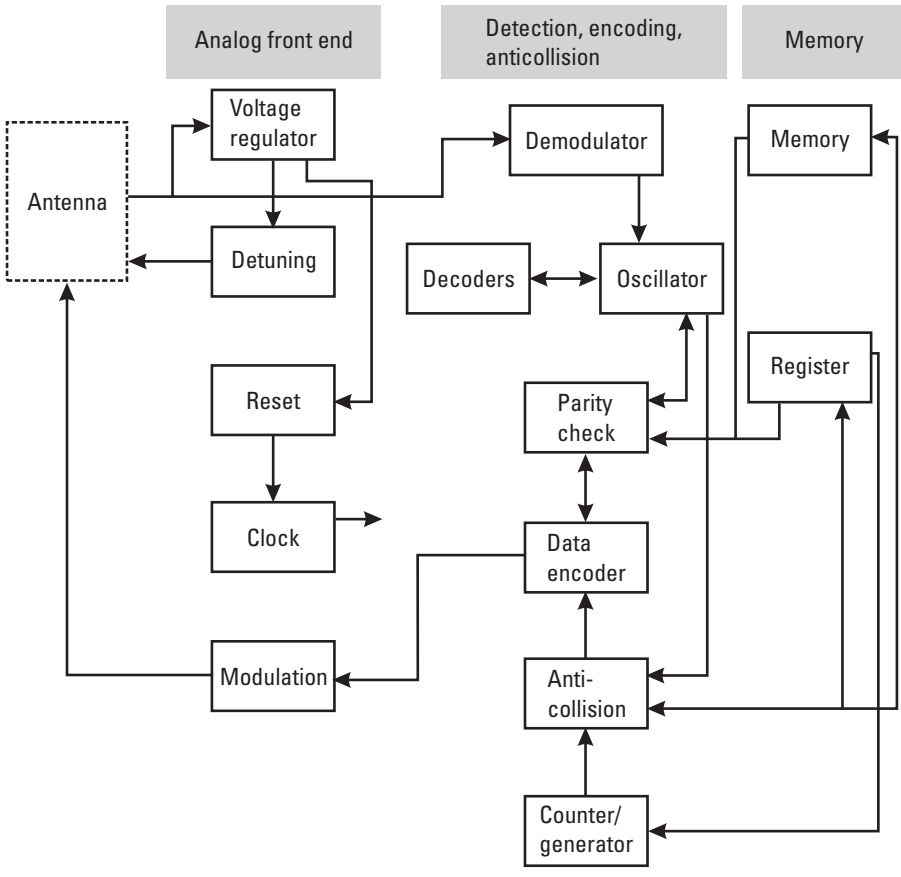


Figure 5.11 RFID tag circuit block diagram.

mask, metal, and poly metal layers. The IC in an RFID tag must be attached to an antenna to operate. The antenna captures and transmits signals to and from the reader. The coupling from the reader to the tag provides both the transmission data and the power to operate the passive RFID tag. Typically, antennas for passive RFID systems can be either simple dipole, 915-MHz RFID tags or more complex coiled shapes for 13.56-MHz systems.

The digital anticollision system is one of the major and most important parts of the tag chip, because it not only implements the slotted ALOHA random anticollision algorithm, but also executes the read/write operation of memory. As we know, the power consumption of memory is very difficult to reduce. Even more, besides the power consumption, the efficiency of the RF front-end rectifier prefers lower output dc voltage. So it is very important to design a low-power, low-voltage digital anticollision system to achieve maximum operating range.

Currently, antennas are made of metals or metal pastes and typically cost as much as 12 cents per antenna to manufacture. However, new methods that range from conductive inks to new antenna deposition and stamping techniques are expected to reduce costs below 1 cent.

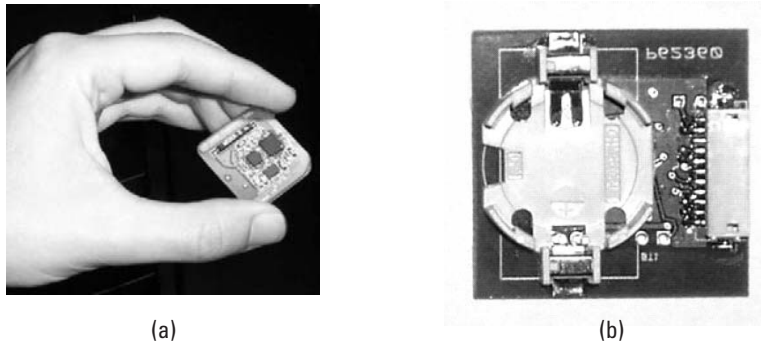
5.3.4 Active Tags

5.3.4.1 Active Tag Description

An active tag usually performs a specialized task and has an on-board power source (usually a battery). It does not require inductions to provide current, as is true of the passive tags. The active tag can be designed with a variety of specialized electronics, including microprocessors, different types of sensors, or I/O devices (Figure 5.12). Depending on the target function of the tag, this information can be processed and stored for immediate or later retrieval by a reader. Active RFID tags, also called *transponders* because they contain a transmitter that is always on, are powered by a battery about the size of a coin and are designed for communications up to 100 feet from the RFID reader. They are larger and more expensive than passive RFID tags, but can hold more data about the product and are commonly used for high-value asset tracking. A feature that most active tags have and most passive tags do not is the ability to store data received from a transceiver.

Active tags are ideal in environments with electromagnetic interference because they can broadcast a stronger signal in situations that require a greater distance between the tag and the transmitter.

The additional space taken up by a battery in an active device necessitates that the active devices be substantially larger than the passive devices. To date, commercially available passive tags are as small as 0.4 mm square and thinner than a sheet of paper. In contrast, commercially available active tags are still only



(a) (b)

Figure 5.12 Active tag (a) front and (b) reverse sides.

as small as a coin, which means that active tags are around 50 times the size of passive ones.

For the *read-only device*, the information that is in the memory cannot be changed by an RF command once it has been written. A device with memory cells that can be reprogrammed by RF commands is called a *read/write device*. The information in the memory can be reprogrammed by an interrogator command.

Although passive tags can only respond to an electromagnetic wave signal emitted from a reader, active tags can also spontaneously transmit an ID. There are various types of unscheduled transmission types, such as when there are changes in vibration or temperature or when a button is pushed.

A semiactive or semipassive (depending on the manufacturer) tag also has an on-board battery. The battery in this case is only used to operate the chip. Like the passive tag, it uses the energy in the electromagnetic field to wake up the chip and to transmit the data to the reader. These tags are sometimes called battery-assisted passive (BAP) tags.

5.3.4.2 Active Tag Classification

Two types of tag systems can generally be recognized within active RFID systems:

- *Wake-up tag systems* are deactivated, or asleep, until activated by a coded message from a reader or interrogator. In the sleep mode, limiting the current drain to a low-level alert function conserves the battery energy. Where larger memories are accommodated, there is also generally a need to access data on an object or internal file basis to avoid having to transfer the entire amount of data so held. These are used in toll payment collection, checkpoint control, and in tracking cargo.
- *Awake tag or beacon systems* are, as the term suggests, responsive to interrogation without a coded message being required to switch the tag from an energy conservation mode. However, they generally operate at lower data transfer rates and memory sizes than wake-up tags, so they conserve battery energy in this way. (A greater switching rate is generally associated with higher energy usage.) This type of tag is the most widely used of the two, and because of lower component costs it is generally less expensive than a wake-up tag system. Beacons are used in most real-time locating systems (RTLS), where the precise location of an asset needs to be tracked. In an RTLS, a beacon emits a signal with its unique identifier at preset intervals, every 3 seconds or once a day, depending on how important it is to know the location of an asset at a particular moment in time.

5.3.5 Active Versus Passive Tags

Active RFID and passive RFID technologies, while often considered and evaluated together, are fundamentally distinct technologies with substantially different capabilities. In most cases, neither technology provides a complete solution for supply chain asset management applications; rather, the most effective and complete supply chain solutions leverage the advantages of each technology and combine their use in complementary ways. Passive RFID is most appropriate where the movement of tagged assets is highly consistent and controlled and little or no security, sensing capability, or data storage is required. Active RFID is best suited where business processes are dynamic or unconstrained, movement of tagged assets is variable, and more sophisticated security, sensing, and/or data storage capabilities are required.

Passive and active tagging systems present very different deployment issues. Active tags contain significantly more sophistication, data management, and security concerns. Active tags generally cost from $10 to $50, depending on the amount of memory, the battery life required, any on-board sensors, and the ruggedness.

5.3.6 Multiple Tag Operation

If many tags are present, then they will all reply at the same time, which at the reader end is seen as a signal collision and an indication of multiple tags. The reader manages this problem by using an anticollision algorithm designed to allow tags to be sorted and individually selected. The many different types of algorithms (Binary Tree, ALOHA, and so on) are defined as part of the protocol standards. The number of tags that can be identified depends on the frequency and protocol used, and can typically range from 50 tags per second for HF up to 200 tags per second for UHF. Once a tag is selected, the reader is able to perform a number of operations, such as reading the tag's identifier number or, in the case of a read/write tag, writing information to it. After finishing its dialogue with the tag, the reader can then either remove it from the list, or put it on standby until a later time. This process continues under control of the anticollision algorithm until all tags have been selected.

When containers of freight are moved on a conveyor or similar equipment in a tag reader system, the reader/writer must read and write data to and from moving tags (Figure 5.13). For successful access, the following conditions must be satisfied (5.32):

$$T_c = A_{cn} \times \left(D_t / D_r \right) \tag{5.36}$$

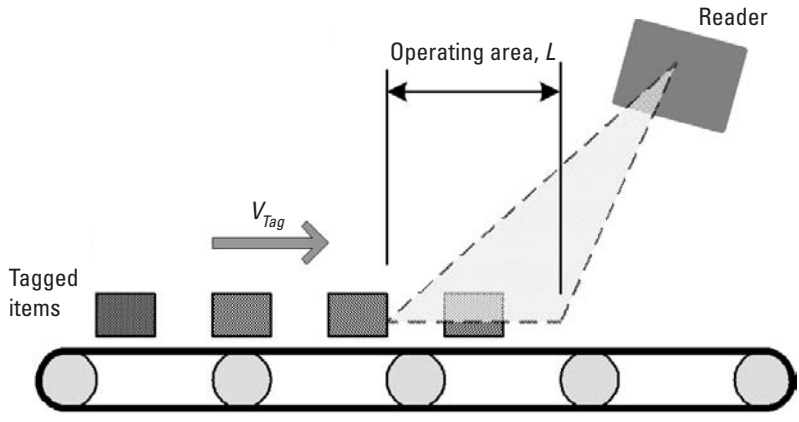


Figure 5.13 Reading moving tags.

This formula shows that when the data transfer volume of the tag D_t increases and the data transfer rate D_r decreases, the tag-reader/writer operation time T_c increases, and operation may fail.

$$T_r = L/V_{Tag} \tag{5.37}$$

Equation (5.37) shows that when the reader/writer operating area decreases, the distance the tag moves (L) decreases, and the tag movement velocity V_{Tag} increases, the amount of time the tag is in the operating area (T_r) decreases, and the operation may fail.

$$T_r \geq T_c + T_d \tag{5.38}$$

Last, (5.38) states that the total amount of time spent in the operating area must be more than the total time taken by the reader/writer and the detection of all tags. If only one type of tag can be used when reading/writing RFID tags attached to freight on a conveyor belt, the reader/writer antenna must have a large operating area to cope with the conveyor belt's speed:

T_r [seconds] = amount of time tag is in operating area;

T_c [seconds] = tag-reader/writer operation time;

D_r [bps] = data transfer rate;

D_t [bit] = data transfer volume;

A_{cn} [count] = average number of tag-reader/writer operations;

V_{Tag} [m/second] = tag movement velocity;

L [m] = distance tag moves within operating area;

T_d [seconds] = amount of time for detecting existence of all tags.

Virtually all high-volume RFID applications require the ability to read multiple tags in the reading field at one time. This is only possible if each RFID tag has a unique ID number. One numbering method is the EPC code, which contains both an item ID number and a serial number. A unique number is the basis for implementing anticollision in any RFID technology. In a multiple-tag operation, where multiple RFID tags are in the reader/writer's operating area, the reader/writer must detect the presence of these multiple tags and read/write each of them consecutively (Figure 5.14). This operating method is generally referred to as the *anticollision protocol* and is different from the single-tag operation protocol.

The effects of operating range, tag orientation, tag movement velocity, and the presence of metallic substances on multiple-tag operating characteristics are basically the same as those on single-tag operating characteristics. One problem with multiple-tag operating characteristics is that the operating time is several times longer than for single-tag operation. Because the reader/writer must read/write each tag, the time increases in proportion to the number of tags. Also,

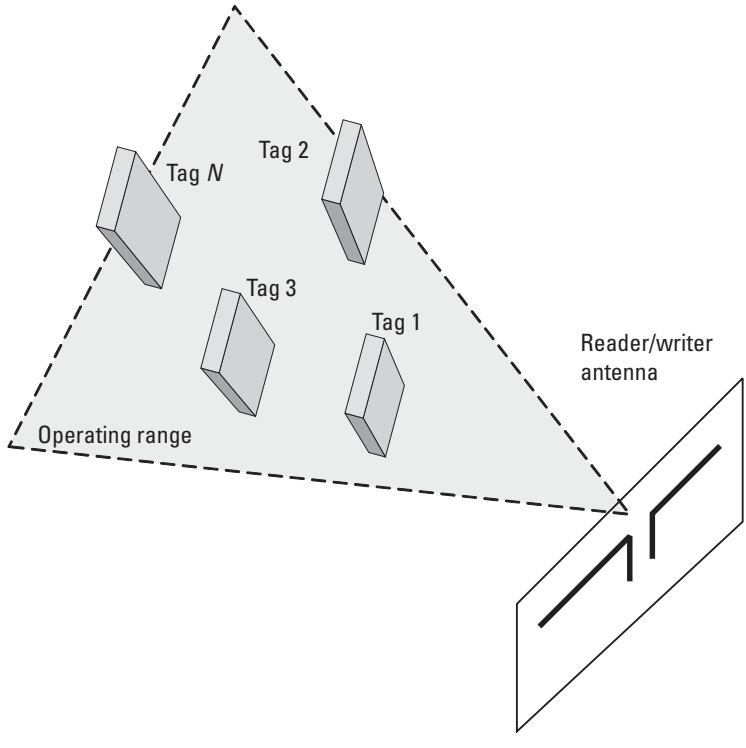


Figure 5.14 Multiple-tag operation.

because multiple tags are used, tags sometimes come into contact or overlap with each other. When there are N tags in the operating area and N_{tag} is the number of tags, the amount of time for which the tags must be in the operating area (T_r) is described by (5.39):

$$T_r \geq \left(T_c + T_{tag}\right) \times N_{tag} \tag{5.39}$$

Although the reader/writer may sometimes read/write stationary tags, in most cases, the tag will be moving. The reader/writer will generally have to read/write RFID tags attached to containers or freight being transported on a conveyor belt or trolleys. When T_c is the operating time for a single tag and T_d is the time required to check for the existence of N multiple tags in an anticollision protocol, (5.40) gives an approximation of the maximum time required for the reader/writer to read/write all N tags (T_N):

$$T_N = \left(T_c + T_d\right) \times N_{tag} \tag{5.40}$$

Example:

Tag information volume = 16 bytes;

Data transfer rate (D_r) = 7.8 Kbps;

$T_c = 0.057$ second;

$T_d = 0.055$ second;

$N = 10$;

$T_N = (0.057 + 0.055) \times 10 = 1.12$ seconds.

Therefore, roughly 1.1 seconds are needed for the reader/writer to finish reading/writing all 10 tags. When the tag information volume is 100 bytes, T_N becomes roughly 7 seconds. For the reader/writer to read/write all the tags, the time required for the tags to pass through the operating area (T_r) must be greater than T_N.

One unfortunate but real fact about RFID tags is that the quality of tags is currently not consistent, and therefore performance is not consistent. Considerable variations are seen in performance from one tag to the next, even among tags from the same manufacturer and model.

5.3.7 Overlapping Tags

In inductive frequency band RFID, the resonance characteristic of the tag antenna coil is used for reader/writer operation. As discussed earlier, a tag's resonant frequency f_0 is calculated by (5.41):

$$f_0 = 1/2\pi\sqrt{LC} \tag{5.41}$$

where L [H] is the inductance of tag antenna coil and C [F] is the capacitance of the tag's tuning capacitor.

If tags overlap, the inductance of their antenna coils is obstructed, and L increases. In this case, the resonant frequency expressed by the formula becomes lower ($f_1 < f_0$). As a result, the electromagnetic waves (current i) generated by the tag's coil become smaller, and the operating area decreases (Figure 5.15).

5.3.8 Tag Antennas

5.3.8.1 Antenna Selection

An antenna is a conductive structure specifically designed to couple or radiate electromagnetic energy. Antenna structures, often encountered in RFID systems, may be used to both transmit and receive electromagnetic energy, particularly data-modulated electromagnetic energy. In the low-frequency (LF) range

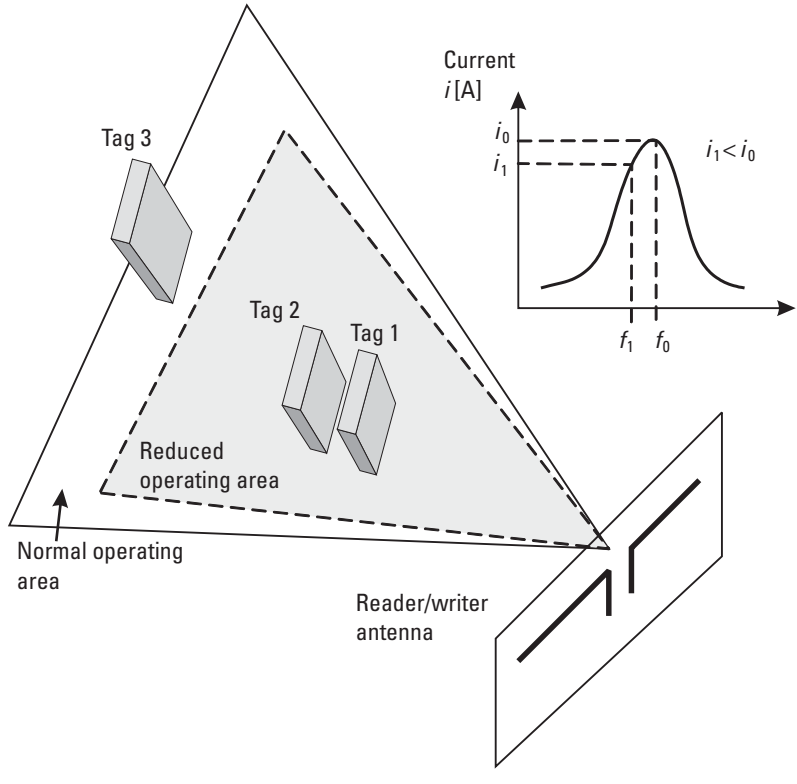


Figure 5.15 Overlapping tags.

with short read distances, the tag is in the near field of the reader antenna, and the power and signals are transferred by means of a magnetic coupling. In the LF range, the tag antenna therefore comprises a coil (inductive loops) to which the chip is attached. In the UHF range, in cases where the read distances are larger, the tag is located in the far field of the reader antenna. The reader and tag are coupled by the electromagnetic wave in free space, to which the reader and tag are tuned by means of appropriate antenna structures.

Good antenna design is a critical factor in obtaining good range and stable throughput in a wireless application. This is especially true in low-power and compact designs where antenna space is less than optimal. It is important to remember that, in general, the smaller the antenna, the lower the radiation resistance and the lower the efficiency. The tag antenna should be as small as possible and easy to produce.

Printed antennas are really very easy to produce. The antenna is attached as a flat structure to a substrate. The next stage in the production process often involves attaching the chip to the substrate and connecting it to the antenna. This assembly is called an inlay. An inlay becomes a tag or transponder when it is fixed to an adhesive label or a smart card. Note, however, that the electromagnetic properties of the materials surrounding the inlay affect the tag's ability to communicate. In extreme cases, tags cannot be read if unsuitable reader antennas are selected.

Another type of usage involves integration into the object that is to be identified. Parts of the object can be shaped to form an antenna and the antenna can be adjusted optimally to suit the object. This significantly increases readability, while simultaneously protecting against counterfeiting.

The size and shape of the tag antenna have a significant effect on tag read rates, regardless of the coupling used for communication. Various types of antennas are available, among which the most commonly used are dipole, folded dipole, printed dipole, printed patch, squiggle, and log-spiral. Among these, the dipole, folded dipole, and squiggle antennas are omnidirectional, thus allowing them to be read in all possible tag orientations, relative to the base antenna. On the other hand, directional antennas have a good read range due to their good resistance to radiation patterns. Care must be taken while choosing an antenna because the antenna impedance must match to the ASIC and to free space. The four major considerations when choosing an antenna are as follows:

- Antenna type;
- Antenna impedance;
- Nature of the tagged object;
- Vicinity of structures around the tagged object.

When individual system performance is not satisfactory, it is advisable to bring redundancy to the system. Low read rates of RFID systems make the deployment of redundant antennas and tags to identify the same object an imperative. Redundant tags are those tags that carry identical information performing identical functions. Dual tags are tags connected to each other that have one or two antennas and are with or without individual or shared memory; *n* tags serving the same purpose as that of dual tags can be used for beneficial use of multiple tags in product identification. It has been observed that both the inductive coupling and backscatter-based tags are dependent on the angle of orientation of the tag relative to the reader. The placement of two tags in two flat planes, three tags in the three-dimensional axes, four tags along the faces of a regular tetrahedron, and so on, can help in achieving the above-mentioned goals.

The choice of an etched, printed, or stamped antenna is a trade-off between cost and performance. For a 13.56-MHz tag, the *Q* factor of the antenna is very important for long read range applications. The *Q* factor is inversely proportional to the resistance of the antenna trace. It has been determined that the etched antenna is less resistive and inexpensive than the printed antenna with conductive material. However, for a very large antenna size (greater than 4×4 inches), both etching and stamping processes waste too much unwanted material. Therefore, printed or wired antennas should be considered as an alternative.

As previously stated, reducing antenna size results in reduced performance. Some of the parameters that suffer are reduced efficiency (or gain), shorter range, smaller useful bandwidth, more critical tuning, increased sensitivity to component and PCB spread, and increased sensitivity to external factors. Several performance factors deteriorate with miniaturization, but some antenna types tolerate miniaturization better than others. How much a given antenna can be reduced in size depends on the actual requirements for range, bandwidth, and repeatability. In general, an antenna can be reduced to half its natural size with moderate impact on performance.

5.3.8.2 Loop Antennas

RFID tags extract all of their power to both operate and communicate from the reader's magnetic field. Coupling between the tag and reader is via the mutual inductance of the two loop antennas, and the efficient transfer of energy from the reader to the tag directly affects operational reliability and read/write range. Generally, both 13.56-MHz and 125-kHz RFID tags use parallel resonant LC loop antennas tuned to the carrier frequency. The RFID circuit is similar to a transformer in which loop inductors magnetically couple when one of the loops, in the case of the reader antenna, is energized with an alternating current, thus creating an alternating magnetic field. The tag loop antenna acts like the

secondary winding of a transformer, where an alternating current is induced in the antenna, extracting energy from the magnetic field. Generally, the larger the diameter of the tag's antenna loop, the more magnetic flux lines that are passing through the coil and increasing the transfer of energy from the reader to the tag.

Loop antennas can be divided in three groups:

1. Half-wave antennas;
2. Full-wave antennas;
3. Series-loaded, short-loop antennas.

where *wave* refers to the approximate circumference of the loop.

The *half-wave loop* consists of a loop approximately one-half wavelength in circumference with a gap cut in the ring. It is very similar to a half-wave dipole that has been folded into a ring, and most of the information about the dipole applies to the half-wave loop. Because the ends are very close together, some capacitive loading exists, and resonance is obtained at a somewhat smaller circumference than expected. The feedpoint impedance is also somewhat lower than the usual dipole, but all of the usual feeding techniques can be applied to the half-wave loop. By increasing the capacitive loading across the gap, the loop can be made much smaller than one-half wavelength. At heavy loading, the loop closely resembles a single-winding, LC-tuned circuit. The actual shape of the loop is not critical, and typically the efficiency is determined by the area enclosed by the loop. The half-wave loop is popular at lower frequencies, but at higher frequencies, the tuning capacitance across the gap becomes very small and critical.

As the name implies, the *full-wave loop* is approximately one wavelength in circumference. Resonance is obtained when the loop is slightly longer than one wavelength. The full-wave loop can be thought of as two end-connected dipoles. Like the half-wave loop, the shape of the full-wave loop is not critical, but efficiency is determined mainly by the enclosed area. The feed impedance is somewhat higher (approximately 120Ω) than the half-wave loop. Loading is accomplished by inserting small coils or hairpins in the loop, which reduces the size. Like the dipole and half-wave loop, numerous impedance-matching methods exist, including gamma matching and tapering across a loading coil or hairpin. The main advantage of the full-wave loop is that it does not have the air gap in the loop, which is very sensitive to load and PCB capacitance spread.

Loaded-loop antennas are commonly used in remote control and remote keyless entry (RKE) applications. The loop is placed in series with an inductor, which reduces the efficiency of the antenna but shortens the physical length.

5.3.8.3 UHF Antennas

A typical inductively coupled feeding structure is shown in Figure 5.16(a) where the antenna consists of a feeding loop and a radiating body. Two terminals of the loop are connected to the chip, and the feed is combined with the antenna body with mutual coupling. For example, if the measured impedance of the selected IC is $73 - j113$, the load antenna impedance should be $73 + j113$ for conjugate matching. To achieve this, the proposed antenna structure is shown in Figure 5.16(b), with the dipole arms bent into an arc shape [6].

Another way to achieve high resistance with an inductively coupled feeding structure is to introduce extra radiating elements. A dual-body configuration is presented in Figure 5.16(c). Two meandering line arms are placed in each side of the feeding loop. The slight decrease of mutual coupling is due to the shorter coupling length. However, strong mutual coupling is now introduced between the two radiating bodies, which can be similarly regarded as being in a parallel connection seen from the feeding loop. In this way, resistance of the radiating body is significantly reduced, resulting in high resistance with meandering line arms.

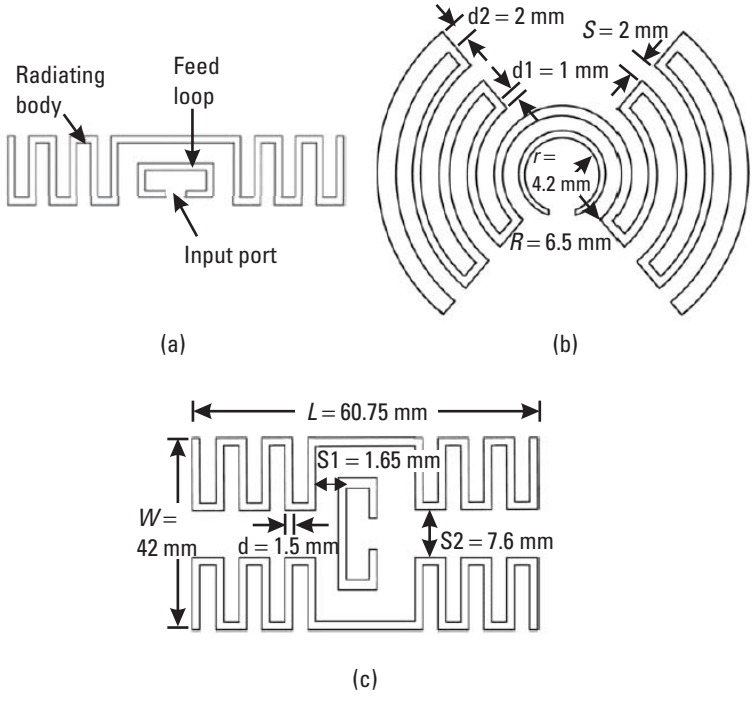


Figure 5.16 UHF antennas: (a) typical configuration, (b) arc configuration, and (c) dual-body configuration.

This antenna can be easily tuned by trimming. Lengths of meander trace and loading bar can be varied to obtain optimum reactance and resistance matching. The trimming is realized by punching holes through the antenna trace at defined locations. Such a tunable design is desirable when a solution is needed for a particular application with minimal lead time.

5.3.8.4 Fractal Antennas

Short reading distances and the fact that the cost per tag is still too high are the major reasons that passive RFID systems have not made their breakthrough yet. One key to greater reading distances is improvements in the tag antenna. Because a passive tag does not have its own power supply, it is important that the tag antenna is able to absorb as much of the energy, radiated from the reader, as possible. Another important parameter to minimize is the size of the tags. Small tags and hence small tag antennas will increase the range of areas in which RFID devices can be employed. The trade-off to designing small, effective antennas is that small antennas are generally poor radiators. A factor that affects the size of the tags is the frequency that is used. Different frequency bands are allocated for RFID and these bands differ in different regions of the world. From an economic point of view, it is highly desirable to be able to use only one type of tag in all of the different regions.

In the study of antennas, fractal antenna theory is a relatively new area. However, fractal antennas and their superset, fractal electrodynamics, are a hotbed of research activity these days. The term *fractal* means linguistically broken or fractured and is from the Latin *fractus.* Fractals are geometrical shapes, which are self-similar, repeating themselves at different scales [7]. Many mathematical structures are fractals, for example, Sierpinski's gasket, Cantor's comb, von Koch's snowflake, the Mandelbrot set, and the Lorenz attractor. Fractals also describe many real-world objects, such as clouds, mountains, turbulence, and coastlines that do not correspond to simple geometric shapes. The terms *fractal* and *fractal dimension* come from Mandelbrot, who is the person most often associated with the mathematics of fractals [8].

Fractal antennas do not have any characteristic size; fractal structures with a self-similar geometric shape consisting of multiple copies of themselves on many different scales have the potential to be frequency-independent or at least multifrequency antennas. For example, it has been shown that a bow-tie antenna can operate efficiently over different frequencies and that the bands can be chosen by modifying the structure. Examples of wideband antennas are the classical spiral antennas and the classical log-periodic antennas, which can also be classified as fractal antennas.

Fractal antennas are convoluted, uneven shapes, and sharp edges, corners, and discontinuities tend to enhance the radiation of electromagnetic energy from electric systems. Fractal antennas, therefore, have the potential to be

efficient. This is particularly interesting when small antennas are to be designed, because small antennas are not generally good at radiating electromagnetic energy. Some fractals have the property that they can be very long but still fit in to a certain volume or area. Because fractals do not have a dimension that is an integer (e.g., it can be something between a line and a plane), they can more effectively fill some volume or area at deposal. Small antennas generally have a very small input resistance and a very significant negative input reactance. This means that small antennas are poor radiators. It has been shown that many small fractal antennas have greater input resistance and smaller input reactance than small traditional antennas. Also, the Q factor of small antennas depends on how effectively the antenna occupies a certain radian sphere. Small fractal antennas can thus be expected to have lower Q factors than their regular counterparts and, hence, higher bandwidths.

Small input resistance and large input reactance also mean that it is difficult (and expensive) to match the antenna input impedance with a matching network. It has been shown that many fractal antennas can even resonate with a size much smaller than the regular ones. Hence, it is possible to reduce or even eliminate the cost associated with input impedance matching. By shaping antennas in certain ways, the directivity can be improved. Fractal antennas are shaped antennas. In some RFID applications where the tag orientation can be controlled, it is possible that antennas with high directivity would be preferred to achieve long reading distances and/or to avoid problems associated with scanning several tags simultaneously. It is also possible that in some applications a small handheld reader with a high directivity antenna would be desirable.

Because frequency-independent antennas and wideband antennas tend to be insensitive to deformations like cutting in the structure and bending of the structure, one might expect some fractal antennas to be resistant against deformations, too.

5.3.9 UHF Tag Circuits

5.3.9.1 Tag DC Supply Voltage Circuitry

The voltage multiplier converts a part of the incoming RF signal power to dc for power supply for all active circuits on the chip (Figure 5.17). The specially designed Schottky diodes with low series resistance allow for high-efficiency conversion of the received RF input signal energy to dc supply voltage.

The voltage multiplier circuit shown here is sometimes also called a *charge pump* in the context of memory ICs. A charge pump is a circuit that when given an input in ac is able to output a dc voltage typically larger than a simple rectifier would generate. It can be thought of as an ac-to-dc converter that both rectifies the ac signal and increases the dc level. It is the foundation of power converters such as the ones that are used for many electronic devices today. In this case, for

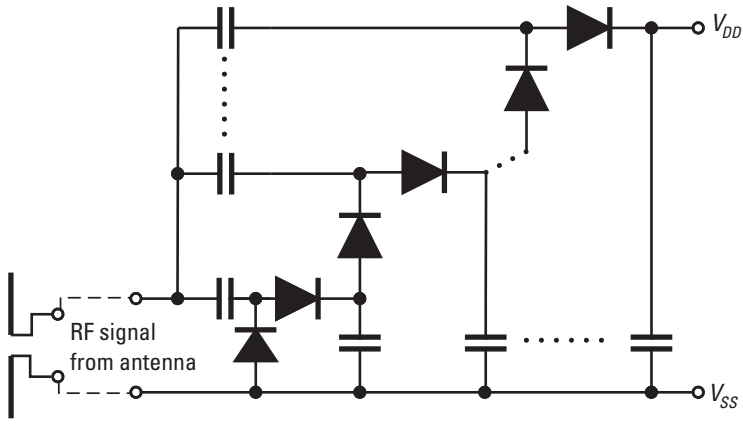


Figure 5.17 Input signal conversion to dc supply voltage.

the RF signal, all of the diodes are connected in parallel (or antiparallel) by the capacitors. For dc, however, they are connected in series to allow a dc current flowing between terminals. The voltage generated between these nodes is approximately equal to:

$$V_{DD} = n\left(V_{\mathrm{RF}} - V_D\right) \tag{5.42}$$

where n is the number of diodes, V_{RF} is the amplitude of the RF input signal, and V_D is the forward voltage of the Schottky diodes (approximately 200 mV at 7 μA).

The input impedance is mainly determined by the junction and substrate capacitances of the Schottky diodes. The real part of the impedance is much lower than the imaginary part and is strongly dependent on the dc current taken from the output. For a typical operating point, the real part of the impedance is approximately 30 times lower than the imaginary (capacitive) part. In other words, the IC's input capacitance has a quality factor of 30, placing high demands on the antenna, which needs to be matched to the IC's input impedance for sufficiently good power efficiency. The design parameters of the voltage multiplier are a trade-off between power efficiency, useful impedance, and operating point (load). Optimization parameters include the number of stages, the size of the Schottky diodes, and the size of the coupling capacitors.

5.3.9.2 Tag Wake-Up Circuit Principles

The interrogation of active RFID tags will inevitably involve the development of a mechanism for turning on the tags because power conservation is an important factor that requires the tags to be turned off when not being interrogated. This

will also be true for active sensors and sensor networks [9]. There are two practical options for turn-on circuit design:

- Rectifier circuits that can produce, from the RF field, a rectified voltage of the order of 1V, which will turn a CMOS transistor from fully off to fully on;
- Rectifier circuits that can produce, from the RF field, a rectified voltage on the order of 5 mV, which when compared to an internal reference voltage can be used to trigger a transistor from the fully off to the fully on state.

Schottky diodes have been widely used in microwave networks because of their excellent high-frequency behavior. For microwave applications, these Schottky diodes are usually fabricated in specialized processes where barrier heights, capacitances, and other parameters can be fully controlled. However, the RFID application demands a low-cost solution, which would tend toward using the standard CMOS processes. However, the problem is that most of the standard CMOS processes do not support the Schottky diode. We have to modify the processes so as to incorporate the Schottky diodes; several research works have been published on standard CMOS processes that are compatible with Schottky diodes. A Schottky junction is relatively delicate and sensitive to excessive RF power, and RFID applications may work in poorly controlled environments where high power may cause the diode to burn out. Hence, in an application it is important to use power limiters to protect the sensitive Schottky diode.

The battery powering active transponders must last for an acceptable time, so the electronics of the label must have very low current consumption in order to prolong the life of the battery. However, due to circuit complexity or the desired operating range, the electronics may drain the battery more rapidly than desired, but use of a turn-on circuit allows the battery to be connected only when communication is needed, thus lengthening the life of the battery. The turn-on circuit shown in Figure 5.18 is adequate and cost effective for a backscattering active tag.

Here, a *p*-channel FET was used as a switch to control the power supply to a label control circuits and can be triggered by the incident RF 915-MHz radiation on the antenna. Thus, the power generated and amplified by the diode resonance can be utilized to turn a *p*-channel FET from an off state to an on state. The turn-on circuit performs adequately at a minimum power of –43 dBW at the resonant frequency of 915 MHz.

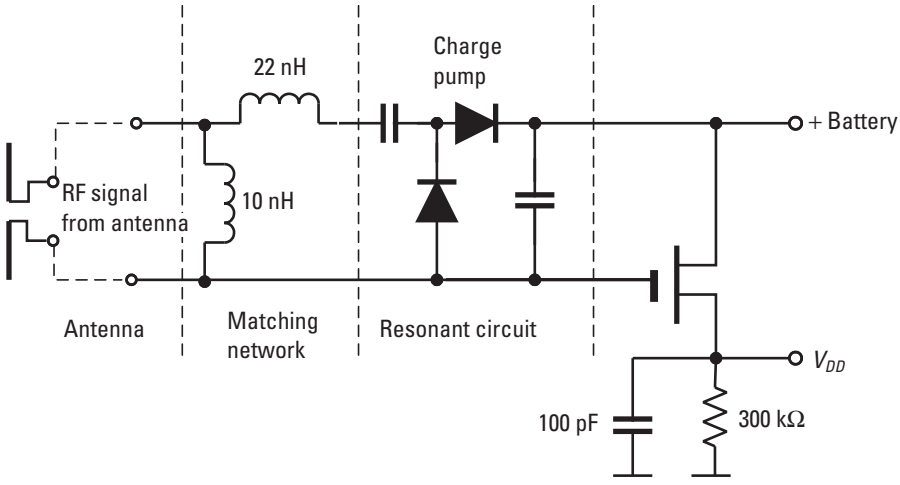


Figure 5.18 Turn-on circuit for the active UHF tag.

5.3.10 Tag Manufacturing Process

The major processes involved in the manufacture of RFID tags are the antenna production process and chip assembly. Both of these activities have gone through serious development in the last few years with the purpose of increasing the throughput and reducing the final cost of the tag.

5.3.10.1 Antenna Production Process

The most popular processes currently on the market for the production of bidimensional antennas are chemical etching and conductive ink printing. *Chemical etching* is an established technology used for the realization of printed electronic circuit boards. To reach the necessary volume requirement, the technology has been modified, adopting a roll-to-roll process where copper tape roll, typically 20 to 40 μm thick, is glued to a polymeric roll substrate, typically PET (polyethylene terephthalate). The copper film is then masked with a photoresist mask and inserted in a chemical bath where the exposed copper is chemically attacked and removed, thereby creating the desired pattern. The photoresist mask is then be removed using a standard process. The strongest advantage of this process, originally invented for the development of flexible circuits, is its availability and well-known manufacturing cost parameters. The currently estimated cost for copper etching is $10 per square meter of produced material that, for 600 antennas per square meter, is equivalent to 1.6 cents per antenna.

Conductive ink printing is certainly the most approachable process on the market. Based on flexographic equipment, the antennas can be printed on a polymer roll in a single step with no need for masking. The major drawback for this technology is the inherent cost of the material used in the process, usually a

conductive ink loaded with 30% silver flake particles, and the higher surface resistivity of the conductive layer, a feature inversely proportional to the final performance of the antenna. Perhaps of even greater concern is the inherent environmental impact of using silver. As RFID tags are used and disposed of, their increasing density within landfills around the world will eventually jeopardize groundwater supplies, leading to requirements for recycling as is presently done for electronics.

5.3.10.2 Chip Assembly

The major challenge in assembling the active component onto the antenna circuit is represented by the small size of the chip itself, the need for a low-temperature attachment process, and the required throughput capacity. Chips for passive RFID tags had their size reduced to submillimeter dimensions in order to optimize the cost of the component, thus increasing the number of chips per wafer. This, of course, creates technical problems when the chip, whose pads are now below the 100-μm size, needs to be connected to the antenna at high speed. To overcome this problem, pick and place (PnP) equipment manufacturers have developed innovative technology enabling the attachment of the chip to the antenna roll using standard flip chip technology and dispensing of a quickly curing conductive adhesive at high speed. The process is quite simple and is capable of attaching about 10,000 components per hour, equivalent to about 70 million tags per year.

Some tag manufacturers have approached the problem from a different point of view, focusing their attention on the production capacity and thus purposefully adopting another process, the *strap attach*. In this solution, the chip is attached to a polymeric carrier film in roll form at high density and high velocity using roll-to-roll equipment. This polymeric carrier has previously been prepared with small caves on the surface to receive the chip, which corresponds to large conductive pads. Proper geometry of the cave and chip die guarantees the proper positioning of the chip on the film. Once this step is completed, it is possible to couple the roll containing the straps with the roll containing the antennas and accomplish the final assembly by using dispensed adhesive. Of course, due to the different density and location of the chip on the strap carrier, each strap needs to be singulated before attachment. The advantage of this process, as stated by its major supporters, is the possibility of assembling chips at a very high velocity. The disadvantage is that the process consists of two-step phases and that the singulation of the strap causes a reduction in the speed of the process. Although it is possible to load the straps at incredible speed, the final assembly of the strap on the antenna is still a process regulated in its throughput by the curing of the adhesive used for attachment.

5.4 Readers

5.4.1 Principles of Operation

RFID tags are interrogated by readers, which in turn are connected to a host computer. In a passive system, the RFID reader transmits an energy field that wakes up the tag and powers its chip, enabling it to transmit or store data. Active tags may periodically transmit a signal, much like a lighthouse beacon, so that data can be captured by multiple readers distributed throughout a facility. Readers may be portable handheld terminals or fixed devices positioned at strategic points, such as a store entrance, assembly line, or toll booth (gate readers.) In addition, readers/interrogators can be mobile; they can have PCMCIA cards to connect to laptop PCs, usually are powered from their own power source (battery) or by the vehicle they are mounted on, and typically have wireless connectivity. The reader is equipped with antennas for sending and receiving signals, a transceiver, and a processor to decode data. Companies may need many readers to cover all of their factories, warehouses, and stores. Readers typically operate at one radio frequency, so if tags from three different manufacturers used three different frequencies, a retailer might have to have multiple readers in some locations, increasing the costs further.

Handheld or portable readers are a very useful resource to supplement fixed readers. Handheld readers can be used instead of a portal reader to record boxes loaded and identify boxes as they are removed; little efficiency is gained relative to barcoded labels, but customer mandates can be accommodated with minimal initial expense.

Handheld readers are also very useful for exception handling of boxes that fail to read at a portal or on a conveyor or that have misplaced or misoriented labels or when identifying boxes of unknown provenance, and so forth. Handheld readers can be useful for inventory cycle counts in storage areas or temporary staging locations, for locating specific cartons in storage, for verifying manifests during assembly, and for specialized applications such as tail-to-tail baggage transfer (moving baggage from one airplane to another in an airport without routing it through the terminal).

RFID readers are used to activate passive tags with RF energy and to extract information from the tag. For this function, the reader includes an RF transmission for receiving and data decoding sections. In addition, the reader often includes a serial communication (RS-232, USB, and so on) capability to communicate with a host computer. Depending on the complexity and purpose of applications, the reader's price range can vary from $10 to a few thousand dollars worth of components and packaging. Typically, the reader is a read-only device, whereas the reader for a read/write device is often called an *interrogator*. Unlike the reader for a read-only device, the interrogator uses command pulses to communicate with a tag for reading and writing data.

The *carrier* is the transmitted radio signal of the reader (interrogator). This RF carrier provides energy to the tag device and is used to detect modulation data from the tag using a backscattering. In read/write devices, the carrier is also used to deliver the interrogator's commands and data to the tag.

The RF transmission section includes an RF carrier generator, an antenna, and a tuning circuit. The antenna and its tuning circuit must be properly designed and tuned for the best performance. Data decoding for the received signal is accomplished using a microcontroller. The firmware algorithm in the microcontroller is written in such a way to transmit the RF signal, decode the incoming data, and communicate with the host computer.

The main criteria for readers include the following:

- *Operating frequency (LF, HF, UHF):* some companies are starting to develop multifrequency readers;
- *Protocol agility:* support for different tag protocols (ISO, EPC, proprietary);
- *Different regional regulations (for example, UHF readers):*
  - UHF frequency agility 902 to 930 MHz in the United States and 869 MHz in Europe;
  - Power regulations of 4W in the United States and 500 mW in some other countries;
  - Manage frequency hopping in the United States and duty cycle requirements.
- *Networking to host capability:*
  - TCP/IP;
  - Wireless LAN (802.11);
  - Ethernet LAN (10base T);
  - RS 485.
- Ability to network many readers together (via concentrators or via middleware);
- Ability to upgrade the reader firmware in the field;
- *Management of multiple antennas:*
  - Typically four antennas per reader;
  - How antennas are polled or multiplexed.
- Adapting to antenna conditions (dynamic auto-tuning);
- Interface to middleware products;
- Digital I/O for external sensors and control circuits.

Certain readers also provide connection options to enable simple process control mechanisms to be implemented, such as digital inputs and outputs with 24V, which can be used to control traffic lights or gates that are released once the tag data has been checked at the goods issue/receipt point. PLC couplings can also be realized using this technology. The higher protocol layers have not been standardized yet, resulting in additional time and effort when it comes to integrating readers across different manufacturers. In addition, readers and antennas at loading gates must be highly tolerant with regard to temperature and must be protected against dust and damp.

Up until the recent surge in developments for the supply chain and EPC tags, readers were used primarily in access control systems and other low-volume RFID applications, which meant that the problem of treating very large numbers of tags and high volumes of data was not such a serious issue. Of course, this is now changing, and many reader manufacturers are starting to develop next generation products to handle the application problems that will be specific to the supply chain and EPC/ISO infrastructure.

5.4.2 Reader Antenna

The reader antenna establishes a connection between the reader electronics and the electromagnetic wave in the space. In the HF range, the reader antenna is a coil (like the tag antenna), designed to produce as strong a coupling as possible with the tag antenna.

In the UHF range, reader antennas (like tag antennas) come in a variety of designs. Highly directional, high-gain antennas are used for large read distances. Regulatory authorities usually limit the maximum power emitted in a given direction; as a result, the transmission power emitted from the reader to the antenna must also be regulated accordingly. One advantage of highly directional antennas is that the reader power often has to be emitted only to the spaces in which the tags that are to read are located.

Generally speaking, physical interdependencies mean that the antenna gain is linked to the antenna size. The higher the gain (or the smaller the solid angle into which the antenna emits), the larger the mechanical design of the antenna will be. It follows, therefore, that highly directional antennas are not used for handheld readers. Antennas typically used for handheld readers include patch antennas, half-wave dipoles, and helix antennas. Larger antenna structures can be used for stationary readers; in the UHF range, they usually take the form of arrays.

All other things being equal, a high-gain antenna will transmit and receive weaker signals farther than a low-gain antenna. Omnidirectional antennas, such as dipole antennas, will have lower gain than directional antennas because they distribute their power over a wider area. Parabolic antennas usually have the

highest gain of any type of antenna but are not really usable in typical RFID applications, except maybe for microwave RFID readers where a long range and narrow radiation pattern is required. A half-wave dipole antenna will have a gain of near unity, or nearly equal the isotropic antenna.

Reader antennas may have different requirements depending on whether they are fixed, portable, or handheld readers. For example, the choice of an antenna for portable devices is dominated by size and weight constraints. Read range and polarization are generally less significant than in the case of fixed readers. Efficient use of RF power to maximize battery life is critical. Highly directive antennas are useful to reduce power consumption, but are generally physically large and thus may not fit in the restricted form factors allowed for portable applications. The fact that handheld reader antennas are small and light constrains the antenna gain.

Antennas that are less than about one-quarter of a wavelength in all dimensions (a quarter wave is about 80 mm [3.2 inches] for UHF operation in the United States) cannot achieve more than about 4 dB of gain. Slightly larger antennas allow up to about 6 dBi of gain, but make the reader somewhat bulky and awkward to carry. The trade-off is important because handheld and portable applications benefit from high antenna gain. The reader is likely to employ less than the maximum allowed transmitting power to improve battery life, so read range is impacted if antenna gain is low. A narrow antenna beam will improve the ability of the user to locate the tag being read by changing the reader orientation and noting the results. The narrow beam of a high-gain antenna, which is undesirable in a stationary-reader application, is often beneficial for a handheld reader, because the user can readily move the beam to cover the area of interest.

5.4.3 Software-Defined Radios in RFID Systems

The problem of continuous change in the EPC market is a vitally important one for all RFID users, and especially those responsible for buying and installing an RFID reader infrastructure. While tags are the consumables of RFID systems, constantly varying, iterating, and regenerating, the RFID reader infrastructure is a deployed capital expense that cannot easily or cost effectively be replaced every time a new tag variant appears. Further, the comings and goings of tags are not neatly synchronized. Generation 1 will not turn into Generation 2 instantaneously; the two generations will coexist, perhaps for as long as few years, and the reality (and hype) surrounding Generation 3 will begin, as well as the introduction of new classes of tag. This type of change is good for the RFID user because it will deliver ever-improving performance and decreasing costs.

Software-defined radio (SDR) uses software for the modulation and demodulation of radio signals. An SDR performs the majority of its signal

processing in the digital domain, most commonly in a digital signal processor (DSP), which is a type of microprocessor specifically optimized for signal processing functions. The advantage of an SDR-based RFID reader is that it can receive and transmit a new form of RFID communication protocol simply by running new software on existing SDR hardware. A software-defined RFID reader consists of an RF analog front end that converts RF signals to and from the reader's antennas into an analog baseband or intermediate frequency signal, and analog-to-digital converters and digital-to-analog converters, which are used to convert these signals to and from a digital representation that can be processed in software running on the reader's digital signal processor.

SDR technology has long been important in the military context, where new radio equipment must interoperate with legacy equipment, much of which is used for many years beyond its design lifetime. Additionally, the U.S. military is often called on to work together with allies that have old, outdated equipment that is incompatible with the more modern U.S. communication hardware. This is exactly analogous to the Generation 1 to Generation 2 (and beyond) transition in RFID. Military SDR projects date back to the early 1990s, and several were fielded in that time frame. Aware of these developments, in 1999 the MIT Auto-ID Center began exploring the idea of using SDR in RFID readers.

5.4.4 Data Transfer Between a Tag and a Reader

5.4.4.1 Signal Transmission

For an RFID system to work, we need three processes: energy transfer, downlink, and uplink. According to this we can divide RFID systems into three groups: full duplex, half-duplex, and sequential. During full duplex and half-duplex operation, the energy is transferred constantly, compared to sequential operation when energy is first transferred by the reader and then the tag responds. In half-duplex systems the information is sent in turns either transferred inductively through load modulation or as electromagnetic backscatter, such as with radar.

In full-duplex systems uplink information is sent on a separate frequency, either a subharmonic or not, so the flow of information can be bidirectional and continuous.

Sequential transfer consists of two phases: First energy is sent to the tag that stores it in a capacitor, then, utilizing the power received, it can function for some time and send its reply. This has the advantage that by extending the charging time and enlarging the capacitor it is possible to acquire more energy for the electronics.

5.4.4.2 Data Transfer Rate

A further influence of carrier frequency is with respect to data transfer, for which it is very important to understand the bit rate (data rate) concept. Whereas in

theory it is possible to transfer binary data at twice the carrier frequency, in practice it is usual to use many cycles of the carrier to represent a binary digit or group of digits. However, in general terms, the higher the carrier frequency, the higher the data transfer rate that can be achieved. So, a low-frequency system operating at 125 kHz may transfer data at a rate of between 200 and 4,000 bps depending on the type of system, while rates up to greater than 100 Kbps (but typically less than 1 Mbps) are possible for microwave systems. It should also be appreciated that a finite bandwidth is required in practice to transfer data, this being a consequence of the modulation that is used. Consequential to transfer capability is the data capacity of the tag. Loosely speaking, the lower the frequency, the lower the data capacity of the tags, simply because of the amount of data required to be transferred in a defined time period. Keep in mind that the capacity can also be determined by the manner in which the tag is designed to be read or written to (for read/write tags), be it in total or part. The choice of data transfer rate has to be considered in relation to system transfer requirements—this is determined by the maximum number of tags that may be expected to be read in a unit interval of time multiplied by the amount of data that is required to be read from each tag. Where a write function is also involved, the number of tags and write requirements must also be considered.

ISO 15693 is an ISO standard for *vicinity cards*, that is, cards that can be read from a greater distance as compared to *proximity cards*. ISO 15693 systems operate at the 13.56-MHz frequency, and offer a maximum read distance of 3 to 4 feet. In ISO 15693 chips, the subcarrier frequency is equal to 423.75 kHz (RF/32) with FSK or OOK modulation and Manchester data coding. The achievable label data transfer rate is up to a relatively fast 26.48 Kbps. Most typical bit rate values in bits per second are RF/8, RF/16, RF/32, RF/40, RF/50, RF/64, RF/80, RF/100, and RF/128. Every tag sends back information with some predefined, usually fixed bit rate. Once a manufacturer programs the data rate, it usually cannot be changed. This data rate is clocked by internal tag frequency. For LF transponders the range is from 100 to 150 kHz, depending on the manufacturer. Consider, for example, a transponder type for which the bit rate is RF/32. This means that the data rate is 32 field clocks (FC) per logic 1 or 0 data bit. The data (bit) rate is a bit time duration and it is defined as field clocks per bit. Taking a field clock equal to 125 kHz and a tag bit rate equal to RF/32, the data rate is 125 kHz/32 = 3.9062 Kbps, so receiving 64 bits of information would take $8\,\mu s \times 32 \times 64 = 16.384$ ms.

The manner in which the tags are interrogated is also important. It can be done singularly (one at a time in the interrogation zone) or as a batch (a number of tags in the interrogation zone at the same time). The latter requires that the tags and associated system have anticontention (anticollision) facilities so that collisions between responses from tags in the zone at the same time can be resolved and contention avoided. Various anticontention protocols have been

devised and applied with various levels of performance with respect to the number of tags that can be handled and the time required to handle them. So, the anticontention performance may be an important consideration in many applications.

5.4.4.3 Read/Write Range

The read/write range is the communication distance between the reader (interrogator) and tag. Specifically, the read range is the maximum distance to read data out from the tag, and the write range is the maximum distance to write data from the interrogator to the tag. The read/write range is, among other effects, mainly related to:

- Electromagnetic coupling of the reader (interrogator) and tag antennas;
- The RF output power level of reader (interrogator);
- Carrier frequency bands;
- The power consumption of the device;
- Antenna orientation;
- The distance between the interrogator and the tag;
- Operating environment conditions (metallic, electric noise, multiple tags, multiple readers, and so on);
- The tag and the tag's dwell time.

The tag's dwell time is the time a tag is in the interrogator's RF field. An RFID interrogator's read range is the distance between the interrogator and the RFID tag at which the signals from the tag can be read properly. Similarly, an RFID interrogator's write range is the maximum distance at which information within the RF signal from the interrogator can be received correctly and stored within the memory of the tag's microchip. More power is needed to write to a tag than to read it; as a result, the tags need to be closer to the antenna to write than to read. The general rule is that the write range is 50% to 70% of the read range of a particular interrogation zone.

Power limitations, as listed in Table 5.3, are imposed by local authority and cannot be chosen arbitrarily. The standardization of RFID technology and the requirements of the local governing bodies are still in progress and change constantly. For that reason, some of the information provided in this book that was correct during the preparation of the manuscript might change by the time it reaches the reader.

The electromagnetic coupling of the reader and tag antennas increases, using a similar size of antenna with high Q on both sides. The read range is improved by increasing the carrier frequency. This is due to the gain in the

Table 5.3
RFID Power Limitations Based on the Region and Frequency

Frequency Band	Power, Limitations, and Region
125 kHz	Inductively coupled RF tags
1.95, 3.25, and 8.2 MHz	Inductively coupled theft tags, worldwide
13.56 MHz	Inductively coupled RFID tags, worldwide
27 and 40 MHz	0.1W ERP, Europe
138 MHz	0.05W ERP, duty cycle < 1%, Europe
402–405 MHz	Medical implants, 25 μW ERP (–16 dBm)
433.05–434.79 MHz	25 mW ERP, duty cycle < 10%, Europe
468.200 MHz	0.5W ERP, Europe
869.40–869.65 MHz	0.5W ERP, duty cycle < 10%, Europe*
902–928 MHz	4W EIRP, America
2400–2,483.5 MHz	ISM band, 0.5W EIRP Europe; 4W America, Bluetooth
5,725–5,875 MHz	25 mW EIRP

Note: EIRP = equivalent isotropic radiated power; ERP = equivalent radiated power.
*To accommodate concerns over the ability of Gen 2 RFID systems to perform under the European regulations, Europe has already increased its available frequency spectrum from 2 to 8 MHz, allowable power output level from 0.5W to 2W, and replaced its 10% duty-cycle restriction with a listen-before-talk requirement. Even with these improvements, work is still under way to further alleviate European regulatory constraints.

radiation efficiency of the antenna as the frequency increases. However, the disadvantage of high-frequency (900-MHz to 2.4-GHz) application is shallow skin depth and narrower antenna beam width causing less penetration and more directional problems, respectively. Low-frequency application, on the other hand, has an advantage in the penetration and directivity, but a disadvantage in the antenna performance. Read range increases by reducing the current consumption in the silicon device. This is because the LC antenna circuit couples less energy from the reader at further distances. A lower power device can make use of less energy for the operation.

For LF and HF (near-field) systems, to increase the magnetic field at the tag's position, the reader/writer antenna coil's radius must be increased, or the current in the antenna coil must be increased, or both. The strength of the magnetic field is attenuated in proportion to the inverse of the cube of distance. By increasing the diameter of the RFID tag's antenna coil, the signal induced in the tag's coil can be increased. Accordingly, for applications that require long-range

operation, the reader/writer antenna coil's radius and tag antenna coil's dimensions must be increased. In tests comparing coin-sized and IC card-sized tags using the same reader/writer, the IC card-sized tag had an operating area several times larger than the coin-sized tag.

The read range of a UHF-based RFID (propagation) system can be calculated by the Friis free-space equation as follows:

$$r = \frac{\lambda \cos\theta}{4\pi} \sqrt{\frac{P_R G_R G_T \left(1 - (\Delta\rho)^2\right)}{P_{th}}} \quad \text{for} \quad 0 \le (\Delta\rho)^2 \le 1 \qquad (5.43)$$

where, G_T is the gain of the tag antenna, $P_R G_R$ is the EIRP for the reader, λ is the wavelength, P_{th} is the minimum threshold power required to power an RFID tag, θ is the angle made by the tag with the reader plane, and $(\Delta\rho)^2$ is the power reflection coefficient, which is the ratio of reflected power to incident power by the tag. Note that the power received by the tag is inversely proportional to the square of the distance between the tag and the reader's antenna. Studies reveal that the orientation of the tag in the RF field affects its read range. In the specific context of a directivity pattern, a perfectly parallel tag, relative to the reader's antenna, yields the maximum read range, whereas a tag perpendicular to the base station antenna's field has minimum to zero read range. Thus, efforts are made to make the tag parallel to the reader antenna by deploying one or more of the following measures:

- Change in orientation of the reader antenna to suit the orientation of the tag antenna;
- Use of redundant antennas for ensuring proper alignment of at least one reader antenna to the tag antenna;
- Increase reader antenna power (of course, within the limits allowed by the local authorities) to reduce the effect of tag orientation;
- Increase the polling rate of the antenna to make more reads in the same sampling time.

The far-field formula is correct, subject to the assumption that the polarization of the reader antenna and the polarization of the tag antenna are perfectly matched. However, in fact, the polarization mismatch is essential and required in most RFID applications. The point is that in the majority of applications the tag is allowed to appear in an almost arbitrary position in the field of the reader antenna while the polarization of the tag antenna is usually linear because of the prerequired small size of the tag. In such a situation, the only way

to fulfill a system requirement is to use a circularly polarized reader antenna. Thus, a sacrifice of 3-dB power loss (at least, although it can be even much higher; see the discussion on polarization mismatch in Chapter 2) due to a polarization mismatch between a circularly polarized reader antenna and a linearly polarized tag antenna overcomes the problem of tag orientation. This is why, nowadays, the major vendors offer mainly circularly polarized reader antennas. At the same time, the linearly polarized antennas are also available in the market for limited RFID applications. In the case of linearly polarized reader and tag antennas, the substantial polarization misalignment may cause a severe power loss, which in its turn can potentially lead to a fault on the part of the RFID system. In the case of a circular-to-linear polarization mismatch the read range r will be $\sqrt{2}$ times shorter than the one calculated by (5.43).

As we can see from the Table 5.4, systems operating in the 915-MHz band may achieve read ranges of 20 feet (6m) or more under current FCC regulations.

5.4.4.4 Environment and Proximity to Other Objects

Up to now, our considerations have focused on data transfer across an uncluttered air interface. However, free-space propagation in which the reader and tag are distanced from any obstructions or other tags, and perfectly aligned relative to each other, is not a realistic situation. In practice, the region between the tag and the interrogator may contain obstacles and materials that can influence the performance of the system. The carrier frequency is one of significance with respect to the effects that the prevailing conditions and clutter factors (obstacles and physical structures) can have. In low- and high-frequency inductive RFID systems, the magnetic field is effectively used to couple data, and such fields are largely unaffected by dielectric or insulator materials (papers, plastics, masonry, and ceramics, for example); the field simply penetrates the materials. Where metals are involved, they can distort the field, depending on how ferrous they are. This will weaken the field strength in the regions of the interrogation zone, in some cases to the extent that system performance is impaired. The range capability may be impaired or the ability to read or write to a tag may be impaired. For uncompensated tag designs operating at resonant frequencies, the presence of metals can often detune the device, in some cases preventing its operation.

At higher frequencies (UHF and above), where, for propagation RFID, the electric component of a field becomes more significant; the higher the frequency the more easily they can penetrate dielectric materials. However, for some materials where energy exchange mechanisms can be identified at or near the carrier frequency, this can result in energy absorption from the propagating wave, hence causing an impairment of range performance. One of the challenges with UHF RFID tags is efficient operation in the presence of water or metal.

Table 5.4
Read Range for Different UHF Reader Powers and Reflection Coefficients

f [MHz]	λ [m]	*P* Reader [W]	*P* Reader [dBm]	Reader Antenna Gain	Tag Antenna Gain	$\Delta\rho$	Angle [°]	Tag Threshold Power [dBm]	Tag Threshold Power [mW]	Read Range [m]
915.00	0.33	0.50	26.99	1.64	1.64	0.40	0.00	−10.00	0.10	2.7731
915.00	0.33	0.50	26.99	1.64	1.64	0.50	0.00	−10.00	0.10	2.6203
915.00	0.33	0.50	26.99	1.64	1.64	0.60	0.00	−10.00	0.10	2.4205
915.00	0.33	2.44	33.87	1.64	1.64	0.40	0.00	−10.00	0.10	6.1259
915.00	0.33	2.44	33.87	1.64	1.64	0.50	0.00	−10.00	0.10	5.7884
915.00	0.33	2.44	33.87	1.64	1.64	0.60	0.00	−10.00	0.10	5.3471

Unfortunately, the human body is made up of mostly water; thus, if the RFID tag is placed close to the human body, performance will suffer.

As far as metals are concerned, they reflect or scatter these higher frequency signals, depending on the size of the metal object in relation to the wavelength of the incident signal. Such effects can impair the range that can be achieved and, in some cases, can screen the reader from the tag and prevent it from being read. Any metal near the tag, such as keys or coins, can also cause the tag to be undetectable.

The proximity of tags may also exhibit a similar effect. Reflections and diffraction effects can often allow pathways around metal objects within an interrogation zone. Because it is difficult to generalize on the effects of clutter within the interrogation zone, it is expedient where possible to avoid clutter and choose a carrier frequency that is appropriate to the conditions to be expected.

Passive RF tags in the UHF and microwave bands have drawn considerable attention because of their great potential for use in many RFID applications [10]. However, more basic research is needed to increase the range and reliability of a passive RF tag's radio link, particularly when the RF tag is placed onto any lossy dielectric or metallic surface. This radio link budget is dependent on the *gain penalty losses* (L_{GP}), a term that quantifies the reduction in RF tag antenna gain due to material attachment.

After combining (5.16) and (5.21), we get the following expression:

$$P_{REC} = \frac{P_R G_R^2 \lambda^2 \sigma}{(4\pi)^3 r^4} = \frac{P_R G_R^2 G_T^2 \lambda^4 (\Delta\rho)^2}{(4\pi)^4 r^4} \tag{5.44}$$

The assumption in this case is that the gains of the reader's transmitting and receiving antennas are the same, which may not always be the case; the reason is that the single-antenna readers are inexpensive and compact but need excellent matching circuits, a high-isolation coupler with extremely high isolation between ports, and electronic circuitry with a wide dynamic range. In the logarithmic form, the same expression for the backscattered power received at the reader looks like this:

$$P_{REC} = P_R + 2G_R + 2G_T + 20\log(\Delta\rho) + 40\log\left(\frac{\lambda}{2\pi}\right) - 40\log r \tag{5.45}$$

where $\Delta\rho$ is a reflection change between switched loads.

Now, we can include in (5.44) additional losses due to the antenna being attached to different types of materials in the form of an adjustment for on-object degradation. In doing so, we get:

$$P_{REC} = P_R + 2G_R + 2G_T + 20\log(\Delta\rho) + 40\log\left(\frac{\lambda}{2\pi}\right) - 40\log r - 2L_{GP} \tag{5.46}$$

A series of measurements was used to measure the far-field gain pattern and gain penalty of several flexible 915-MHz antennas when attached to cardboard, pine plywood, acrylic, deionized water, ethylene glycol, ground beef, and an aluminum slab. It has been shown that the gain penalty due to material attachment can result in more than 20 dB of excess loss in the backscatter communication link.

From the reader's perspective, handheld and portable antennas are very likely to operate in proximity to people's hands and arms, as well as other obstacles. The amount of power reflected from the antenna, measured by its reflection coefficient or return loss, should ideally be unaffected by such obstacles unless they are actually within the antenna beam. The best return loss performance in the presence of near-field objects is usually obtained from balanced antennas, in which the two halves of an antenna are driven by precisely opposed currents and there is no large ground plane. However, such antennas are relatively large compared to single-ended (nonbalanced) antennas and require a balanced–unbalanced transformer (balun) to connect them to ground-referenced antenna cables or circuit board connectors. With a balun, the antenna is less sensitive to the presence of near-field objects.

5.4.5 UHF Reader Electronic Circuitry

To shrink the size of the RF portion of an RFID reader, it is necessary to increase the functions in each element. Figure 5.19 shows a typical block diagram for an RFID reader and shows one possible way of integrating elements into a chipset. Each module is briefly described in the following sections.

5.4.5.1 UHF Reader Source Module

The purpose of the *source module* is to provide a synthesized *local oscillator* (LO) for transmitting (Tx) and receiving (Rx) paths in an RFID reader. The updated FCC standard requires frequencies to be within 10 ppm over the operating temperature ranges. It is necessary to amplify the signal after the synthesizer, in order to provide adequate LO input to the Tx and Rx signal paths due to typical synthesizer output powers and the loss of the power divider. For a source module, it is critical that the part be adaptable so that a single PC-board footprint can be used to handle all of the different bands. Using an integrated synthesizer/voltage-controlled oscillator (VCO) IC, it is possible to center the VCO bands by using different inductor values. The Japanese band requires a faster

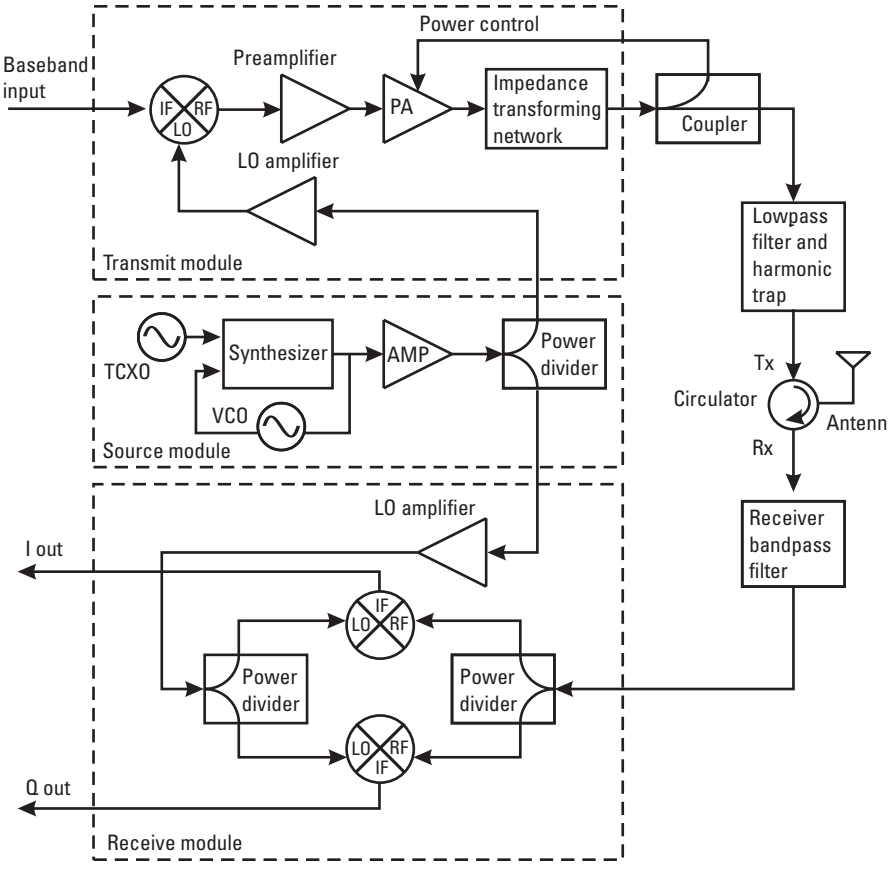


Figure 5.19 UHF RFID chipset block diagram.

switching speed than the U.S. and European bands, which can still be realized with a 5-kHz bandwidth loop filter, but with different component values. It is desirable to have isolation on the order of 20 dB in the power divider. For cost reasons, monolithic narrowband power dividers are generally used and are not optimal for covering 850 to 960 MHz. To optimize the isolation for each, tuning inductors and/or capacitors are used to recenter the power divider isolation. To shrink the size and reduce the overall component count, it is necessary to combine somewhat diverse parts to create a source module. An additional requirement that is typical of synthesizer/source modules is that shielding is required for loop stability and minimization of phase noise.

5.4.5.2 UHF Reader Transmitting Module

As depicted in Figure 5.19, a typical transmitting module would include a double balanced modulator (DBM), LO amplifier, preamplifier, power amplifier, and

impedance transforming network (ITN). The high level of integration, with over 50 dB of available small-signal gain, requires careful module layout. To maintain stability, it is necessary to keep the preamplifier located as far as possible from the power amplifier. The DBM provides a means to modulate the carrier signal. An LO amplifier is included to raise the signal available from the source module to a level sufficient to drive the mixers. Additionally, having the LO amplifier provide a 50Ω interface allows for simple interconnection to the source module. The modulated RF output from the mixer goes to a preamplifier and then to a power amplifier. The preamplifier has a gain of 17 dB and the power amplifier, implemented as a three-stage device, provides a small signal gain of 35 dB.

Also included in the transmit module is an impedance transforming network (ITN). The purpose of this network is to transform the 50Ω load impedance to a level that the power amplifier needs to drive in order to produce the desired output power at the available supply voltage. For a typical supply of 3.6V, this impedance is only a few ohms, creating large circulating currents. These low impedances necessitate proper handling of circuit parasitics. This circuit requires careful design and implementation from performance and reliability perspectives. To be able to provide the desired 1W RF level at the antenna terminals typical for UHF readers, the power amplifier needs to be capable of providing sufficient power output capability to overcome the signal losses introduced between the transmitting module output and the antenna. These losses would include any coupler, filter, circulator, connector, and cabling used in the path to the antenna. It is desirable to control the power in order both to set the output level to various requirements and to implement a commonly used form of carrier amplitude modulation called *pulse-interval modulation*, which is used to interrogate tags. The modulation bandwidth must be sufficient for the intended data rates without significant distortion, but as much circuitry as possible should be broadband.

The transmitter transmits encoding data with ASK modulation, including DSB-ASK, and SSB-ASK for forward link, and sends an unmodulated carrier for the return link. The maximum output power from the PA is restricted to 30 dBm (1W).

5.4.5.3 UHF Reader Receiving Module

Most modern wireless communication systems use digital modulation/demodulation techniques, and there is a good reason for this. They provide increased channel capacity and greater accuracy in the transmitted and received messages in the presence of noise and distortion. In digital communication systems, a finite number of electrical waveforms or symbols are transmitted, where each symbol can represent 1 or more bits. It is the job of the receiver to identify which symbol was sent by the transmitter even after the addition of noise and distortion. Distortion in wireless communication can be caused by several things

such as passing a signal through filters having insufficient bandwidth or inefficient switching of nonlinear elements. Ultimately, the effects of such events within communication systems are termed *intersymbol interference* (ISI). In addition to ISI, there are other types of distortion more notably termed *delay spread* and noise. Delay spread occurs when multiple versions of the same signal are received at different times. This occurs when the transmitted signal reflects off multiple objects on its way to the receiver (multipath). System designers are focusing their attention on their transceivers in search of a method or components that might help them achieve a superior signal-to-noise ratio, resulting in a lower bit error rate. It is widely projected that one of the reasons for the delay in wide-scale adoption of RFID systems has been the unacceptable bit error rate of RFID tag reading.

In addition, RFID systems operating in the UHF band have unique attributes; during operation, the reader antenna emits electromagnetic energy in the form of radio waves that are directed toward an RFID tag. The tag absorbs energy, and through its built-in microchip/diode, uses it to change the load on the antenna, which in turn reflects an altered signal to the reader. This method is known as backscatter and is the basis by which a passive RFID tag identifies its presence. These backscattered signals are essentially at the same frequency as that of the transmitted signal. The backscattered signal antenna received is sent to the receiver through a directional coupler. The receiver front end must be designed to withstand high-interference signal levels without introducing significant distortion spurs. The receiver noise needs to be low enough that the system has sufficient dynamic range to allow error-free detection of low-level responding tag signals.

Homodyne detection, whereby a sample of the transmitted signal prior to modulation is used as the LO source for the receiving I/Q demodulator, is utilized. Having both the transmitted and received signals at the same frequency exacerbates the difficulty of recovering the weak reflected signal, because it has to be identified in the presence of the higher powered carrier frequency. Consequently, it is an advantageous to choose transceiver components that help improve the overall signal-to-noise ratio as well as minimize LO carrier leakage. The I/Q demodulator is a key element that can be used to maximize the signal-to-noise ratio and to minimize LO carrier leakage. Direct conversion to baseband frequency with the lowest bit error rate and the highest sensitivity possible is crucial, not only for reader accuracy, but also to its range of usage.

5.5 RFID Power Sources

RFID tags need power to sense, compute, and communicate, which is further classified into three categories: storage (batteries, capacitors), energy harvesting

mechanisms (vibrations/movement, photovoltaic, thermal gradient, and so on), and energy transfer (inductive coupling, capacitive coupling, backscatter). Because many of these devices are expected to operate with a minimum of human intervention, optimizing power consumption is a very important research area. RFID tags may derive the energy to operate either from an on-tag battery or by scavenging power from the electromagnetic radiation emitted by tag readers. *Storage* refers to the way devices store power for their operation, done by means of batteries or capacitors. Batteries are used when a longer life is required, and capacitors are used in applications that require energy bursts for very short durations [11].

5.5.1 Power Harvesting Systems

Power harvesting (sometimes termed *energy scavenging*) is the process of acquiring energy from the surrounding environment (ambient energy) and converting it into usable electrical energy; the self-winding watch is a historical example of a power-harvesting device. The watches were wound by cleverly extracting mechanical energy from the wearer's arm movements. In medical devices, for example, a patient's normal daily activities could power an implantable pump that delivers insulin to a diabetic. The use of piezoelectric materials to harvest power has already become popular. Piezoelectric materials have the ability to transform mechanical strain energy into electrical charge. Piezo elements are being embedded in walkways to recover the "people energy" of footsteps. They can also be embedded in shoes to recover walking energy.

Energy transfer is the way by which passive RF devices are powered. The energy transfer mechanisms are inductive coupling, capacitive coupling, and passive backscattering. Inductive coupling is the transfer of energy between two electronic circuits due to the mutual inductance between them. Similarly, capacitive coupling is the transfer of energy between two circuits due to the mutual capacitance between them. Passive backscattering is a way of reflecting back the energy from one circuit to another.

Passive RFID tags obtain their operating power by harvesting energy from the electromagnetic field of the reader's communication signal. The limited resources of a passive tag require it to both harvest its energy and communicate with a reader within a narrow frequency band as permitted by regulatory agencies. A passive tag's power comes from the communication signal either through inductive coupling or far-field energy harvesting. Inductive coupling uses the magnetic field generated by the communication signal to induce a current in its coupling element (usually a coiled antenna and a capacitor). The current induced in the coupling element charges the on-tag capacitor that provides the operating voltage, and power, for the tag. In this way, inductively coupled

systems behave like loosely coupled transformers. Consequently, inductive coupling works only in the near field of the communication signal.

Far-field energy harvesting uses the energy from the interrogation signal's far-field signal to power the tag. The signal incident on the tag antenna induces a voltage at the input terminals of the tag. This voltage is detected by the RF front-end circuitry of the tag and is used to charge a capacitor that provides the operating voltage for the tag. In the far field, tag-to-reader communication is achieved by modulating the RCS of the tag antenna (backscatter modulation.)

There is a fundamental limitation on the power detected a distance away from a reader antenna. In a lossless medium, the power transmitted by the reader decreases as a function of the inverse square of the distance from the reader antenna in the far field. A reader communicates with and powers a passive tag using the same signal. The fact that the same signal is used to transmit power and communicate data creates some challenging trade-offs. First, any modulation of the signal causes a reduction in power to the tag. Second, modulating information onto an otherwise spectrally pure sinusoid spreads the signal in the frequency domain. This spread, referred to as a *sideband,* along with the maximum power transmitted at any frequency, is regulated by local government bodies in most parts of the world. These regulations limit the rate of information that can be sent from the reader to the tag. RFID systems usually operate in the free ISM bands, where the emitted power levels and the sideband limits tend to be especially stringent.

The signaling from the tag to the reader in passive RFID systems is not achieved by active transmission. Because passive tags do not actively transmit a signal, they do not have a regulated limit on the rate of information that can be sent from the passive tag to the reader. Passive tags obtain impinging energy during reader interrogation periods, and this energy is used to power tag ICs. In the near field, tag-to-reader communication is achieved by modulating the impedance (load modulation) of the tag as seen by the reader [12]. For the maximum reading range, one has to ensure maximum power transfer efficiency from the reader to the tag. What makes the problem challenging is that in the case of an inductively coupled reader tag, the reader must deal with a changing effective load due to the location-dependent mutual coupling effect between the reader and tag as well as the unpredictable number of tags in the read zone of the reader.

The powering of and communication with passive tags with the same communication signal places restrictions on the functionality and transactions the tags are capable of. First, very little power is available to the digital portion of the integrated circuit on the tag, thus limiting the functionality of the tag. Second, the length of transactions with the tag is limited to the time for which the tag is expected to be powered and within communication range. Governmental regulations can further limit communication timings. In the United States'

915-MHz ISM band, regulations require that, under certain operating conditions, the communication frequency change every 400 ms. Because every change in frequency may cause loss of communication with a tag, transponders must not be assumed to communicate effectively for longer than 400 ms. Finally, it is important to minimize state information required in passive tags. In many practical situations, power supplied to the tag may be erratic, and any long-term reliance on state in the tag may lead to errors in the operation of a communications protocol.

5.5.2 Active Power Sources

5.5.2.1 Batteries

Battery-assisted backscatter tags have their own power source to preenergize the silicon chip. The data is otherwise sent and received from the reader in the same way as a passive tag. This is of benefit when many tags are present in an interrogation zone; if they are all passive, they all need a lot of energy initially to reach sufficient voltage to turn on. With metals and fluids near tags, this is even harder due to interference and blind spots in the field. An on-board power source on each tag helps to overcome this.

Primary lithium has been a favorite option in this market, because the chemistry offers several positive factors including high-energy density, long life (approximately 10 years), and long storage life. Additionally, this chemistry is ideal for RFID tag applications because it is lightweight. For RFID tag systems, primary lithium/manganese dioxide (Li-MnO_2) and lithium-thionyl chloride (Li-$SOCL_2$) are the two types of batteries that are most common. These lithium batteries offer a set of performance and safety characteristics that is optimal for RFID tag applications. Li-MnO_2 is relatively safe, compared to volatile lithium batteries, such as lithium-sulfur dioxide (Li-SO_2) and lithium-thionyl chloride (Li-$SOCL_2$), and does not develop any gas or pressure during battery operation.

However, one main disadvantage is that a single Li-MnO_2 cell cannot operate at voltages greater than 3.0V. These are typical in high-pulse applications, which Li-SO_2 and Li-$SOCL_2$ can satisfy. Li-MnO_2 cells are best suited for applications that have relatively high continuous or pulse current requirements. However, because most electronic components used in RFID tags require a minimum operating voltage of 3V, at least two Li-MnO_2 cells must be connected in series to ensure a proper margin of safety for system reliability. This requirement adds weight and cost while potentially decreasing reliability due to increased part count.

Overall, the Li-MnO_2 chemistry has a high energy density and has the ability to maintain a high rate of discharge for long periods of time. It can be stored for a long time (typically between 5 and 10 years) due to its low self-discharge rates. It also has the capability to supply both pulse loads and

maintain a constant discharge voltage. Li-MnO_2 cells can operate in temperatures ranging from −20°C to +70°C, although storage in temperatures exceeding +55°C is not highly recommended, and operation will be below full energy capacity at low temperatures. Their nominal voltage is typically 3.0V, which is 2 times the amount of that found in alkaline manganese batteries.

Li-$SOCL_2$ is a low-pressure system that is considered superior to lithium-sulfur dioxide systems in terms of high-temperature and/or unusual form-factor applications. Due to its low self-discharge rate, Li-$SOCL_2$ has a maximum shelf life of 10 to 15 years. This service life is the same for all construction, whether cylindrical, coin, or wafer. This chemistry also has the highest open circuit voltage of 3.6V.

For most applications, only one cell of Li-$SOCL_2$ is required to maintain sufficient operating voltage. This is true as long as one cell can provide enough current to uphold the operating lifetime. RFID tag applications require very low continuous current and moderate pulse current, which Li-$SOCL_2$ batteries have no problem providing.

5.5.2.2 Other Power Sources

In the design of mobile electronics, power is one of the most difficult restrictions to overcome, and current trends indicate that this will continue to be an issue in the future. Designers must weigh wireless connectivity, CPU speed, and other functionality versus battery life in the creation of any mobile device. Power generation from the user may alleviate such design restrictions and may enable new products, such as batteryless on-body sensors. Power may be recovered passively from body heat, arm motion, typing, or walking or actively through user actions, such as winding or pedaling. In cases where the devices are not actively driven, only limited power can generally be scavenged (with the possible exception of tapping into heel-strike energy) without inconveniencing or annoying the user. That said, clever power management techniques combined with new fabrication and device technologies are steadily decreasing the energy needed for electronics to perform useful functions, providing an increasingly relevant niche for power harvesting. Current and historical devices have shown that such mechanisms can be practical and desirable, yet much work remains in the creation and exploitation of these microgenerators.

RFID technology was originally thought to be a passive technology because the tags had no batteries; they just collected energy from the reader and sent back their information (limiting in this way the distance between the reader and the tags). New advancements in the technology, however, have allowed the development of enhanced tags (active RFID) whose function fills the gap between the RFID traditional field and wireless sensor networks field. Recent research has revealed nuclear power as a possible source of power for wireless sensor and RFID networks [13]. Although certainly a little bit frightening at

first thought, note that the isotopes used in the actual prototypes penetrate no more than 25 μm in most solids and liquids, so in a battery they could be safely contained by means of a simple plastic package. (Most smoke detectors and some emergency exit signs already contain radioactive material.) The huge amount of energy these devices can produce is illustrated by these figures: The energy density measured in mWh/mg is 0.3 for a lithium-ion battery, 3 for a methanol-based fuel cell, 850 for a tritium-based nuclear battery, and 57,000 for a polonium-210 nuclear battery. The current efficiency of a nuclear battery is around 4%, and current research projects (e.g., as part of the new DARPA program called Radio-Isotope Micropower Sources) aim at 20%. To make a little more sense out of these figures, for example, with 10 mg of polonium-210 (contained in about 1 mm^3) a nuclear battery could produce 50 mW of electric power for more than 4 months.

5.6 Review Questions and Problems

1. The reader produces a magnetic field that triggers the tag as shown in Figure 5.20. When the reader receives the transmitted data, it interprets the data and takes appropriate action. When the transponder enters the field produced by the reader, the coil produces a voltage inside the tag. In a passive transponder, this voltage can be used to power the tag. In an active transponder, the voltage is used to wake the tag and use its internal battery. Active transponders generally have longer read distances, have a shorter operational life, and are larger and more costly to manufacture. Passive transponders are generally smaller, have a longer life, and are less expensive to manufacture. For optimum performance, the transponder coil is used in a parallel LC circuit designed to resonate at the operating frequency of the reader.

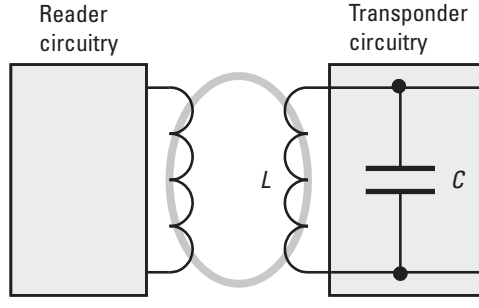


Figure 5.20 RFID system and a resonant-frequency calculation.

a. Calculate the capacitor value for a 4.9-mH transponder coil operating at 125 kHz. (*Answer:* $C = 331$ pF.)
b. What would be the resonant frequency f_1 if the overlapped tags have a new total inductance of 5.5 mH? (*Answer:* $f_1 = 117.96$ kHz.)

2. What do you think about the idea of passive RFID devices for locating small children?
3. You want an RFID tag that supports longer distance communications and does not rely on the reader to provide power to the tag. What kind of tag do you need?
4. Are there any health risks associated with RFID and its radio waves? Discuss.
5. Formula for EIRP in dBm:

$$\text{EIRP} = \frac{E^2 \cdot r^2}{30\Omega}$$

where EIRP is in watts, E is in volts per meter, and r is in meters.

a. Show the EIRP in dBm, using E in dBμV/m and r in meters.
(*Answer:* $\text{EIRP}_{[\text{dBm}]} = E_{[\text{dB}\mu\text{V/m}]} + 20 \log r_{[\text{meters}]} - (10 \log 30 + 90)_{[\text{dB}]}$.)

b. In standard test setups, the electrical field strength is often measured at a distance of 3m. Show that in this case we can use the simplified formula:
$\text{EIRP} = E_{[\text{dBm}]} = E_{[\text{dB}\mu\text{V/m}]} - 95.23_{[\text{dB}]}$

6. The chip device turns on when the antenna coil develops 4 V_{PP} across it. This voltage is rectified and the device starts to operate when it reaches 2.4 V_{DC}.

a. Calculate the B-field to induce a 4-V_{PP} coil voltage with an ISO Standard 7810 card size (85.6 × 54 × 0.76 mm), using the coil voltage equation. Frequency = 13.56 MHz, number of turns is 4, the Q of the tag antenna coil is 40, and $\cos \alpha = 1$ (normal direction, $\alpha = 0°$).
b. Calculate the induced voltage, assuming that the frequency of the reader was 1,000 Hz off from the resonant frequency of the tag. What conclusion can you make from this calculation?

$$V_{Tag} = 2\pi f NSQB \cos\alpha$$

From here we have:

$$B = \frac{V_{tag}}{2\pi f NSQ \cos\alpha}$$

Tag coil size = (85.6×54) mm^2 (ISO card size) = 0.0046224 m^2

$$B = \frac{4/\sqrt{2}}{2\pi \cdot 13.56 \cdot 10^6 \cdot 4 \cdot 4.6 \cdot 10^{-3} \cdot 40 \cdot 1}$$

$$B = 0.045[\mu T] *$$

For the reader frequency offset, instead of frequency f, we use the following expression to calculate f_1:

$$f_1 = \frac{f_0}{1+\Delta f} = \frac{13.56 \text{ MHz}}{1+10^{-3}} = 13.546 \text{ MHz} \tag{5.47}$$
$$V_{tag} = 2\pi f NSQB \cos\alpha = 2.83\,[V]$$

**Note:* The tesla (symbol T) is the SI-derived unit of magnetic flux density (or magnetic induction). It is used to define the intensity (density) of a magnetic field. It is named in honor of world-renowned inventor, scientist, and electrical engineer Nikola Tesla. The tesla, equal to 1 weber per square meter, was defined in 1960.

7. The use of the elecromagnetic field for energy scavenging has been considered [14]. Research and calculate how far you have to be from a cellular base station in order to achieve successful energy scavenging. Are there any other urban areas offering a similar level of electromagnetic field sufficient for energy scavenging?

References

[1] Lehpamer, H., *Microwave Transmission Networks; Planning, Design, and Deployment,* New York: McGraw-Hill, 2004.

[2] Jiang, B., "Energy Scavenging for Inductively Coupled Passive RFID Systems," *IMTC Instrumentation and Measurement Technology Conference,* Ottawa, Canada, May 2005.

[3] Karthaus, U., and M. Fischer, "Fully Integrated Passive UHF RFID Transponder IC with 16.7-μW Minimum RF Input Power," *IEEE J. of Solid-State Circuits,* Vol. 38, No. 10, October 2003.

[4] Microchip, "13.56 MHz RFID System Design Guide," 2004.

[5] Swamy, G., and S. Sarma, "Manufacturing Cost Simulations for Low Cost RFID Systems," white paper, Cambridge, MA: Auto-ID Center, Massachusetts Institute of Technology, February 2003.

[6] Yang, L., et al., "Design and Development of Novel Inductively Coupled RFID Antennas," Atlanta, GA: School of Electrical and Computer Engineering, Georgia Institute of Technology, 2006.

[7] Felber, P., "Fractal Antennas," project, ECE 576, Chicago: Illinois Institute of Technology, December 12, 2000.

[8] Mandelbrot, B., *The Fractral Geometry of Nature,* New York: W. H. Freeman and Company, 1983.

[9] Hall, D., et al., "Turn-On Circuits Based on Standard CMOS Technology for Active RFID Labels," Adelaide, Australia: School of Electrical and Electronic Engineering, University of Adelaide, 2005.

[10] Griffin, J. D., et al., "RF Tag Antenna Performance on Various Materials Using Radio Link Budgets," *Antennas and Wireless Propagation Letters,* Vol. 5, No. 1, December 2006, pp. 247–250.

[11] Cheekiralla, S., and D. W. Engels, "A Functional Taxonomy of Wireless Sensor Network Devices," Cambridge, MA: Auto-ID Laboratory, Massachusetts Institute of Technology, 2005.

[12] Jiang, B., et al., "Energy Scavenging for Inductively Coupled Passive RFID Tags," *Instrumentation and Measurement Technology Conference,* Ottawa, Canada, May 2005.

[13] Dulman, S., "Data-Centric Architecture for Wireless Sensor Networks," Ph.D. Dissertation, University of Twente, Enschede, the Netherlands, 2005.

[14] Yang, G. -Z., *Body Sensor Networks,* New York: Springer-Verlag, 2006.

6

RFID System Design Considerations

6.1 RFID System Key Considerations

6.1.1 Configuration Design

In practice, determining number, type, and placement of readers, and the manner in which they are connected to other sensors (e.g., motion detectors) and actuators (e.g., conveyor belt speed controls) is part of a large design challenge. As an example, suppose we wish to use RFID tags to keep track of rare books in a large bookstore; perhaps the most straightforward design is to assign a reader to each bookshelf in order to determine the books in its vicinity. However, the number of readers required by this design, and the implied size of higher level infrastructure to support the data rate from them may not be economically feasible. An alternate design is to assign readers to the points of entry and exit from aisles between bookshelves; in this case, we can infer the current location of a book based on the location of the reader that read its tag most recently. In the former case, tag readers provide *state information* (book x is at location y), whereas in the latter case, readers provide *change-of-state* (event) information (book x just entered aisle z). This design choice at the lower layers of the architecture would affect the amount and nature of data that must be stored at other layers, as well as the complexity and cost of the system. In the state-based design, if all past sensor readings for book x are somehow lost (perhaps due to a system malfunction) the book can still be very easily located by simply issuing a query for its EPC. In the event-based design, this option may not be available because the current location of book x is out of the range of all sensors [1].

Although hardware configurations (placement of readers, interconnections, and so on) are difficult to change on a frequent basis, software

configurations, which handle how readings are interpreted and routed, can be changed without much labor. This possibility provides the opportunity to rapidly incorporate new business processes into the RFID infrastructure.

During the system design stages, we have to keep in mind some of the basic constraints of RFID systems and incorporate those into our approach (Figure 6.1).

When designing the optimum read range for an inductively coupled RFID system (125 kHz and 13.56 MHz), we should primarily consider the reader's power, the tag's power consumption, the tag's quality factor (*Q*), the tag's tuning, the reader's antenna aperture, and the tag's antenna aperture. Secondary considerations include the tag's modulation depth, the reader's signal-to-noise ratio (SNR), the tag's power-conversion efficiency, the reader's antenna tuning and carrier accuracy, the reader's filter quality, how well the reader's driver matches the antenna, the microcontroller's speed and code efficiency, and the tag's data rate. Sometimes, the modulation type also affects the read range. Phase-shift-keying (PSK) and frequency-shift-keying (FSK) systems are inherently more immune to noise than amplitude-shift-keying (ASK) systems, because PSK and FSK systems use a subcarrier that noise cannot easily duplicate. In ASK systems, any sufficiently wide noise spike can look like data and corrupt a bit, so we must use checksums, parity schemes, or a cyclic-redundancy check (CRC) to counteract the noise. In PSK systems, 0° or 180° phase shift represents a binary bit (1 or 0) during the entire bit time; in FSK systems, two different subcarrier frequencies represent 1 or 0. However, in a passive system, the tag does not transmit anything, so there is no true subcarrier, only variations of AM. The use of checksums or CRCs and the range factors mentioned earlier affect the read range so dramatically that any benefits of using FSK or PSK usually become insignificant.

The application environment can also affect the read/write range. Key factors include the proximity of metal to the tag or reader antennas, the presence of in-band noise sources, whether the tag and reader are stationary or moving, and

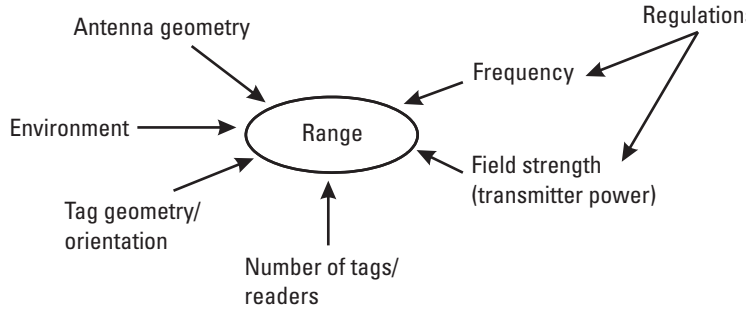


Figure 6.1 Constraints on read/write range.

the angle of the tag with regard to the reader's H-field. Another environmental factor is whether the system is enclosed; a system in a shielded tunnel, for example, can use more power than one in the open air. Some of the additional challenges for RFID systems are large populations of tags, a dynamic tag population, random orientation of tagged objects, and a very high reading speed.

Tag power consumption, turn-on voltage, and modulation depth vary dramatically from model to model and manufacturer to manufacturer. In addition, chips for different bands typically have very different power requirements; for example, a typical 13.56-MHz chip powers up at 4 V_{PP} and typically draws 7 μA, whereas the 125-kHz chip powers up at 9 V_{PP} and draws 10 μA. Power consumption differs widely for systems operating at 13.56 MHz, because CMOS devices consume more current proportionally as their clocking frequency increases. This frequency-dependent consumption is not a problem in synchronous tags operating at 125 kHz; however, a tag that is deriving its clock from a 13.56-MHz carrier has at least one gate that consumes 100 times more current than its counterpart in the 125-kHz tag. The rest of the divider chain draws as much or more than the fastest gate.

As a summary, we can say that the range of passive RFID systems is limited by such factors as tag characteristics, propagation environment, and RFID reader parameters. For example, high-frequency systems have better propagation characteristics, but poorer range in clear air. Pallets, for example, may use UHF tags, but a box of strawberries may need an HF tag. Typically, reader sensitivity is high, and the tag limitation prevails. Tag range can be maximized by designing a high-gain antenna that is well matched to the chip impedance, but this is a task for electronics circuit design engineers and not system designers and/or integrators.

6.1.2 System Design Checklist

Recognizing opportunities for applying any technology, in this case RFID, is largely a matter of being aware of its capabilities and being able to see how those capabilities relate to your own business operations, processes, services, and products. For an RFID project to be successful, it is necessary to approach the business problem and potential RFID solution using a systems approach. During the design process, the designer must look at all of the processes, plan for the future, and think creatively on how you can improve on each operation. RFID systems should be conceived, designed, and implemented using a systematic development process in which end users and specialists work together to design RFID systems based on the analysis of the business requirements of the organization. Implementing an RFID-based system is like implementing any system; the following checklist will help define requirements:

Systems

- Why are you implementing RFID?
- Do you have a mandate to do so or are you looking at improving your internal operation?
- Is there a requirement or preference for standards?
- Is your market domestic, international, or both?

Tags

- Do you require disposable tags or are reusable tags acceptable?
- What type of tag is required (read-only, R/W, WORM)?
- What is the maximum amount of data to be stored in the tag (data capacity)?
- What data format will be used?
- How and where will the tags be applied?
- What do you do when a tag is read?
- What do you do if a tag is not read?

Reader

- What is the required read zone (width, height, and depth)?
- How may tags will the reader read or write to at one time?
- What are the possible location(s) for the tag (redundancy requirements)?
- What is the orientation of the tags and distance between tags?
- At what speed and direction will the tags be traveling?
- What error control and correction will be required?
- Do you require any data security?
- What is the required distance between different reader antennas?
- What is the distance between antenna location and the reader?
- Is portability a requirement?
- Do you need a data interface and protocol for the reader/interrogator (batch, online, wireless, Ethernet, and so on)?

Environment

- What is the proximity of the tags and reader antenna to metals, liquids, and so forth?
- What temperature and humidity will the equipment normally be exposed to? What about exposure to chemicals, UV and X-rays, mechanical stress, splash conditions, dust, and so forth?

Business

- What is the average cost per tag?
- How is the RFID implementation going to affect the bottom line?
- What is the return on investment (ROI)?

Note that the larger the coverage area in the environment, the greater the implementation challenges. Therefore, the longer the reading distance between the tag and the reader, the more noise and interference with which the designer has to contend. Problems with electrical noise are rare, but it is better to perform a site survey before commencing antenna design than to struggle with solving a noise problem later. In general, electrical noise tends to influence the receiver performance and results in reduced reading ranges. Slight changes in antenna orientation to the noise source, additional grounding, or shielding can all help to reduce the effects of noise.

Not many of the off-the-shelf interrogators will survive all of the extreme conditions to which it might be subjected. Making the choice of a proper interrogator for the specific environment is critical in reducing the costs involved with replacing frequently damaged equipment and the downtime associated with hardware failure. A thorough environmental study is always recommended, even if the conditions seem to be readily apparent.

6.1.3 Carrier Frequency and Bandwidth

The carrier frequency and channel bandwidth are key considerations in RFID systems for a number of reasons, practical and legislative. Data is carried on a carrier frequency within a channel defined by government regulations. The channel is characterized by the carrier frequency and the associated bandwidth or range of spectral allocation to accommodate the frequency spread relating to the data-modulated signal. The process of modulation invariably generates symmetrical so-called sidebands, represented in stylized form as the triangles in the diagram on Figure 6.2.

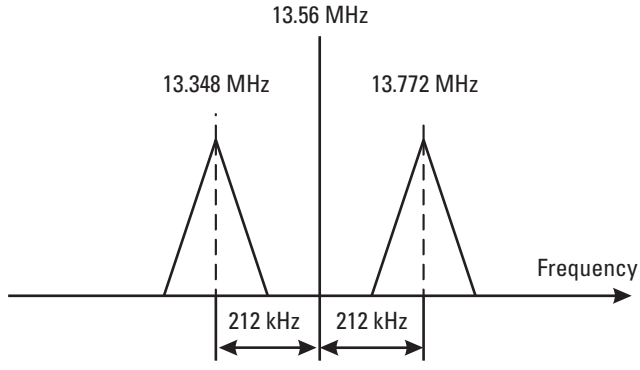


Figure 6.2 13.56-MHz carrier frequency with subcarriers.

Depending on the type of modulation, subcarrier components are used or produced as a result of the modulation process. For example, a technique often used for 13.56-MHz, high-frequency, RFID systems uses a 212-kHz subcarrier to accommodate the baseband coded (source data encoded to accommodate vagaries within the communication channel) data, resulting in two subcarrier modulation products that are close to the reader carrier frequency but sufficiently distant to allow more effective detection and separation from the reader carrier. The bandwidth and the associated sensitivity to frequency components within the band, characterized for both tag and reader, are important for a number of reasons. They largely determine the performance of the transfer system, including susceptibility to interference. These quantities also have to be appropriately controlled to meet regulatory requirements, to ensure they do not interfere with other spectrum users.

A number of channels may be specified for use within a regulatory directive. Where this is so, the channels are sufficiently separated to avoid interference but also require protocols to allow access to these channels without contention. Where the bandwidth and sensitivity are specified, the density of readers for realizing coband operation (i.e., the minimum distance between readers) also becomes a consideration.

From a practical standpoint, the choice of frequency, together with the strength or power of the carrier, has a bearing on the range of communication that can be achieved between tag and reader. To work, the tag has to receive a signal of sufficient magnitude and the reader must be sufficiently sensitive to pick up the tag's response. Any carrier is subject to a reduction in strength the further it is detected from the source. Other factors can also influence the magnitude of the signal over distance, including objects and materials in the region between the tag and reader and mechanisms that add noise or interference to the signal being communicated, making it difficult at the receiver end to distinguish the data-carrying signal from the noise and interference signals.

6.1.4 Frequency Band Selection

In practice, the region between the tag and interrogator may contain obstacles and materials that can influence the performance of the system (Figure 6.3). The carrier frequency is one of significance with respect to the effects that the prevailing conditions and clutter factors (obstacles and physical structures) can have. In LF and HF inductive RFID systems, the magnetic field is effectively used to couple data, and such fields are largely unaffected by dielectric or insulator materials (papers, plastics, masonry, and ceramics, for example); the field simply penetrates the materials.

We mentioned earlier that metals can distort the field dependent on how ferrous they are. This will weaken the field strength in regions of the interrogation zone, in some cases to the extent that system performance is impaired. The range capability may be impaired or the ability to read or write to a tag may be impaired. For uncompensated tag designs operating at resonant frequencies, the presence of metals can often detune the device, in some cases preventing it from operating. At higher frequencies (UHF and above) where, for propagation RFID, the electric component of a field becomes more significant, the higher the frequency the more easily they can penetrate dielectric materials. However, for some materials where energy exchange mechanisms can be identified at or near the carrier frequency, this can result in energy absorption from the propagating wave, thus causing an impairment of range performance. Water molecules for example can have a significant effect on microwave transmissions.

As far as metals are concerned, they reflect or scatter these higher frequency signals depending on the size of the metal object in relation to the

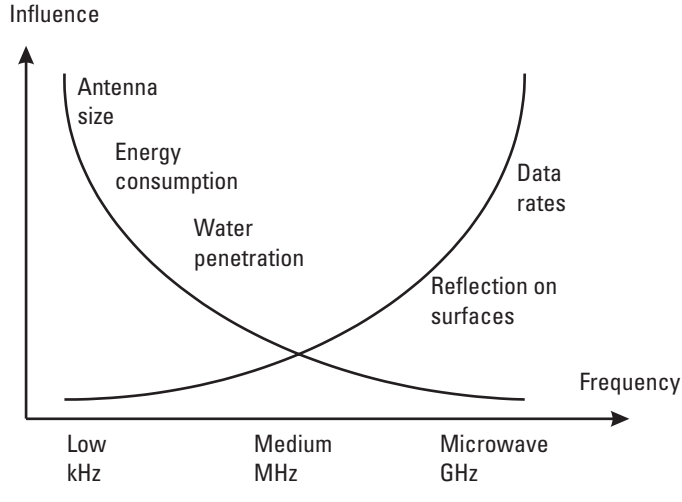


Figure 6.3 Influence of frequency on RFID performance.

wavelength of the incident signal. Such effects can impair the range that can be achieved and in some cases can screen the reader from the tag and prevent it from being read. The closeness of tags may also exhibit a similar effect. Reflection and diffraction effects can often allow pathways around metal objects within an interrogation zone. Because it is difficult to generalize about the effects of clutter within the interrogation zone, it is expedient where possible to avoid clutter and choose a frequency band that is appropriate to the expected conditions.

6.1.5 Power and Range

From what has been said thus far, the simple expedient to extending the range would be to increase the power to interrogate the tag and/or the power available within the tag to affect a response. Indeed this can be done, but only within specified and regulated limits. The extent to which a tag or reader is subject to noise and interference is essentially governed by their respective bandwidth and sensitivity ratings. The greater the sensitivity, the smaller the signals it can respond to, providing they are within the bandwidth of the receiving device. The greater the bandwidth of the receiving device, the greater the susceptibility to noise and interference. However, mitigation techniques may be used to help improve the performance in avoiding or rejecting unwanted signals. The sensitivity, together with channel selection, can also have a bearing on the relative positioning and density of readers. Where readers are operating within the same channel, without any access or anticontention management facilities, the allowable distance between readers is determined by the reader's transmission signal strength and the receiver's sensitivity.

For a given power or operational field strength, the greater the receiver sensitivity, the greater the separation has to be between readers. This, in turn, sets the limit on the density of readers that can be accommodated within a particular application environment. To achieve effective functionality where readers are in range of each other, the readers must operate using appropriate mitigation or communication management techniques. These techniques include channel selection to avoid coband coincidence, operational duty cycles, or access management algorithms.

In the United States, FCC regulations limit the amount of power that can be transmitted between 902 and 928 MHz to 30-dBm maximum transmitter power output and a maximum of 36-dBm effective radiated power. A 6-dBi gain antenna (typical for RFID antennas) is added to 30-dBm transmitter power output to yield a 36-dBm ERP. Sometimes we find that the reader is actually transmitting 32.5 dBm of power, not 30 dBm as required by the FCC. The typical loss in antenna cables is about 2.5 dB. So if we start with 32.5 dBm coming

out of the reader and subtract a 2.5-dB loss in the cables, we realize that 30.0 dBm of power arrives at the antenna. Generally speaking, we can say:

$$\begin{aligned}\text{EIRP}[\text{dBm}] = {} & \text{Tx}[\text{dBm}] - \text{loss in transmission line}[\text{dB}] \\ & + \text{antenna gain}\ [\text{dBi}]\end{aligned} \tag{6.1}$$

It is possible in the United States to use a higher-gain antenna, such as an 8-dBi antenna, as long as the transmitter power output is reduced by 2 dBm, so that the effective radiated power stays under the 36-dBm ERP limit. Generally, an 8-dBi gain antenna has a narrower beamwidth than a 6-dBi antenna, so doing this may be useful in specific situations where a designer wants a longer but narrower read field. In general, users tend to want the largest possible read field; given FCC constraints, that is accomplished with a 6-dBi antenna. It is not a good idea to change the power settings, cabling, or antenna that come with the reader, because this could violate FCC or other local regulations. Check your reader's documentation or ask the manufacturer about changing power settings, cabling, and antennas that comply with relevant regulations.

6.1.6 Link Budget

Due to the indirect power supply in RFID equipment, it is essential to carefully calculate the power budget. Here, we are going to use a simplified calculation using decibels (more detailed formulas and their derivations are presented in Chapter 5). Contributions from the transmitter and the reader antenna are relatively easy to evaluate; however, in the operational environment, the transmission setup and antenna are subject to strong variations due to the strongly variable environment. Modern systems are targeted for communication with up to a few hundred tags, thus requiring good reliability (i.e., safety margins have to be included). In modeling a real situation with a reader, many tags and other objects in between are important parts, of the RFID system design, and the use of RF simulation tools in all parts of the system helps to predict the range of reliable operation. Typical UHF operating parameters are as follows:

- Reader transmitting power $P_R = 33$ dBm (2W);
- System operating wavelength $\lambda = 0.33$m (915 MHz);
- Reader receiver sensitivity $S_R = -80$ dBm (10^{-11}W);
- Reader antenna gain $G_R = 4$ (6 dBi);
- Tag power (sensitivity) requirement $P_T = -14$ dBm (40 μW);
- Tag antenna gain $G_T = 1.26$ (1 dBi);
- Tag backscatter efficiency $E_T = 0.01$ or 1% (−20 dB) calculated as $(\Delta\rho)^2$.

The differential reflection coefficient, $\Delta\rho$, is described in Chapter 5 in more detail.

6.1.6.1 Forward Link Budget

The signal received at the tag (Figure 6.4) has to be bigger or, at the worst case, equal to the tag sensitivity threshold. In the case where the tag is at the far edge of the interrogation zone, we can say that:

$$\begin{aligned} P_{Tag} &= P_R + G_R + G_T - \sum \text{Losses} - FSL \\ FSL &= P_R + G_R + G_T - \sum \text{Losses} - P_{Tag} \\ FSL &= (33+6+1) - (3+3+1) - (-14) = 47 \text{ dB} \end{aligned} \tag{6.2}$$

$$r = 10^{\frac{FSL - 31.75}{20}} \tag{6.3}$$

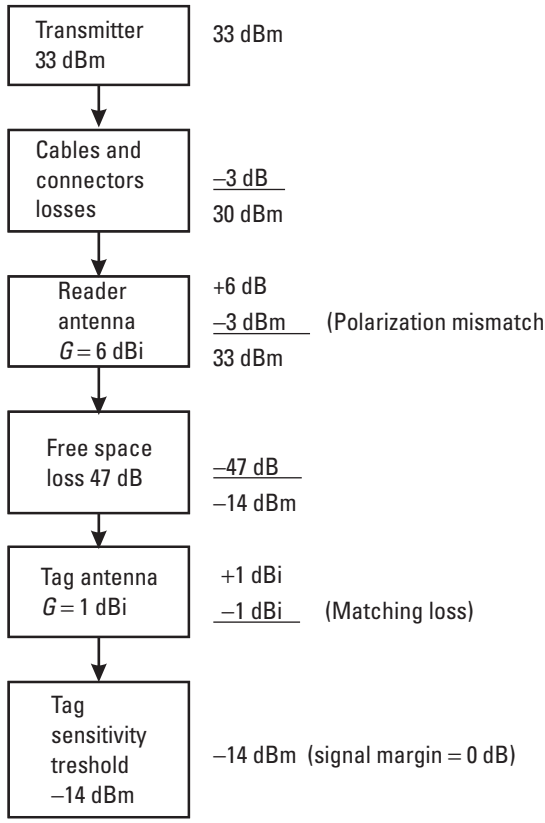


Figure 6.4 UHF RFID forward link budget (reader to tag).

The maximum free-space loss (FSL) allowed for this case is 47 dB. From the Friis formula for FSL, we can calculate the maximum distance at 915 MHz to be 5.8m (or approximately 19 feet).

6.1.6.2 Reverse Link Budget

The backscatter communication radio link budget describes the amount of modulated power that is scattered from the RF tag to the reader, as shown in (6.4). The antenna gains include losses due to mismatch (E_T):

$$P_{REC} = \frac{P_R G_R^2 G_T^2 \lambda^4 E_T}{(4\pi)^4 r^4}$$
$$P_{REC} = \frac{10^3 (2)^2 (1)^2 (0.33)^4 (0.01)}{(4\pi)^4 (5.8)^4} = 0.0000168\ \mu\text{W} \tag{6.4}$$

Using decibels (6.5), we can write:

$$P_{REC} = P_{Tag} + G_R + G_T - \sum \text{Losses} - FSL - E_T$$
$$P_{REC} = -14 + 6 + 1 - (3 + 1) - 47 - 20 \tag{6.5}$$
$$P_{REC} = -78\ \text{dBm (the same result as before)}$$

So, because $S_R = -80$ dBm and $P_{REC} > S_R$, we still have about a 2-dB signal margin at the reader's receiver. So, the question is this: In the UHF read range is the tag sensitivity limited or is the reader sensitivity limited? Well-designed passive systems are always limited by the tag's sensitivity (Figure 6.5).

In practice, the maximum theoretical activation range is decreased by four types of additional losses:

- *Absorption:* Because most RFID systems are deployed indoors and there is not always a line-of-sight path between the tag and the reader, the free-space assumption is usually not valid. The electromagnetic wave supplying tags with power is, for example, completely reflected by perfect conductors and partially reflected by perfect dielectrics. Real-world lossy dielectrics between the reader antenna and the tag antenna will also absorb some of the incident radiation. The result is that, in practice, depending on the material between the tag and the reader, the read range can be significantly less than predicted.
- *Multipath fading:* Even if there is a line-of-sight path between the reader antenna and the tag, small-scale fading effects can increase and decrease the read range. Multipath fading is caused by interference between two

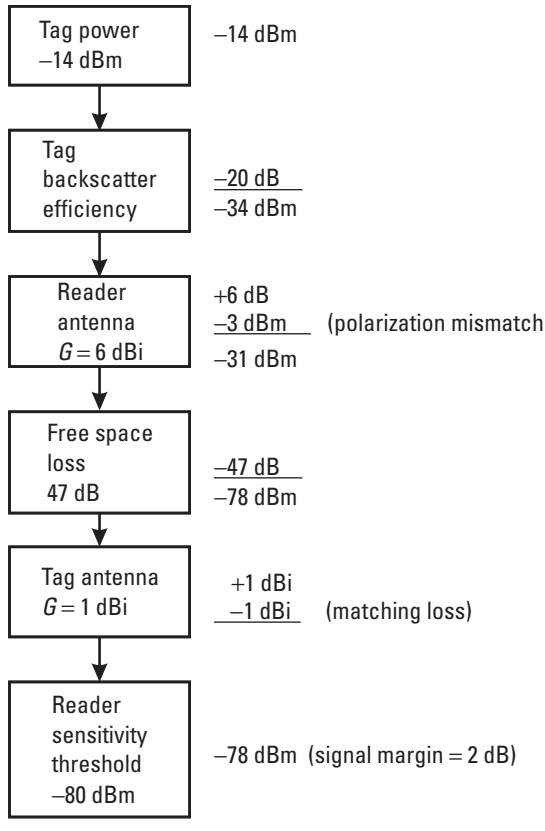


Figure 6.5 UHF RFID reverse link budget (tag to reader).

or more versions of the transmitted reader signal, which arrive at the receiver at slightly different times. These multipath waves combine at the receiver to result in a signal that can vary widely in amplitude and phase. Due to the constructive and destructive effects of multipath waves, a tag moving past a reader antenna can pass through several fades in a small period of time. If the tag passes through such a field null, it will lose power and possibly also its state.

- *Polarization losses:* The activation range is further significantly reduced by polarization losses, because the precise orientation of tags relative to the reader antenna is usually not known. Even when the reader is transmitting with a circularly polarized antenna, the transponder fails to be adequately powered when the axis of the tag dipole antenna is aligned with the propagation direction of the emitted electromagnetic wave. Circularly polarized antennas also introduce an additional loss of 3 dB. A promising approach to alleviate this orientation dependence is the use

of two tag antennas that are orthogonally polarized and attached to the same microchip.

- *Impedance mismatch:* The activation range predicted by the Friis transmission equation could, in practice, be further reduced by impedance mismatch between the tag antenna and microchip. In most calculations, this fact is neglected and a perfect match is assumed.

The *EIRP* determines the power of the signal transmitted by the reader in the direction of the tag. The maximum allowed EIRP is limited by national regulations.

Chip sensitivity threshold is the most important tag limitation. It is the minimum received RF power necessary to turn on an RFID chip. The lower it is, the longer the distance at which the tag can be detected. Chip sensitivity is primarily determined by the RF front-end architecture and fabrication process; RFID chips may also have several RF inputs connected to different antenna ports. Antenna gain is another important limitation; tag range is highest in the direction of maximum gain, which is fundamentally limited by the frequency of operation and the tag size.

Tag detuning is due to the fact that antenna characteristics change when the tag is placed on different objects or when other objects are present in the vicinity of the tag. Tag detuning degrades antenna gain and impedance match and thus affects the tag range.

Reader sensitivity is another important parameter that defines the minimum level of the tag signal that the reader can detect and resolve. The sensitivity is usually defined with respect to a certain SNR or error probability at the receiver. Factors that can affect reader sensitivity include receiver implementation details, communication protocol specifics, and interference, including signals from other readers and tags.

Figure 6.6 shows that, although calculations and even RF measurements indicate that correct tag readings should be achieved without problem, at increased distances they will become increasingly unreliable. After a certain distance, the number of correct readings of the 60-tag pallet of Figure 6.6 will decrease sharply with increasing distance.

6.1.7 Collision Avoidance

A major problem with RFID systems is that a tag might not be read, in spite of being in the reader's range, due to collisions. A collision is said to have occurred when various devices interfere with each others' operations, or their simultaneous operations lead to loss of data. The reading process is not efficient due to various types of collisions, which are classified as follows:

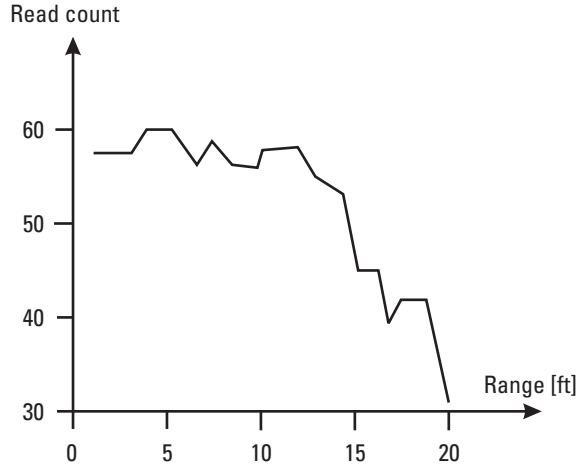


Figure 6.6 The tag sensitivity limitation in practice.

- *Single reader–multiple tags collision:* Multiple tags are present to communicate with the reader. They respond simultaneously and reader is not able to interpret the signal.
- *Single tag–multiple readers collision:* A single tag is in the range of two or more readers. Tags are mainly passive entities, so they do not have enough power to differentiate between the frequency ranges of the readers. Tag interference is more common among active tags, in that they have a greater range and are more likely to interact with multiple readers at a given instant in time. When this problem exists in isolation, it is said to be a resource-constrained scheduling problem and is solved using optimization methods.
- *Reader–reader collision:* Two or more readers within the same frequency range interfere with each others' operations.

These problems need to be resolved to provide efficient solutions for tag identification and these are the major research areas where work needs to be done to practically implement RFID systems. Several metrics can be used to judge the quality of anticollision algorithms: performance, range, bandwidth requirements, implementation costs, noise and error tolerance, and security.

6.1.7.1 Tag–Tag Collision

In many existing applications, a single-read RFID tag is sufficient; animal tagging and access control are examples. However, in a growing number of new applications, the simultaneous reading of several tags in the same RF field is absolutely critical; library books, airline baggage, garment, and retail

applications are a few. To read multiple tags simultaneously, the tag and reader must be designed to detect the condition that more than one tag is active. Otherwise, the tags will all backscatter from the carrier at the same time, and the AM waveforms would be garbled. This is referred to as a collision and no data would transfer to the reader.

With several entities communicating on a same channel, it is necessary to define some rules to avoid collisions and therefore to avoid information loss. The required rules are known as the *collision avoidance protocol*. The tags' computational power is very limited and they are unable to communicate with each other. Therefore, the readers must deal with the collision avoidance themselves, without help from the tags. Usually, the readers' solution consists of querying the tags until all identifiers are obtained. The process of addressing and isolating a single tag is referred to as *singulation*. We say that the reader performs the *singulation* of the tags because it can then request them selectively, without collision, by indicating the identifier of the queried tag in its request [2].

The collision avoidance protocols that are used in current RFID systems are often proprietary (i.e., not open-source) algorithms. Therefore, obtaining information on them is difficult. Currently, several open-source standards are available and they are being used more often instead of proprietary solutions.

We distinguish the EPC family from the ISO family, but regardless of whether they are EPC or ISO, there are several collision avoidance protocols. Choosing one of them depends, in part, on the frequency used. EPC proposes standards for the most used frequency, that is, 13.56 MHz and 860 to 930 MHz. ISO proposes standards from 18000-1 to 18000-6, where 18000-3 corresponds to the 13.56-MHz frequency and 18000-6 corresponds to the 860- to 960-MHz frequency. There are two main classes of collision avoidance protocols: the deterministic protocols and the probabilistic protocols. Usually, we use the probabilistic protocols for systems that use the 13.56-MHz frequency (the United States' regulations in this band offer significantly less bandwidth), and the deterministic protocols for systems using the 860- to 960-MHz frequency because they are more efficient in this case.

Deterministic protocols rely on the fact that each tag has a unique identifier. If we want the singulation process to succeed, the identifiers must stay unchanged until the end of the process. In the current tags, the identifiers are set by the manufacturer of the tag and written in the tag's ROM. In normal RFID systems, there is no exchange after the singulation because the reader has obtained the expected information, that is, the identifiers of the tags which are in its field.

The *probabilistic protocols* are usually based on a time-division multiple access (TDMA) protocol, called ALOHA. The ALOHA protocol is a simple protocol originally developed for use in radio communication systems, but it

can be applied in every system where uncoordinated information is send over the same channel. The original protocol has two rules:

- Whenever you have something to send, send it.
- If there is a collision when transmitting (i.e., another entity is trying to send at the same time), try to resend later. This also applies in the case of transmission failure.

In the tag–reader context, tags avoid collisions with other tags by randomly delaying their responses. If a collision does occur, the reader will inform all nearby tags and the culprits will wait another, usually longer, random interval before continuing. Higher densities of tags will result in a higher collision rate and degraded performance. The ISO 15693 standard for RFID supports a slotted ALOHA mode of anticollision.

Slotted ALOHA is a more advanced, but still simple, protocol, where the receiving entity sends out a signal (called a beacon) at equally spaced intervals, thus dividing time into slots. The beacon announces the start of a new slot and thereby the time to start sending the next packet for any entity having one ready. The version of slotted ALOHA applied in RFID collects a number of consecutive slots into groups. At the beginning of each group the reader announces that only transponders with IDs starting with a specified substring are to answer now. Each tag thus activated picks a random number and waits for that many slots before transmitting.

In general, the number of slots is chosen randomly by the reader, which informs the tags that they will have n slots to answer to its singulation request. Each tag randomly chooses one slot among the n and responds to the reader when its slot arrives. If n is not sufficiently large with regard to the number of tags present, then some collisions occur. To recover the missing information, the reader interrogates the tags one more time. It can mute the tags that have not caused collisions (switched-off technique) by indicating their identifiers or the time slots during which they transmitted. Also, according to the number of collisions, it can choose a more appropriate n.

A simple deterministic algorithm used to solve the tag–tag collision problem is the *tree-walking algorithm* (TWA), which is generally used in UHF readers (Figure 6.7). In this protocol, the reader splits the entire ID space into two subsets and tries to identify the tags belonging to one of the subsets, recursing along the way until a subset has exactly one tag or no tags at all. To describe how a tree walk is performed, a simple example is given, in which the transponders IDs only consist of 3 bits. Three transponders with the IDs 001, 011, and 110 are introduced into the reader's scanning area.

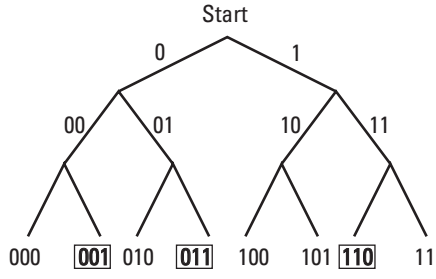


Figure 6.7 The tree-walking algorithm.

The reader first asks if any transponders have a 0 as the first bit. The 110 transponder does not and goes into a sleep state, while the other two transponders answer. The reader then asks if any transponders have a 0 as the second bit. Again this is confirmed by the 001 transponder, but the 011 transponder goes into a sleep state. Then the reader asks for transponders with a 0 as the third bit. Nobody answers, and 001 goes into the sleep state. Because nobody answers, the reader backs up one step and asks all transponders that confirmed their presence at the second bit to wake up. This reactivates 001. The reader now asks for transponders with a 1 as the third bit; 001 answers and is now fully identified. By continuing this back-up-one-step and forward-one-step sequence a number of times, all three transponders are identified.

Due to larger turnaround times at lower frequencies, TWA is not deemed suitable for HF readers. Instead, the HF readers use a slotted termination adaptive collection (STAC) protocol somewhat similar to the framed ALOHA protocol.

The binary tree walking anticollision algorithm discussed here has an inherent security problem due to the asymmetry between forward and backward channel strengths. Every bit of every singulated tag is broadcast by the reader on the forward channel. At certain operating frequencies, a long-range eavesdropper could monitor these transmissions from a range of up to 300 feet (100m) and recover the contents of every tag. A variant of binary tree walking, which does not broadcast insecure tag IDs on the forward channel and does not adversely affect performance, originally appeared under the name *silent tree walking*. Assume that a population of tags shares some common ID prefix, such as a product code or manufacturer ID. To singulate tags, the reader requests all tags to broadcast their next bit. If there is no collision, then all tags share the same value in that bit. A long-range eavesdropper can only monitor the forward channel and will not hear the tag response. Thus, the reader and the tags effectively share a secret bit value. When a collision does occur, the reader needs to specify which portion of the tag population should proceed. If no collisions

occur, the reader may simply ask for the next bit, since all tags share the same value for the previous bit.

Because we assumed the tags shared some common prefix, the reader may obtain it as a shared secret on the uplink channel. The shared secret prefix may be used to conceal the value of the unique portion of the IDs. Suppose we have two tags with ID values $b_1 b_2$ and $b_1 \overline{b_2}$. The reader will receive b_1 from both tags without a collision; then it will detect a collision on the next bit. Because b_1 is hidden from long-range eavesdroppers, the reader may send either $b_1 \oplus b_2$ or $b_1 \oplus \overline{b_2}$ to singulate the desired tag without revealing either bit. Eavesdroppers within the range of the backward channel will obviously obtain the entire ID. However, this blind tree walking scheme does effectively protect against long-range eavesdropping of the forward channel with little added complexity. Performance is identical to regular tree walking, since a tag will be singulated when it has broadcast its entire ID on the backward channel.

A number of other variants of the same idea have been described in the literature.

6.1.7.2 Reader–Tag Collision

A reader–tag collision occurs when the signal from a neighboring reader interferes with tag responses being received at another reader. This problem has been studied in the EPCglobal Class1 Gen 1 and Gen 2 standards for UHF readers. In the Gen 1 standard, the reader–tag collision problem is mitigated by allowing frequency hopping in the UHF band or by TDMA. In Gen 2 the readers and tags operate on different frequencies so that the tag response does not interfere or collide with reader signals. Either solution requires fairly sophisticated technology [3].

6.1.7.3 Reader–Reader Collision

In the future large-scale RFID deployments will most likely involve multiple readers due to the fact that each RFID reader has a limited interrogation range. The reader can communicate with any tag within its interrogation range. The size and shape of the interrogation range of a particular reader is determined by many factors including antenna characteristics, radio transmitting power, radio obstructions (due to, say, packaging) and wireless interferences. It also depends on the characteristics of the tag. It is not uncommon for a single reader to be unable to provide appropriate coverage for the entire region of interest. Also, in several applications such as warehousing or manufacturing, a large area must be perfectly covered. This motivates the use of multiple RFID readers geographically dispersed and networked in some fashion (in an ad hoc network, for example) and performing tag reading concurrently. Use of multiple readers not only improves coverage, but also improves read throughput by virtue of concurrent

operation. However, several collision problems might occur when multiple readers are used in proximity to each other.

Colorwave is a distributed algorithm based on TDMA and one of the first works to address reader–reader collisions [4]. In particular, it considers an interference graph over the readers, wherein there is an edge between two readers if they could lead to a reader–reader collision when transmitting simultaneously. It then tries to randomly color the readers such that each pair of interfering readers has different colors. If each color represents a time slot, then the above coloring should eliminate reader–reader collisions. If conflicts arise (i.e., two interfering readers pick the same color or time slot), only one of them wins (i.e., sticks to the chosen color); the others pick another color again randomly. Some authors suggest coloring of the interference graph using k colors, where k is the number of available channels. If the graph is not k-colorable using their suggested heuristic, then the authors suggest removal of certain edges and nodes from the interference graph, using other methods that consider the size of the common interference regions between neighboring readers.

A query is said to be successfully sent if it is sent by a reader and is successfully received by all the tags in the read range; that is, if it does not collide with any other query in the network. Hence, if the reader does not receive any offline messages for a query, the query is considered to have been sent successfully. We define the system throughput as follows:

$$\text{System throughput} = \frac{\text{total successful query (ALL readers)}}{\text{total time}} \tag{6.5}$$

In general, the tag identification is through a query–response protocol in which the reader sends a query and the tag responds with its unique identification number. The higher the number of queries sent successfully, the higher the throughput and, hence, the higher the number of tags identified by the readers. Thus, throughput and efficiency together define the effectiveness of the anticollision protocol.

Engels et al. presented two algorithms called distributed color selection (DCS) and variable-maximum distributed color (VDCS) [5].

6.1.8 Tag Reading Reliability

Ghost reads occur when an RFID reader gathers information from a noisy environment and reports on a tag that does not exist. Reporting of phantom tags consumes processing and network time that may impact system reliability and performance. Statistically, there is always a chance of a ghost read; however, Gen 2 was specifically designed to address and minimize the occurrence of ghost reads. Statistics reveal that roughly one ghost read occurs per thousand tags. Gen

2 virtually eliminates the potential for these phantom reads, even in a noisy environment, by providing five sequential mechanisms that serve as checks for a tag's validity. Only tags that pass all of these five tests are designated as valid tag reads and entered into the system:

1. *Tag response time:* Tags must respond to a reader within a very short, defined time frame. If the tag response is not timed exactly from the beginning of a response to the end of the response, the probability that the tag is a phantom is high, so the reader will ignore the tag. The reader may try to reread the tag at a later time, possibly under different conditions.
2. *Preamble:* For each and every response, tags first send a signal called a preamble. When a reader receives a valid preamble, it knows the signal is valid (from a real tag), and not simply noise. If the preamble is not valid, the reader discontinues communication and moves on to the next tag signal.
3. *EPC format check:* If the preamble is validated, the reader then examines the bitstream transmission to ensure that it is in a valid EPC format. If the EPC format is validated, communication between tag and reader continues. If the EPC format is not validated, the reader begins communication with another tag.
4. *Bit match:* The reader compares the number of bits the Gen 2 tag reported it would be sending to the number of bits received; if it is not a match, the information is discarded.
5. *Cyclic redundancy check:* The CRC checks for bit errors in transmission by comparing the number of bits the tag stated it would send with the number of bits actually sent. If the correct number of bits was received, the transmission is verified as accurate. If the correct number of bits was not received, the data is rejected and the reader moves on to begin communication with the next tag.

6.2 RFID Reader–Tag Communication Channel

Irrespective of the mode of coupling (inductive or propagation), the means of effectively transferring data between tag and reader relies on a dialogue between the two. Based on command data within the reader and the tag response signals, the dialogue is generally initiated by the reader (readers talk first). However, in some RFID systems, the tag may operate in a beacon-type mode and effectively talk first. Suffice it to say at this stage that the object in either case is to transfer data; so the dialogue essentially requires the tag to be identified and data

requested, acknowledged, and sent. To achieve this, it is necessary for data to be written or encoded into a tag in a particular way:

1. With other data elements added to facilitate identification, the source data (the data required to be carried and used at the receiver end of communications) components (so-called *metadata*) and, as appropriate, elements used for error detection and correction, contention management, and communications dialogue are included. This is generally known as *source encoding*.
2. The source encoded message is structured by certain baseband techniques to better match the signal form of the message to be sent to the characteristics of the transmission channel or medium through which it is to be transmitted. This is process is often referred to as *channel encoding*.

The data, communicated between tags and readers, must be sent in a reliable manner. With data encoded in this way, the final conditioning that is used to facilitate transmission is the modulation of the encoded data using a suitable frequency-defined carrier signal. At the receiver end the reverse of these processing elements is performed (demodulation, together with channel and source decoding) to recover the source data. A variety of techniques are used for channel encoding and modulation distinguished by particular performance and cost characteristics.

The combination of coding and modulation schemes determines the bandwidth, integrity, and tag power consumption. The coding and modulation used in RFID communications is limited by the power and modulation/demodulation capabilities of the tags. Another limiting factor is the bandwidth occupied by the signal; RFID tags that are passive do not transmit signals actively and therefore can use more bandwidth than a reader. A reader has its own power source and therefore is required by regulations to use less bandwidth.

6.2.1 Data Content and Encoding

Line coding involves converting a sequence of 1s and 0s to a time-domain signal (a sequence of pulses) suitable for transmission over a channel. The following primary factors should be considered when choosing or designing a line code:

1. *Self-synchronization:* Timing information should be built into the time-domain signal so that the timing information can be extracted for clock synchronization. A long string of consecutive 1s and 0s should not cause a problem in clock recovery.

2. *Transmission power and bandwidth efficiency:* The transmitted power should be as small as possible, and the transmission bandwidth needs to be sufficiently small compared to the channel bandwidth so that intersymbol interference will not be a problem.
3. *Favorable power spectral density:* The spectrum of the time-domain signal should be suitable for the transmission channel. For example, if a channel is ac coupled, it is desirable to have zero power spectral density near dc to avoid dc wandering in the pulse stream.
4. *Low probability of error:* When the received signal is corrupted by noise, the receiver can easily recover the uncoded signal with low error probability.
5. *Error detection and correction capability:* The line code should have error detection capability, and preferably have error correction capability.
6. *Transparency:* It should be possible to transmit every signal sequence correctly regardless of the patterns of 1s and 0s. If the data is coded so that the coded signal is received correctly, the code is transparent.

Two broad categories of codes are used in RFID: level codes and transition codes. *Level codes* represent the bit with their voltage level. *Transition codes* capture the bit as a change in level. Level codes, such as nonreturn-to-zero (NRZ) and return-to-zero (RZ), tend to be history independent; however, they are not very robust. Transition codes can be history dependent, and they can be robust. Transmitter (tag) is responsible for encoding, that is, inserting clocks into the data stream according to a select coding scheme while the receiver (reader) is responsible for decoding, that is, separating clocks and data from the incoming embedded data stream. Readers are capable of transmitting at high power but are limited to narrow communication bands by communications regulations; therefore, the encoding used from reader to tag usually needs to occupy a low bandwidth. Passive tags, however, do not actively transmit a signal; therefore, the encoding used for tag-to-reader communication can occupy a high bandwidth. There are many types of line codes but we will only discuss a few of them (Figure 6.8) that are important for our RFID discussion. More general information about data encoding can be found in [6].

6.2.1.1 Nonreturn-to-Zero Coding

The simplest channel code is the one known as NRZ, or nonreturn-to-zero. Simple, combinational logic signals are a good example of NRZ, where a logic 1 is coded as one dc level and logic 0 as another. NRZ requires time coordination but longs strings of 0s and 1s do not produce any transitions that could create problems in error detection and recovery. NRZ produces a high dc level (average of

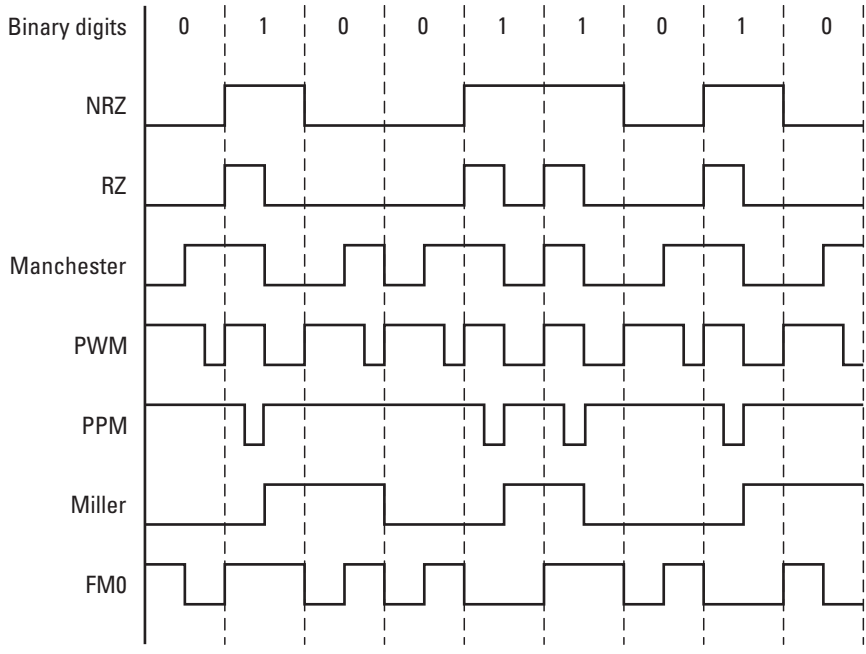


Figure 6.8 Examples of several coding schemes.

0.5V). NRZ is rarely used for serial transmission, except for relatively low-speed operations such as those associated with modem transfers.

Serial transmission in wired systems generally consists of at least two transmission lines: one carries the data, and the other the clock to which the data is synchronized. Wireless transmission, however, is an entirely different situation; wireless data has only one medium to travel through, the air, and as such cannot support separate transmission of data and clock. Therefore, traditional NRZ data, in which a logic 1 is a high signal for one clock period, and a logic 0 is a low signal for one clock period, cannot be used. A data stream, in general, can contain long strings of 1s or 0s; these would be represented in an NRZ stream by very long dc values, during which the receiving system may lose synchronization with the transmitter's clock. To combat this, the Manchester code, for example, can be used.

6.2.1.2 Return-to-Zero Coding

Return-to-zero (RZ) coding describes a line code used in telecommunications signals in which the signal drops (returns) to 0 between each pulse. This takes place even if a number of consecutive 0s or 1s occur in the signal. This means that a separate clock does not need to be sent alongside the signal, but suffers

from using twice the bandwidth (data rate) compared to the NRZ format. The signal is self-clocking.

Although RZ coding contains a provision for synchronization, it still has a dc component resulting in baseline wander during long strings of 0 or 1 bits, just like the NRZ line code. A variant, RZ inverted, swaps the signal values for 1 and 0.

6.2.1.3 Biphase Mark (Manchester) Coding

Manchester coding incorporates a transition in the middle of every transmitted bit. A logic 1 is represented by a transition from low to high, and a logic 0 is represented by a transition from high to low. Because the transmitted signal must change at least once for every bit transmitted, the problem of transmitting long dc values is eliminated. The Manchester code provides for efficient communication because the bit rate is equal to the bandwidth of the communication. Biphased data streams generally have a signal change in the middle of each bit, independent of the value. Therefore, the signal does not necessarily return to zero. The biphase method has these characteristics:

- *Synchronization:* Because the transition for each bit is predictable, the receiver can synchronize on this edge. These codes are also known as self-clocking.
- *Error immunity:* To cause an error, the noise must invert the signal both before and after the transition.
- A 1-to-0 transition represents a 0 bit.
- A 0-to-1 transition represents a 1 bit.
- The mid-bit transition is used as a clock as well as data.
- The residual dc value is eliminated by having both polarities for every bit.
- The bandwidth required could be twice the bit rate (the efficiency of this code can be as low as 50%).

Unlike NRZ data with a separate clock signal, the clock is not provided explicitly to the device receiving Manchester-encoded data; instead, it is given, encoded, in the transmitted data so the clock must be recovered from the data. Traditionally, this is accomplished via a phase-locked loop (PLL) that takes in the received data stream and outputs the transmitter's clock.

6.2.1.4 FM0 Coding

EPCglobal Class 1 Gen 2 provides multiple options for tag coding and the simplest approach is FM0 coding. In FM0 coding (biphase space), a transition has

to occur at the end of each bit period, but for a 0 bit, an addition transition in the middle is required. The duty cycle of a 00 or 11 sequence, measured at the modulator output, is a minimum of 45% and a maximum of 55%, with a nominal value of 50%. FM0 encoding has memory; consequently, the choice of FM0 sequences depends on prior transmissions. FM0 signaling always ends with a dummy data 1 bit at the end of a transmission.

6.2.1.5 Miller Coding

Miller coding is also called *delay modulation* and provides a transition for every bit. In this code a 1 is encoded as a transition occurring at the center of the bit cell, while consecutive 0s have a transition at the cell boundary between them. This means that a pattern such as 10101 has no transitions at the cell boundaries.

- There is a transition in the middle of a bit period if it is a 1 bit.
- There is a transition at the start of the bit period if the 0 bit is followed by a 0 bit.
- For a 0 followed by a 1 or a 1 followed by a 0, no transition occurs at the symbol interval.

This code is very efficient in terms of the desired bandwidth (half of the desired bandwidth of Manchester coding). Miller coding is self-clocking and has a relatively low LF content but is not dc-free, however.

Miller squared coding (so called because it was the result of a modification of Miller coding by a second, quite separate Miller!) has one additional rule. This states that the final transition of an even number of 1s occurring between two 0s is omitted (i.e., 01110 occupies five cells and has three transitions, whereas 011110 occupies six cells but also has three transitions), thus making the code dc-free.

6.2.2 Modulation

The data coding scheme determines how the data is represented in a continuous stream of bits. How that stream of bits is communicated between the tag and the reader is determined by the modulation scheme. Each RFID standard employs one modulation scheme for the forward link (reader-to-tag) and another for the reverse link (tag-to-reader). The modulation schemes reflect the different roles of reader and tag. The reader must send enough RF power to keep the tag powered. A passive tag does not transmit its own signals but modulates by changing the phase or amplitude of the reader's transmitted signal that is being backscattered (in UHF system) from its antenna.

For convenience, RF communications typically modulate a high-frequency carrier signal to transmit the baseband code. RFID systems usually employ modulation techniques and coding schemes that are simple to produce. The three classes of digital modulation are ASK, FSK, and PSK. The choice of modulation is based on power consumption, reliability requirements, and bandwidth requirements. All three forms of modulation may be used in the return signal, although ASK is most common in load modulation at 13.56 MHz, and PSK is most common in backscatter modulation. The choice is essentially determined by performance requirements and cost. For example, ISO 18000 Type C (also known as EPC Class 1 Gen 2) calls for double sideband-ASK (DSB-ASK), single sideband-ASK (SSB-ASK), and phase reversal-ASK (PR-ASK). Amplitude-shift-keyed digital modulations are spectrally inefficient, requiring substantial RF bandwidth for a given data rate. Bandwidth efficiencies of 0.20 bit/Hz of RF bandwidth are not uncommon for DSB-ASK.

One approach to improving bandwidth efficiency is to use SSB-ASK. This is particularly important in European countries where bandwidth restrictions may preclude DSB-ASK. The power efficiency of DSB-ASK and SSB-ASK is dependent on the modulation index. With a modulation index of 1, or on-and-off keying (OOK) of the carrier, the lowest carrier-to-noise ratio (C/N) required to achieve a given bit error rate (BER) is obtained for DSB-ASK and SSB-ASK. Unfortunately, this also provides the least amount of RF power transport on the downlink to supply the tag with energy. Ideally, the off time of the carrier should be minimized, so that the tag doesn't run out of power. The C/N requirements should also be minimized to maximize the ID read range. For many modulations, these are conflicting goals.

One such modulation that can minimize the C/N requirement in a narrowband while maximizing the power transport to the tag is PR-ASK. Similar to a PSK signal, PR-ASK changes phase 180° each time a symbol is sent. PR-ASK also creates an AM depth of 100% or a modulation index of 1, as the phase vector of the old symbol and the new symbol cross and briefly sum to a zero magnitude. This provides an easily detected clock signal, as the amplitude briefly goes to zero, but minimizes the time the carrier power is off, so power transport to the passive tag is optimized. PR-ASK has C/N and bandwidth requirements that more closely match PSK than DSB-ASK, making it attractive for narrowband and longer range applications.

DSB-ASK is the least bandwidth efficient modulation, but the easiest to produce by OOK of the carrier signal. ASK modulation specifications often have a modulation depth as well as rise and fall time requirements. The rise and fall time is typically related to the bandwidth filtering, whereas the modulation depth is set by the attenuation difference between the keying states.

Systems incorporating passive RFID tags operate in ways that may seem unusual to anyone who already understands RF or microwave systems. There is

only one transmitter; the passive tag is not a transmitter or transponder in the purest definition of the term, yet bidirectional communication is taking place. The RF field generated by a tag reader (the energy transmitter) has three purposes:

1. Induce enough power into the tag coil to energize the tag. Passive tags have no battery or other power source; they must derive all power for operation from the reader field. The 125-kHz and 13.56-MHz tag designs must operate over a wide dynamic range of carrier input, from the very near field (in the range of 200 V_{PP}) to the maximum read distance (in the range of 5 V_{PP}).
2. Provide a synchronized clock source to the tag. Many RFID tags divide the carrier frequency down to generate an on-board clock for state machines, counters, and so forth, and to derive the data transmission bit rate for data returned to the reader. Some tags, however, employ on-board oscillators for clock generation.
3. Act as a carrier for return data from the tag. Backscatter modulation requires the reader to peak-detect the tag's modulation of the reader's own carrier.

Although all of the data is transferred to the host by amplitude modulating the carrier (backscatter modulation), the actual modulation of 1s and 0s is accomplished with three additional modulation methods:

- *Direct modulation:* The amplitude modulation of the backscatter approach is the only modulation used. A high in the envelope is a 1 and a low is a 0. Direct modulation can provide a high data rate but low noise immunity.
- *FSK:* This form of modulation uses two different frequencies for data transfer; the most common FSK mode is $F_c/8$ to $F_c/10$. A 0 is transmitted as an amplitude-modulated clock cycle with period corresponding to the carrier frequency divided by 8, and a 1 is transmitted as an amplitude-modulated clock cycle period corresponding to the carrier frequency divided by 10. The amplitude modulation of the carrier thus switches from $F_c/8$ to $F_c/10$ corresponding to 0s and 1s in the bitstream, and the reader has only to count cycles between the peak-detected clock edges to decode the data. FSK allows for a simple reader design and provides very strong noise immunity, but suffers from a lower data rate than some other forms of data modulation.
- *PSK:* This method of data modulation is similar to FSK, except only one frequency is used, and the shift between 1s and 0s is accomplished

by shifting the phase of the backscatter clock by 180 degrees. Two common types of PSK are (1) change phase at any 0 or (2) change phase at any data change (0 to 1 or 1 to 0). PSK provides fairly good noise immunity, a moderately simple reader design, and a faster data rate than FSK.

Regardless of what method of carrier modulation is implemented, any voltage modulation sequence controlled through the tag's memory or circuitry will result in the transmission of a bit pattern that mimics the modulation sequence. Therefore, essentially any binary information stored on the tag can be wirelessly transmitted back to the receiver.

As already mentioned, a problem unique to RFID systems is the vast difference in power between the outgoing signal from the reader and that returning to the reader as reflected from the tag. In some situations, this difference may be in the range of 80 to 90 dB, and the return signal may be impossible to detect. To avoid this problem, the return signal is sometimes modulated onto a subcarrier, which is then modulated onto the carrier. For example, in the ISO 15693 standard for RFID, a subcarrier of 13.56/32 = 423.75 kHz is used.

As defined by ISO/IEC 18000-3 (13.56 MHz), the modulation used is typically ASK (either 10% or 100%) for the forward link (reader to tag) and load modulation for the reverse link with a rate defined as a division of the carrier. The load modulation produces subcarriers that utilize binary phase-shift-keying (BPSK) modulation. Load modulation appears in the frequency domain as sidebands offset by the subcarrier frequency from the transmission frequency. Figure 6.9 illustrates this approach.

Future developments in RFID devices will include more complex modulation schemes, for example, technologies such as software defined radio (SDR) are being implemented in dedicated short-range communications (DSRC) and other applications.

6.2.3 Data Encryption

RFID readers in public places can read the RFID data and connect to networks that provide real-time data about the owners and thus infringe on privacy. Encryption of data will help to solve this problem to a certain extent. Encryption of the RFID tag contents will ensure that unauthorized readers are not able to access the data or, if they can get the encrypted content, they will not have access to the encryption key. Many encryption algorithms are available that can be suitably used to encrypt the data on RFID.

EPCglobal has recently ratified its Gen 2 global standard that uses frequency and power in a way that complies with the major regional regulatory

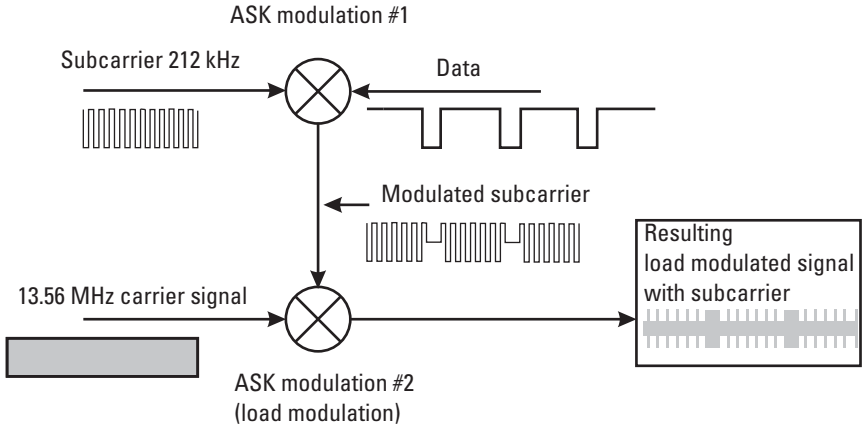


Figure 6.9 Load modulation diagram.

environments. In addition to improvements in security of the data on the tag, the standard includes the ability to lock the identification fields in the tag, so that they cannot be spoofed or changed without a password. It also includes a strong kill mechanism, so retailers and others have the option of automatically erasing all data from the tag as it passes through a reader. However, the standard does not allow for encryption, because one of the user requirements for the standard was that the tags be inexpensive. But security issues will continue to be addressed in the hardware and policy working groups [7] and, in the meantime, implemented as proprietary solutions (AES, for example) by some of the equipment suppliers.

Current implementations of secure RFID rely on digital cryptographic primitives in the form of hashes and block ciphers. The presence of these blocks is motivated by privacy requirements, but they increase the overall processing latency, the power consumption, and the silicon area budget of the RFID tag. In addition, existing passive RFID systems rely on simple coding and modulation schemes using narrowband RFs, which can be easily eavesdropped on or jammed.

6.2.3.1 Data Encryption Standard

The *Data Encryption Standard* (DES*)* algorithm is the most widely used encryption algorithm in the world. For many years, and among many people, secret code–making and DES have been synonymous. DES works on bits, or binary numbers—the 0s and 1s common to digital computers. Each group of 4 bits makes up a hexadecimal, or base 16, number. Binary 0001 is equal to the hexadecimal number 1; binary 1000 is equal to the hexadecimal number 8; 1001 is equal to the hexadecimal number 9; 1010 is equal to the hexadecimal number A; and 1111 is equal to the hexadecimal number F.

DES works by encrypting groups of 64 message bits, which is the same as 16 hexadecimal numbers. To do the encryption, DES uses keys, which are also apparently 16 hexadecimal numbers long or apparently 64 bits long. However, every 8th key bit is ignored in the DES algorithm, so that the effective key size is 56 bits. But, in any case, 64 bits (16 hexadecimal digits) is the round number on which DES is organized.

In cryptography, *Triple-DES* is a block cipher formed from the DES cipher by using it three times. Given a plaintext message, the first key is used to DES-encrypt the message. The second key is used to DES-decrypt the encrypted message, and because the second key is not the right key, this decryption just scrambles the data further. The twice-scrambled message is then encrypted again with the third key to yield the final ciphertext. In general Triple-DES with three different keys has a key length of 168 bits: three 56-bit DES keys (with parity bits Triple-DES has the total storage length of 192 bits), but due to the meet-in-the-middle attack, the effective security it provides is only 112 bits.

Triple-DES is slowly disappearing from use, as it is largely replaced by its natural successor, the *Advanced Encryption Standard* (AES). One large-scale exception is within the electronic payments industry, which still uses these methods extensively and continues to develop and promulgate standards based on it. This guarantees that Triple-DES will remain an active cryptographic standard well into the future. By design, DES and, therefore, Triple-DES suffer from slow performance in software; on modern processors, AES tends to be around 6 times faster. Triple-DES is better suited to hardware implementations (e.g., VPN appliances and the Nextel cellular and data network), but even there AES outperforms it. Finally, AES offers markedly higher security margins; a larger block size, potentially longer keys, and hopefully freedom from cryptanalytic attacks.

6.2.3.2 Advanced Encryption Standard

AES is a symmetric key encryption technique that, as mentioned, will eventually replace the commonly used DES. It was the result of a worldwide call for submissions of encryption algorithms issued by NIST in 1997 and completed in 2000. The winning algorithm, Rijndael, was developed by two Belgian cryptologists, Vincent Rijmen and Joan Daemen. AES provides strong encryption and has been selected by NIST as a Federal Information Processing Standard in 2001, and in 2003 the National Security Agency (NSA) in the United States announced that AES is secure enough to protect classified information up to the top secret level, which is the highest security level, and defined as information that would cause exceptionally grave damage to national security if disclosed to the public.

The AES algorithm uses one of three cipher key strengths: a 128-, 192-, or 256-bit encryption key (password). Each encryption key size causes the

algorithm to behave slightly differently, so the increasing key sizes not only offer a larger number of bits with which you can scramble the data, but also increase the complexity of the cipher algorithm. AES is fast in both software and hardware, is relatively easy to implement, and requires little memory. As a new encryption standard, it is currently being deployed on a large scale.

6.3 Testing and Conformance

6.3.1 Test Equipment

RFID engineers today face a variety of design and test challenges to bring a product to market. First, the product must meet local frequency regulations to emit energy into the spectrum; next, the interrogator and tag interaction must reliably work together. To accomplish this, both the interrogator and tag must comply with the appropriate industry standard. Finally, to be competitive, the RFID system's performance must be optimized to appeal to a particular market segment. This could mean maximizing the number of transactions per second, operating in a dense reader environment, or stretching the reader's ability to communicate over longer distances.

RFID systems, particularly those with backscattering passive tags, present some unique challenges for test and diagnostics. Timing measurements are of particular concern, because system readers can be required to read the ID data from many tags very quickly without error. Most RFID systems use transient time division duplexing (TDD) schemes, in which the interrogator and tags take turns communicating on the same channel. To read many ID tags within a very short period of time with a serial TDD multiplexing scheme, the standards call for very precise timing. Timing measurements on the data interchange thus present a unique RFID challenge. The transient RFID signals often contain spectrally inefficient modulations using special PCM symbol encoding and decoding. Troubleshooting the homodyne interrogators or tags that receive these unusual signals requires special signal analyzer capabilities. Traditionally, swept tuned spectrum analyzers, vector signal analyzers, and oscilloscopes have been used for wireless data link development.

The spectrum analyzer has historically been the tool of choice to characterize the RF spectral output of a transmitter to ensure compliance with regulatory emission restrictions. The traditional swept tuned spectrum analyzer was developed primarily for the analysis of continuous signals, not the intermittent RF transients associated with modern RFID products. This can lead to a variety of measurement issues, particularly the accurate capture and characterization of transient RF signals.

Similarly, the vector signal analyzer possesses little ability to capture transient RF signals, also being initially developed for CW signals. Though most

vector signal analyzers have extensive demodulation ability for popular spectrally efficient modulations, current offerings have virtually nothing to support the spectrally inefficient RFID modulations and their special PCM decoding requirements.

The oscilloscope has long been a valuable tool for analysis of baseband signals. In recent years some oscilloscopes have extended their sampling speed to very high microwave frequencies. They are, however, still suboptimal tools for UHF or higher frequency measurements on RFID systems. Relative to the modern real-time spectrum analyzer, the fast oscilloscope has substantially less measurement dynamic range and lacks modulation and decoding capability.

The real-time spectrum analyzer (RTSA) solves the limitations of the traditional measurement tools to provide a substantially more efficient test and diagnostic experience for the RFID engineer [8]. Pulsed tag reads and writes require an RF analyzer optimized for transient signals. The RTSA has the digital processing speed necessary to transform the input signal from time-domain samples into the frequency domain with a real-time FFT prior to capturing a recording of data. This enables the RTSA to compare spectral amplitudes to a frequency mask set by the user in real time. The RTSA can then trigger a capture on a spectral event of interest for subsequent detailed off-line analysis. Many RFID and near-field communications (NFC) devices use proprietary communications schemes that are optimized for specific market applications, so the test equipment should offer a variety of flexible modulation measurements that enable testing of the proprietary systems with manually configured measurements. The instrument should allow a user to define the modulation type, decoding format, and data rate.

Once the basic specifications are met, it is important to optimize some of the RFID product's features to gain a competitive advantage in a particular market segment. One such example is optimizing the number of tag reads possible in a given amount of time, resulting in the overall system capacity increase, and thus making it more appealing to (lucrative) high-volume applications. An important element in maximizing capacity is minimizing the turnaround time for each tag reply; available RF power, path fading, and altered symbol rates can lengthen the time it takes for the tag to reply to the interrogator's query. The slower the reply, the longer it will take to read a large number of tags.

6.3.2 Frequency- and Bandwidth-Related Measurement

As mentioned earlier, bending tags or placing them in proximity to conductive objects and other tags can detune the tag antenna, preventing them from going into resonance and thus either becoming inoperative or significantly reducing the operating range. Consequently, *frequency deviation measurements* are critical to ensuring compliance with various RFID and transmitter standards. For

example, frequency accuracy for a 13.56-MHz RFID interrogator is typically specified at ± 7 kHz.

Like frequency deviation measurements, *occupied bandwidth measurements* ensure compliance with standards designed to prevent interference with other signals. RFID readers and tags are intentional transmitters that fall under regional regulations, such as FCC 47 Part 15 in the United States, EN 300 330 in Europe, and ARIB STDT60/ T-82 in Japan. As RFID heads toward global acceptance, the most stringent of these regulations will apply. Also under close scrutiny is the effect of human exposure to RFID electromagnetic fields, as spelled out in documents such as IEEE C.95-1-1991 and EN50364:2002 ("Limitation of Human Exposure to Electromagnetic Fields from Devices Operating in the Frequency Range 0 Hz–10 GHz, Used in Electronic Article Surveillance (EAS), Radio Frequency Identification (RFID), and Similar Applications"). Various regulations will define limits using different units. The power flux density S [mW/cm^2], electric field strength E [V/m], and magnetic field strength H [A/m] are interchangeable according to the following equation:

$$S = E^2/3{,}700 = 37.7\,H^2 \tag{6.6}$$

6.3.3 Polling and Timing Measurements

When the RFID reader/interrogator searches for a tag it is referred to as *polling.* Associated with polling is a number of timing measurements called out in various RFID standards. One key timing measurement is *turnaround time,* for both the transmit-to-receive and receive-to-transmit modes. Other timing measurements are *dwell time* or *interrogator transmit power on ramp, decay time* or *interrogator transmit power down ramp,* and *pulse pause* timing.

6.3.4 Collision Management

ISO/IEC 18000-3 Section 6.2.7.9 calls for reading 500 tags within 390 ms. It also calls for reading 50 words of data within 930 ms from static tags, and 944 ms from active tags. If you are using only a PC to time stamp the interactions and test for compliance, there is no way to know why the interaction takes so long, at which point a collision is occurring, or where the particular tag is that is being problematic. However, by using test equipment and monitoring the over-the-air interface during polling, it is possible to troubleshoot when a collision occurs and determine the cause (interference, faulty tag, hopping pattern error, and so on).

6.3.5 Multivendor Interoperability

If RFID is to successfully penetrate into large open systems, RFID interoperability is a necessity. Not only must tags from any vendor be able to communicate with readers from any vendor, but a given tagged object must be able to be identified by readers of any user in a wide variety of application conditions. Today, RFID systems are primarily comprised of systems that may not always interoperate due to the mix of RF propagation technologies and information protocols. The formation of EPCglobal, a joint venture formed by UCC and EAN, is expected to drive global retailer and manufacturing adoption of RFID technologies, especially in the supply chain management applications.

Key to market proliferation of UHF Gen 2–compliant products is the EPCglobal program to certify the hardware that implements the standard. This includes the testing of tag chips, readers, and printer/encoders with embedded reader modules. Products that pass the tests conducted by MET Laboratories (an independent third-party lab contracted by EPCglobal) earn the EPCglobal certification marks, a seal of approval, which indicates the products' adherence to the stringent requirements of the standard. EPCglobal has defined three phases of certification:

- Compliance;
- Interoperability;
- Performance.

Compliance testing verifies that products comply with the UHF Gen 2 standard, and products bearing the certification mark are your assurance that they have been rigorously tested against EPCglobal standards. Some of the relevant electromagnetic compatibility (EMC) and safety standards are:

United States

- U.S. EMC—FCC Rule Part 15 or 90;
- Safety—UL 60950 for tag interrogators and NRTL certification.

Canada

- Canada EMC—RSS-210;
- Safety—CSA 60950 and SCC certification body.

Europe

- Europe EMC testing in accordance with ETSI EN 301 489-1 and ETSI EN 301 489-3;
- Radio testing in accordance with ETSI 300-220;
- Safety testing in accordance with EN 60950;
- Declaration of conformity for CE marking requirements.

The FCC recently classified passive RFID chips as unintentional radiators. This implies that passive RFID chips are exempt from the same clearance tests as other technical devices with electromagnetic fields.

Interoperability testing builds on the compliance certification and verifies the ability of different compliance-certified Gen 2 components to work together. Despite the obvious advantages of ensuring interoperability among products, not all RFID vendors have been able to achieve this level of certification.

Clearly, certified UHF Gen 2 interoperability is a major milestone in the development of RFID systems. As important as that is, though, *performance* is still what matters most to RFID deployments. RFID hardware must have a high degree of receptivity, meaning both tags and readers are not only extremely sensitive to each other's signals, they are also able to reject the interference from other RF sources operating in the area. EPCglobal, recognizing the critical importance of receptivity to system performance, created a working group to address these and other issues. In the process of defining minimum requirements, they will address not only the performance of tags applied to various classes of products, such RF-friendly materials as paper, plastic, wood, and so forth, as well as more problematic materials such as liquids and metals, but also the key aspects of tag performance, such as sensitivity, interference rejection, orientation, and electrostatic discharge (ESD). Once the objectives are defined, the EPCglobal Hardware Action Group will draft the specifications for performance testing and thus complete the UHF Gen 2 standard.

Although all products submitted to interoperability testing must first be certified for compliance to the Gen 2 standard, it is not uncommon for some manufacturers to have misinterpreted certain elements of the specification, preventing their tags, for example, from operating with other Gen 2 devices. More insidiously, certain tags and readers may be interoperable with each other, but not with all other Gen 2 devices. As such, the scope of interoperability tests should be designed to exercise, as much as possible, the full functionality of the Gen 2 spec (including operation at timing limits) with a prime objective of ensuring true multivendor compatibility.

Interoperability problems between the various industries are on the horizon because cross-industry requirements usually play a minor role within a given industry. However, to fully exploit the RFID technology's potential, cross-industry standards have to ensure interoperability. Therefore, cross-industry consultation is necessary to prevent the emergence of different standards from impeding the use of RFID across industries.

6.3.6 Test Labs

We have mentioned a few times already that RFID system calculations are just a first approximation of the real-world environment. RFID systems do not always necessarily work as well as theoretically described, and sometimes as advertised. This can be due to interference in the environment, unforeseen reflections, or simply an installation that does not account for the peculiarities of the technology. Theoretical formulas and analysis, even computer simulations, may not be always sufficient to represent customer environments and predict the RFID system behavior.

Independent test labs simulating the real customer environment could be the answer; services offered usually include the identification, evaluation, and integration of prototypes, support for middleware and applications, and development of hardware including antennas and tags. These labs will allow companies to test RFID systems in real customer environments and iron out any potential problems. In addition, to ensure that different scenarios can be tested at the center, some of these labs have built prototypes for a number of industries, including pharmaceuticals, retail, logistics, manufacturing, electronics, government, and transportation.

Standards that will define test methods for certain group (and application) of RFID systems are under development. For example, ISO/IEC 18046:2006, "Information Technology—Automatic identification and data capture techniques—Radio frequency identification device performance test Methods," First edition, 2006-11-01, was prepared by Joint Technical Committee ISO/IEC JTC 1, Information Technology, Subcommittee SC 31, Automatic Identification and Data Capture Techniques. This first edition of ISO/IEC 18046 cancels and replaces ISO/IEC TR 18046:2005, which has been technically revised.

This ISO standard defines test methods for performance characteristics of RFID devices (tags and interrogation equipment) for item management and specifies the general requirements and test requirements for tag and interrogator performance, which are applicable to the selection of the devices for an application. Of particular significance is that these tests are defined for RFID devices having one antenna only; it does not apply to testing in relation to regulatory or similar requirements.

6.4 Review Questions and Problems

1. An RFID tag may be used in situations where tagged objects such as pallets or boxes travel on a conveyor belt at speeds up to 10 feet/second (3.048 m/s.) The tag spends a small amount of time in the read field of the RFID reader, meaning that a high read rate is required. In such cases, the RFID system must be carefully planned to ensure reliable tag identification. The other potential problem could be the *Doppler effect*. Calculate the Doppler shift [using (6.7), where f is the transmitting frequency, Δf is the change in the frequency of the reflected signal, and c is the speed of light] for this case, at 915 MHz, and determine if the operation could be affected.

$$\Delta f = \frac{f \cdot v}{c} \tag{6.7}$$

   (*Answer:* The Doppler shift is around 9 Hz and will not affect the correct operation of the system.)
2. Describe one application of active and passive RFID technology and discuss and compare the following aspects: range, multiple-tag operation capabilities, data storage, sensor capabilities (temperature, humidity, shock, temper detection, security), business process impact, country specific and global standards, and coexistence with other technologies.
3. One example of a battery-assisted (semipassive) UHF RFID system is one that uses 50% duty cycle Manchester encoding for the forward link; this is possible because the RF signal does not have to power the tag. The tag is designed not for power gathering, but for optimized signal detection. This allows much smaller signals to be picked up and amplified on the chip. FSK backscatter signaling uses different frequencies to signal a high or low, making it easier to discern the signal from noise, because the reader must merely find a frequency, not a signal edge. Discuss the design summary for this system and decide where and when you would use this not-all-that-inexpensive system.
4. To provide information on an otherwise invisible tag detection process, some propose use of a so-called watchdog tag that would provide required transparency. Simply speaking, the watchdog tag is a sophisticated version of an ordinary tag, because it features an additional battery, a small screen, and potentially even a long-range communication channel. The watchdog tag's main task is to decode the commands transmitted by a reader and make them available on the screen of the

device for inspection by the user, or to log all data transfers and provide consumers with detailed summaries whenever needed. Although the watchdog tag could be carried by the user as a separate device, its functionality could also be integrated into a mobile phone, allowing it to leverage the existing display, battery, memory capacity, and long-range communication features of the phone. How much complexity would this additional feature add to reader–tag protocols? Would the read speed and the number of tags read be affected? What are your thoughts on usefulness of watchdog tags?

5. A reasonable definition of privacy is necessary to consider various policy choices. However, privacy is a notoriously difficult concept to define. The United Nations codifies that "no one shall be subjected to arbitrary interference with his privacy" as a basic human right. It has even been suggested that "all human rights are aspects of privacy." In 1890, Supreme Court Justice Louis Brandeis famously articulated privacy as the "right to be left alone." Ruth Gavison of the *Yale Law Journal* defined three core aspects to privacy: secrecy, anonymity, and solitude. Considering these issues, Simson Garfinkel has written an "RFID Bill of Rights" based on the U.S. Department of Health and Education's Code of Fair Information Practices [9]. Garfinkel's RFID Bill of Rights reads as follows:

> Users of RFID systems and purchasers of products containing RFID tags have…
> a. The right to know if a product contains an RFID tag,
> b. The right to have embedded RFID tags removed, deactivated or destroyed when a product is purchased,
> c. The right to first class RFID alternatives: consumers should not lose other rights (e.g., the right to return a product or to travel on a particular road) if they decide to opt-out of RFID or exercise a RFID tag's "kill" feature,
> d. The right to know what information is stored inside their RFID tags and what information is associated with those tags in associated databases. If this information is incorrect, there must be a way to correct or amend it.
> e. The right to know when, where and why an RFID tag is being read.

Which one of these rights will be most difficult to implement and why? List these rights from the most important to the least important ones. Do you agree with all of them? Add a few other requirements/rights that are important for you and the society in which you are living.

6. *Operating outside the allocated spectrum:* In many jurisdictions it may be possible to obtain special permission from the regulators of the radio spectrum to use equipment that operates outside the prevailing legislation. For example, the testing and development of new forms of RFID equipment or the temporary use of noncompliant RFID equipment may be allowable on a case-by-case basis. A large number of factors will typically be taken into account in such circumstances, such as the location of the equipment to be deployed and the potential for interference with legitimate users of the spectrum. This approach has been used by some companies in Europe that desire to start their RFID trials with equipment conforming to North American legislation. In these cases, the anticipated use of the noncompliant equipment is for a limited period and operation of the equipment has been modified to minimize the chance of it causing interference.

   Assume that you are working as a RFID system designer for a large company that would like to test and potentially implement a newly designed RFID system in a frequency band that was not originally allocated to RFID. What kind of approvals and licenses would you require in the area in which you are living? Where would you start your regulatory battle to implement a nonstandardized system? What would happen if you decided next year to open warehouses in another country and wanted to use the same RFID system over there? Is international harmonization a good idea? Could international harmonization slow down or expedite your company's project? Discuss the answers to these questions in detail.

7. Item management includes both the identification of an item and its location. Whereas RFID provides a means of radio identification, RTLS provides a means of radio location. The *real-time locating system* (RTLS) is being applied in a number of industries, ranging from health care to manufacturing. RTLS has an important role to play as far as real-time data is concerned and where assets are required to be located in transit. According to an analysis, it is expected that the revenue generated by RTLS will reach $1.26 billion by 2011. Shortcomings in the existing systems have initiated the need for a real-time locating system. The ability to generate real-time data is the major driver behind RTLS.

   a. What type of RFID system would you use for the application shown in Figure 6.10? What other components of the network are required to make the system work? Discuss the project.

   b. Condition monitoring is another valuable function that can be added to supply chain or asset management applications. Tempera-

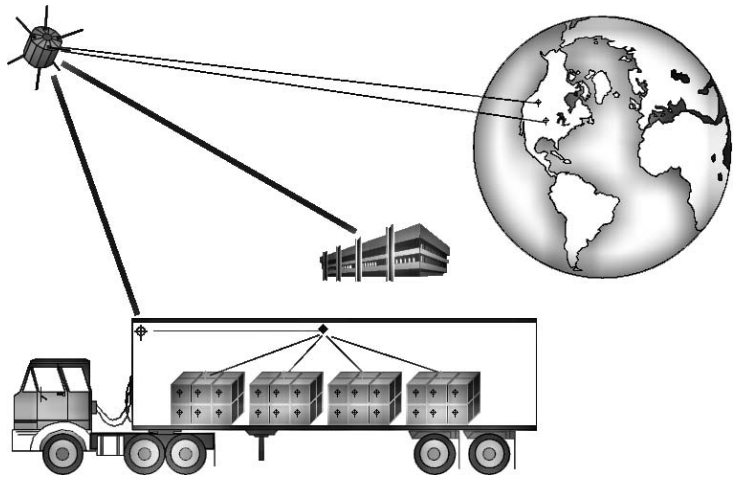


Figure 6.10 Satellite/RFID system for the tracking of transported goods.

ture, humidity, access, and use of assets can be monitored. Discuss adding condition monitoring to the previous project of location monitoring.

8. One priority for RFID use is hardware certification since RFID interoperability is a condition for successful open applications. Tags from any vendor must communicate with readers from any vendor, and tagged objects must be readable in a variety of application conditions. Discuss the RFID certification process in the country in which you are living.
9. The EU's *Directive on Waste Electrical and Electronic Equipment* (WEEE 2002/96/EC Directive) does not explicitly rule out the possibility that RFID chips will be seen as waste electrical and electronic equipment. With RFID technology becoming increasingly more widely adopted, do you think that RFID chips should be treated as dangerous material? Does the size of RFID chips play a role in making that decision? Should active tags be discarded differently than passive tags?

References

[1] Chawathe, S. S., et al., "Managing RFID Data (Extended Abstract)," Computer Science Department, University of Maryland, 2005.

[2] Avoine, G., and P. Oechslin, "RFID Traceability: A Multilayer Problem," Lausanne, Switzerland: Ecole Polytechnique Fédérale de Lausanne, 2006.

[3] Jain, S., and S. R. Das, "Collision Avoidance in a Dense RFID Network," Stony Brook, NY: Computer Science Department, Stony Brook University, 2006.

[4] Kim, S., et al., "Reader Collision Avoidance Mechanism in Ubiquitous Sensor and RFID Networks," Sungbuk-ku, Seoul, Korea: Department of Electronics Engineering, Korea University, 2006.

[5] Engels, D. W., et al., "Colorwave: An Anticollision Algorithm for the Reader Collision Problem,"*IEEE Int. Conf. on Communications,* Vol. 2, 2002, pp. 1206–1210.

[6] Stallings, W., *Data and Computer Communications,* 6th ed., Upper Saddle River, NJ: Prentice-Hall, 2000.

[7] Juels, A., "RFID Security and Privacy: A Research Survey," RSA Laboratories, September 28, 2005.

[8] Tektronix, "RFID and NFC Measurements with the Real-Time Spectrum Analyzer," Application Note, 2005.

[9] Weis, S. E., "Security and Privacy in Radio-Frequency Identification Devices," Cambridge, MA: Massachusetts Institute of Technology, 2003.

7

RFID Sociocultural Implications

7.1 Market Trends and Usage

Many new and promising RFID applications are in the works. For example, medicinal products can be tagged and traced to combat drug counterfeiting, and logging tagged items into and out of your refrigerator can help you track when certain products are out of stock or whether certain products have gone beyond their expiration date. An RFID chip can provide useful information over the whole life cycle of its tagged product. That is why RFID chips can improve customer relations through better after-sales services. There is a large field of applications when it comes to medical services and services for people with other types of needs. Even the tracking of criminals on parole from prison is imaginable. One of the most popular new developments is a contact-less payment solution called SmartPay for small payments in the United States. It is also an efficient technique for reducing thefts from shops, in addition to stock keeping functions. Looking toward the future, RFID and smart tags will allow the creation of an Internet of things, where objects and locations may be directly related to one another. These objects will also be capable of increasingly intelligent interaction.

Apart from its expected benefits, the more intensive and extensive use of RFID in the future also raises major issues in the areas of privacy, security, technological reliability, and international compatibility. One key challenge for decision makers is to create a common vision and a set of goals on how RFID can keep companies (and countries) more innovative and competitive in the world economy. At the same time, citizens must have the tools and freedom of

choice they need to protect their privacy and security. At present, technological challenges such as the lack of global frequency standards, low reading rates, interference with other radio sources, insufficient encryption capabilities, and the cost of implementation and end-user concerns prevent wide adoption of RFID technology. So, the main issues to be addressed are consumer privacy, standards and interoperability, harmonization of the frequency spectrum, intellectual property rights, and future research needs.

What will be the social consequences of a world full of embedded RFID tags and readers? Will our privacy be eroded as RFID technology makes it possible for our movements to be tracked and allows our personal information to be available in unprecedented detail? These and many other questions must be answered before RFID systems become commonplace. One of the major worries for privacy advocates is that RFID tags identifying individual items purchased with credit or debit cards will link buyers to the specific items in the card's or the store's databases. Marketers could then use these data to keep track of exactly what particular people bought, down to the color, size, style, and price—more information than UPC barcodes reveal. In an amplification of the way that phone and direct-mail solicitors use similar, less accurate data to target people for sales pitches, those equipped with RFID-derived data might hone in on consumers with very specific sales pitches.

Another concern is that RFID equipment will produce automatic audit trails of commercial transactions: In a totally tagged world, it will be easier to detect when we lie about how we spent our time or what we did and where. This capability could have great consequences for the workplace, and the legal system might look to using logs kept by tag readers as courtroom evidence. We may need laws to specify who can access data logs and for what purpose. In Europe, the Data Protection Act already limits access to computer records of this kind, and the United States will probably enact similar legislation.

The problem does not lie with RFID technologies themselves; it is the way in which they are deployed that raise privacy concerns. Privacy and security must be built in from the outset (i.e., at the design stage); just as privacy concerns must be identified in a broad and systemic manner, so too must technological solutions be addressed systemically. A thorough privacy impact assessment is critical. Users of RFID technologies and information systems should address the privacy and security issues early in the design stages, with a particular emphasis on data minimization. This means that wherever possible, efforts should be made to minimize the chance to identify, observe, and link the RFID tags with personal information and other associated data. Use of RFID information systems should be open and transparent, and offer individuals as much opportunity as possible to participate and make informed decisions.

7.1.1 Price Barriers to Adoption

National and international regulatory authorities are trying to enforce at least some regional interoperability and, in some cases, international interoperability, but progress is slow and competition for valuable bandwidth is strong. There are currently two principal barriers to global implementations for low-power tags: the regulations and the laws of physics.

The saying "high speed—high frequency—high capability—high cost" still applies. With these constraints, RFID systems can now generally be grouped into two types, although with some exceptions:

- Low-cost, but capable, passive tag systems;
- High-cost specialized tag systems for operation at high speeds and long distances.

RFID supporters envision a world where RFID reader devices are everywhere—in stores, cars, clothes, factories, and even in our home refrigerators. But RFID tags will not become ubiquitous in consumer products as long as the price of the tags is viewed as prohibitively expensive by many businesses. RFID tags currently cost from U.S. 20 cents to $1 each, which still makes them impractical for identifying millions of items that cost only a few dollars. Some experts predict that in quantities of 1 billion, RFID tags would approach 10 cents each. The holy grail of 5-cent tags, which is the stated primary goal of the Auto-ID Center, would be attained in lots of 10 billion. More recent technological developments may put a 1-cent tag within reach, which, in turn, would fuel demand for RFID comparable to that for barcodes.

Makers of RFID chips and so-called inlays, which include the chip, antenna, and substrate, have been trying for years to reduce prices for RFID tags to 5 cents. The type of materials and assembly methods used to package tags impact the final cost directly (around 30%) and to some extent the communication performance. In the supply chain, the cost of tags is one of the main considerations for mass adoption, with the 5-cent tag being the much-talked-about target. How to achieve this figure is currently one of the great debates. Traditionally, chip die size has always been the key focus, and IC companies have managed to get die sizes (chip area) down to around 0.3 mm^2 for UHF chips, resulting in a manufacturing cost of about 1 to 2 cents, depending on the silicon process, leaving 3 cents for the rest of the costs. This is where the real challenge now seems to be.

However, these less expensive tags may lack the capabilities of their more expensive counterparts. As a result, most manufacturers have been content to cut prices at a steady rate of about 5% to 10% per year since 2000 while improving the technology. As a result, users of the tags are employing them in

applications no one dreamed of a decade ago, despite their inability to reach the elusive nickel price.

At McCarran International Airport in Las Vegas, for example, operators attach bag tags with dual dipole antennas to luggage to ensure that RFID readers in the handling system can communicate with all bags, regardless of their orientation on conveyor belts. The technology integrates two antennas 90° from one another; thus, the RFID tags can communicate with the airport's RFID readers, no matter how baggage handlers toss the luggage onto conveyor belts. Such dual-antenna tags have not reached rock-bottom prices, but at roughly 20 cents each, they offer capabilities nickel tags cannot match. Similarly, retailers have begun using tags with specialized antennas to enable garments buried in stacks to successfully talk to RFID readers. Again, cheaper tags are unlikely to achieve such feats.

There seems to be a lot of attention given to the cost of tags. The benefits that can be derived from implementing RFID technology can far outweigh the cost of the tags. An investment in time and money is required; however, to be successful and get the best return on investment it is important to understand the technology and how it can benefit the organization. In other words, by understanding the technology, it is possible to reduce the overall cost of implementation.

7.1.2 Globalization

Widespread deployment of RFID relies on the availability of either dedicated or unlicensed bands. Current interest is in the UHF frequencies, which offer a good balance between antenna size and path loss. However, the requirements for these bands vary widely around the world, frustrating attempts to deploy systems in an era of global trade. Frequency band is just one of the challenges; given often-conflicting global constraints, with the implied requirement to recognize tags wherever goods might flow, the challenge is to build in support for multiple data rates, modulation formats, and interference environments through a flexible and programmable air interface. From the technical perspective, in order to make RFID a truly global technology, some basic requirements have to be fulfilled:

1. *Compatibility:* Suitability of products, processes, or services for use together under specific conditions to fulfill relevant requirements without causing unacceptable interactions. Interchangeability, interoperability, and noninterference are differing levels (or degrees) of compatibility.
2. *Interchangeability:* The condition that exists between devices or systems that exhibit equivalent functionality, interface features, and

performance to allow one to be exchanged for another, without alteration, and achieve the same operational service.

3. *Interoperability:* The condition that exists between systems, from different vendors, to execute bidirectional data exchange functions, in a manner that allows them to operate effectively together. A guarantee of a certain level of compatibility has to be achieved between different implementations of the same standard. The desired level of compatibility is specific to a given standard, and can be limited to basic services. *Interconnection and interoperability are the main objectives of standardization.*
4. *Noninterference:* The condition that exists when standard-compliant components of various types or of different vendor origins coexist within the same space without serious detrimental effect on one another's performance. Components are not necessarily required to communicate with one another as part of a common infrastructure, but merely to peacefully coexist.

7.2 RFID Security and Privacy Aspects

7.2.1 Access to Information

If you are lost in an airport or parking lot, an RFID-based system that can guide you to your gate or car would be appealing. So too would be the ability to return items to shops without receipts, either for refunds or warranty servicing, and RFID-enhanced medicine cabinets that ensure that you have remembered to take your medications. The concern is the effect on individual privacy of RFID-enabled computing systems that can automatically see everyday objects—the clothing on your person, the medical implants in your body, the prescription drugs you are carrying, the payment devices in your pocket, and perhaps even individual pieces of paper, such as banknotes and airline tickets [1] (see Figure 7.1).

To increase consumer acceptance of RFID technology, RFID advocates must promote and implement comprehensive security measures, along with consumer education, enforcement guidelines, and research and development of practical security technologies. Technical organizations (such as EPCglobal) are developing standards for the electronic product code, including EPCglobals' *Guidelines on EPC for Consumer Products.*

It is useful to understand the difference between on-tag and off-tag access control. As the name implies, on-tag access control mechanisms are located on the RFID tags themselves. On-tag access control is the most common type of RFID access control, with mechanisms including tag deactivation,

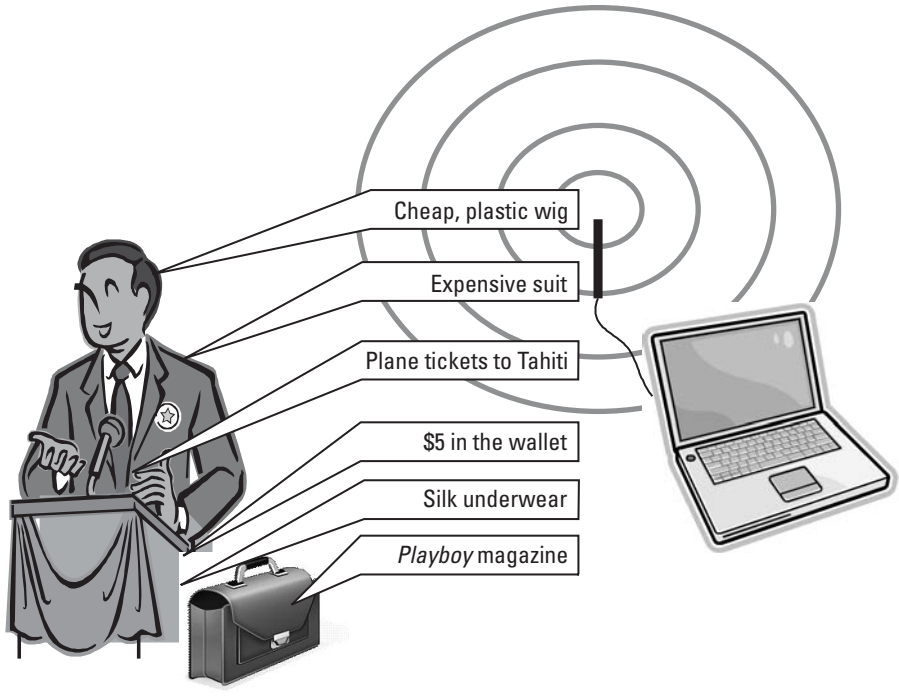


Figure 7.1 Potential RFID privacy threats.

cryptography, and tag–reader authentication. In contrast, off-tag access control mechanisms put the access control mechanism on a device external to the RFID tag. Examples of this include the RSA blocker tag, a special RFID tag designed to prevent readers from performing unwanted scanning and tracking of people or goods, without any disruption to normal RFID operation (developed by RSA Laboratories), and external re-encryption. Off-tag access control has the advantage that it can protect low-cost RFID tags (such as EPC tags), because the access control does not require any extra complexity (hence, extra cost) in terms of the RFID tag itself.

7.2.2 Privacy Threats and Protection

The impending ubiquity of RFID tags requires not only support mechanisms to provide adequate performance, but also measures to address privacy concerns associated with unobtrusive tags on everyday items. When people envisioned computing capabilities everywhere, embedded in the environment in such a way that they can be used without being noticed, they also acknowledged that the invisible nature of the computing devices will make it difficult to know what is controlling what, what is connected to what, and where information is flowing.

This tension between the contradicting requirements of control and privacy on the one hand, and usability and performance on the other, are well illustrated by the privacy concerns associated with the planned deployment of RFID technology in supermarkets and retail outlets. Two notable privacy issues complicate adoption of RFID systems:

1. *Leaking information pertaining to personal property:* If a generic dumb RFID system is used, anyone can read, without restriction, the connection between the product and the tag and obtain information regarding the tagged contents of, say, a purse or any tagged item worn on the body in a manner about which the possessor is unaware.
2. *Tracking the consumer's spending history and patterns and physical whereabouts:* If a product ID is specific to an individual (when, say, tags are used in clothes and other personal belongings, such as shoes, watches, handbags, and jewelry), tracking the person's movements over an extended period becomes an option. Not only can physical location be tracked, an individual's personal information (stored on multiple independently managed databases) might also be accessible based on a unique ID.

These RFID privacy threats follow from the basic functionality of RFID technology: An ID can be read without permission, is constant and unique, and contains potentially sensitive data. A number of proposed RFID privacy protection schemes are classified based on the new functionality they implement in RFID technology; they range from adding only memory to adding lightweight circuits. Each involves a trade-off between the cost of the tag and the value of privacy protection. Several approaches are discussed next.

7.2.2.1 Kill Function

The EPCglobal standard specifies that tags must be equipped with at least one nullification function as a way to address public opposition. This function, called the *kill command,* disables the functionality of the tag after consumers purchase a product. It involves a high degree of consumer privacy protection at negligible cost; however, because the disabling process is performed manually by millions of individual consumers, human error is always a possibility. Moreover, the major problem in killing the tag is that the various RFID stakeholders would no longer be able to take advantage of the future emerging services that would rely on the millions of RFID tags likely to be dispersed throughout the consumer environment.

This simple countermeasure, a built-in option designed to kill the functionality of an RFID tag when the consumer leaves the store, has been incorporated into the EPCglobal standard (Class 1 Generation 2 UHF Air Interface

Protocol). For consumers, its purpose is easy to understand and thus easy to accept. However, killing a tag's functionality curtails the future potential use of RFID in consumer services, such as in smart refrigerators that automatically reorder food products, expiration date and product recall alarms, and personal library management.

7.2.2.2 Normal Tags and Smart Tags

Other privacy protection schemes generally reflect two main approaches: normal tags and smart tags. The *normal tag* approach protects individual consumer privacy without having to modify the existing tag or cost the user organization more money. The normal tag approach achieves privacy protection by preventing the unauthorized reading of the output from the tag, blocking electric waves with aluminum foil, or jamming waves to interfere with a tag's ID being read by an adversary's unauthenticated reader. *Smart tags* are equipped with additional components, such as rewritable memory, basic logic circuits, hash function units (turning data into a relatively small number that may serve as a digital fingerprint of the data), and common-key/public-key encryption units.

Tag cost, security level, and scalability are likely to be the key factors in any trade-off equation calculated by any organization thinking about implementing these schemes. When the tag incorporates rewritable memory, the reader rewrites the information in the tag to achieve privacy protection. This approach is notable for its low cost, because the tag requires only rewritable memory.

On the other hand, a lightweight circuit is incorporated into the tag, and a re-encrypted ID to the reader is calculated by the circuit. Although public-key cryptosystems come close to providing good privacy protection, they are not suitable for tags because public-key primitives are complex and costly. A noteworthy scheme employing this technology is the *hash-chain scheme,* in which a hash function circuit is embedded in the tag, and the tag response is calculated by the hash function. The scheme holds down the cost of the tag because the hash function is lightweight, pseudorandom, and one way. Here, *pseudo-random* means that the output of the hash function is computationally indistinguishable from a true random value. Being *one way* means it is computationally unfeasible to compute the input of the hash function from the output of the hash function. The scheme addresses ID leakage and tracing problems through the pseudorandomness of the hash function, which prevents leakage and tracing. Moreover, the scheme is forward secure; that is, even after the tag's secret is exposed through tampering, the tag's past history cannot be traced due to the hash function being only one way. The drawback to the hash-chain scheme is that the load on the server is proportional to the number of tags, though the load can be reduced through advanced computation.

7.2.3 The Blocker Tag

The RFID blocker tag takes a different approach to enhancing RFID privacy. It involves no modification to consumer tags. Rather, the blocker tag creates an RF environment that is hostile to RFID readers. The blocker tag is a specially configured, ancillary RFID tag that prevents unauthorized scanning of consumer items. In a nutshell, the blocker tag spams misbehaving readers so they can't locate the protected tags' identifiers. At the same time, it permits authorized scanners to proceed normally [2].

The blocker tag spoofs the Tree-Walking Protocol into thinking that all tags, that is, all identifiers, are present. To do this, it simply emits both a 0 and a 1 in response to all reader queries. The result is that the reader attempts to traverse the entire identifier tree, believing that all possible tag identifiers in the world are present. The reader stalls because the tree is far too big to be fully scanned (for Class 1 EPC tags, the tree would have 2^{96} nodes).

7.2.4 Reader Signal Energy Analysis

One approach to RFID privacy does not rely on logical protocols at all. A system has been proposed, based on the premise that legitimate readers are likely to be quite close to tags (such as at a checkout counter), whereas malicious readers are likely to be far away (such as a competitor in the parking lot). In preliminary experiments, it was found that a reader signal's SNR decreases as the distance increases; that is, the farther away a reader is, the greater the noise level in the signal a tag receives. With some additional circuitry, therefore, an RFID tag might be able to obtain a rough estimate of the querying reader's distance and change its behavior accordingly. A tag interacting with a distant reader might only reveal the type of product it is attached to, a pair of trousers, for example. When interacting with a nearby reader, however, the tag might also reveal its unique identifier. A more sophisticated, multitiered approach is also possible, in which tags furnish increasing amounts of information as readers get closer.

Of course, distance alone does not provide an ideal trust metric. But distance could be combined with traditional access control techniques, such as a challenge–response protocol between the reader and tag, to achieve a more comprehensive approach to RFID tag privacy. Indeed, the distance-measurement approach is complementary to blocker tags.

7.2.5 Protecting the Public

The Federal Trade Commission held a hearing in June 2004 to "facilitate discussion of the public policy issues surrounding the use of RFID and to encourage the development of best practices for RFID that do not compromise consumers' privacy and security." Their March 2005 released report [3] finds

that the privacy issues associated with RFID are linked to database security, and that industry can play an important role in addressing privacy concerns raised by some RFID applications. The report emphasizes the importance of industry self-regulatory programs, meaningful accountability provisions to help ensure compliance, and implementation of reasonable and appropriate measures to protect data collected by RFID systems.

Legislators in several states, recognizing privacy concerns stemming from the use of RFID, have introduced bills that seek to respond to the increasingly rapid adoption of RFID technology. Although none of the proposed pieces of legislation has been passed into law, the introduction of these bills signifies that RFID-related technologies appear to be generating concerns within the legislative branches of state and federal governments. According to some, legislation restricting RFID use at this early stage would likely stifle the technology and delay deployment in the marketplace. It would be more productive to monitor the technology during the next few years, while engaging with the business and government sectors regarding their respective use of RFID and their policies on maintaining RFID privacy and security. The Competitive Enterprise Institute (a nonprofit public policy organization advocating nonregulatory, market-based solutions) states that as RFID technology comes into full use, various social forces would constrain it more suitably than government regulation.

A number of governments around the world, citing a need to ensure the protection of individual privacy rights, have raised concerns that the collection, storage, transfer, and use of personal information through RFID technology could possibly violate individuals' privacy rights. For example, in the European Union, the EU Article 29 Working Party of Member State Data Protection Authority has recently expressed its concern that RFID technology may contravene the requirements of the EU Directive on Data Protection. Accordingly, in recent months, the European Commission has held a number of workshops and issued inquiries concerning the privacy implications of RFID. The Asia-Pacific Economic Cooperation forum is considering the relationship of RFID privacy to its recent privacy guidelines. In particular, South Korea has called for the development of RFID privacy guidelines in the forum's Electronic Commerce Steering Group. In 2004, Japan also issued privacy guidelines for RFID. Finally, the Organization for Economic Cooperation and Development's (OECD's) Working Party on Information Security and Privacy is currently reviewing the scope of policies and concerns with the global use of RFID on security.

The U.S. Department of State is implementing its conversion program for the RFID-based electronic passports, or *e-passports*, despite warnings from security experts that these passports could be accessed or tracked by the wrong individuals. In fact, some security experts feel that the technology contained in this type of passport could be used by terrorists to construct a bomb designed to target anyone of their choosing. There is also some concern that e-passports do not

have enough security embedded to resist hackers and the advancement of technology. Nevertheless, in August 2006, the State Department began issuing e-passports containing RFID chips out of its Denver and Washington, D.C., passport offices, with full production to begin by mid-2007. The plan is for all U.S. passports to include RFID chips containing personal biometric information by 2017. New U.S. e-passports contain a 64-kbit RFID chip with personal information about the passport holder. U.S. Department of Homeland Security officials claim that the passports must be held within 10 cm of a reader to have their data read.

7.2.6 Fair Information Practices

The Fair Information Practices (FIP), published by the OECD in 1980, are a well-established set of guidelines for consumer privacy (http://www.oecd.org). They have their roots in a 1973 report from the U.S. Department for Health, Education, and Welfare and were drawn up by the OECD to better facilitate the cross-border transfer of customer information as part of trade between its member states. The eight principles can be summarized as follows:

1. *Collection limitation:* Data collectors should only collect information that is necessary, and should do so by lawful and fair means, that is, with the knowledge or consent of the data subject.
2. *Data quality:* The collected data should be kept up-to-date and stored only as long as it is relevant.
3. *Purpose specification:* The purpose for which data is collected should be specified (and announced) ahead of the data collection.
4. *Use limitation:* Personal data should only be used for the stated purpose, except with the data subject's consent or as required by law.
5. *Security safeguards:* Reasonable security safeguards should protect collected data from unauthorized access, use, modification, or disclosure.
6. *Openness:* It should be possible for data subjects to learn about the data collector's identity and how to get in touch with him or her.
7. *Individual participation:* Data subjects should be able to query data collectors as to whether or not their personal information has been stored, and, if possible, challenge (i.e., erase, rectify, or amend) this data.
8. *Accountability:* Data collectors should be accountable for complying with these principles.

The FIP form the basis for many of today's privacy laws, such as EU Directive 95/46/EC (April 1995), which provides the framework for the national privacy laws of all EU member states. For example, Article 6 of the directive requires data collectors to collect only as much information as necessary (also called the proportionality principle or the principle of data minimization), while Article 7 requires them to obtain the unambiguous consent of the data subject before collection. It is undisputed that the act of reading out one or more RFID tags can constitute a data collection process, meaning that existing privacy laws also apply to the communication between tags and their readers. This has also been recently pointed out by the International Community of Data Protection and Privacy Commissioners; at the outset, this would mean that RFID readers would need to be openly announced with the help of public signs and placards explaining the purpose and extent of the data collection, as well as the identity of the data collector.

More comprehensive unofficial text of the EU Directive 95/46/EC can be found at the Center for Democracy and Technology's Web site (http://www.cdt.org).

7.3 Health Risks from RFID

Electricity and electromagnetic fields (EMFs) bring countless benefits to society. We cannot do without them, yet we do not know the consequences of long-term exposure to EMFs, if indeed there are any. Therefore, research is needed to understand the risks and set appropriate safety standards. Scientists researching the health effects of nonionizing EMFs have to contend with an enormous electromagnetic spectrum, reaching from static electric and magnetic fields to EMFs at frequencies in the terahertz range. They also have to contend with the complexity of the human organism. The number and nature of the ways in which these two systems interact can at present only be guessed at. Behind the scenes, dynamic social and political forces are at work among an enormous array of stakeholders with different interests in the field: the scientists and research groups themselves; policy makers, politicians, and governments and regulatory bodies at the local, national, and international levels. And we must not forget health professionals, industrialists, investors, trade associations, trade unions, marketing professionals, users of devices reliant on EMFs, patients, and so on; we should not overlook for a moment the modern media as well.

Today, hundreds of thousands of RFID scanners and EAS systems are in use. All of these systems utilize EMFs to detect and scan tags. According to some sources, RFID systems pose no threat to the health of ordinary people; this might be true, but prolonged occupational exposure may occur. The report by

the International Commission on Non-Ionizing Radiation Protection (ICNIRP) examines the effects of EM radiation on humans [4]. The report describes mechanisms of thermal and nonthermal interaction between EMFs and biological systems. Thermal interaction is the heating of tissue, which can cause damage. The most notable nonthermal interaction is brain stimulation, which could alter the membrane potentials at a cellular level and might have adverse effects on the nervous system. The report states that the high frequencies of the EAS/RFID systems produce no heating or thermoregulatory stress. However, EAS and RFID devices may interact with medical devices such as pacemakers, which can cause dangerous situations and indirect health hazards. The report recommends that further studies should be made on this area and that device manufacturers should provide information needed for health risk assessments. There is also a need to continue to collect exposure data, especially for occupational groups. If possible, low-frequency and high-frequency exposure should be differentiated.

It is a well-known fact that strong EMFs can interfere with electronics. Hospitals contain many devices that are critical, for example, life-support equipment that must not be disrupted. Today many hospitals have banned the use of cellular phones because of the risks of interference. How are RFID applications compared to this? Are they a potential source of interference, too?

Since 1993 there has existed an electromagnetic compatibility (EMC) standard for medical devices, IEC 60601-1-2. This standard, however, was originally only specified for frequencies lower than 1 GHz; in 2001, it was updated to cover frequencies up to 2.5 GHz. In the standard, it is specified that life-support equipment must be able to operate in the presence of field strengths up to 10 V/m and equipment that does not support life, up to 3 V/m [5].

RFID transmitters and similar short-range radio devices are regulated differently in Europe and the United States. In Europe regulations are country specific but based on the CEPT regulations. In the United States, radio devices have to conform to the rules of the FCC.

Recommendations

- Continue to monitor current research on the health effects of exposure to radio frequencies from RFID equipment to verify whether new information indicates health risks.
- Ensure that any RFID equipment purchased meets the FCC requirements protecting users from any possible thermal effects and includes all possible measures to minimize the risks of emissions from the system interfering with electrically powered active medical devices, such as pacemakers.

- Purchase RFID equipment that meets the human exposure specifications of nonregulatory agencies such as ANSI, the IEEE, the NCRP, the American Conference of Industrial Hygienists (ACGIH), and the ICNIRP.
- If RFID equipment is installed, conduct measurements of the RF emissions levels generated from systems to ensure that the emissions are below those recommended under the OSHA nonionizing standard (CFR 29, Section 1910.97) and by the nonregulatory agencies listed in the previous point.
- If RFID equipment is purchased, ensure that it is installed and maintained according to the manufacturers' specifications.

7.4 Ethical and Moral Dilemmas of Technology

The human species is the only species on this planet with the capacity to direct its own evolution (genetic engineering, cloning, and so on) at its own accelerated pace and without being confined any longer by the will of nature or the excruciatingly slow pace of evolution. This may be the ultimate challenge for the next generation of scientists (and humans in general). It is also critical to realize that technology has no moral value; it is neither good nor evil. Rather, it is the application of the technologies that raises the moral and ethical issues [6].

Prostheses (which started with the humble limb or hip prostheses or cardiac pacemaker) are becoming intelligent, with embedded microsensors, and are being custom manufactured for individuals, providing capabilities beyond natural systems. Some prostheses are being controlled by direct connections to the brain, as in the experiment in which a monkey brain implant controls a remote robotic arm. Considerable research is being devoted to neural (brain) implants or a direct brain-machine interface [7]. What are the implications and complications of communicating directly with computers or connecting to the Internet or, perhaps, of direct brain-to-brain interactions? Who will pay for the expensive bioreactors to grow, and tissue banks to store, the organs? If people receive supranormal sensory organs (eyes, ears, muscles), will they be treated as anomalies or freaks? Politically, will there be a new class of people who can afford the technology, who will live longer, and who will perhaps be politically more powerful? And finally, what will it mean to be human if nearly all of your body is made of replacement parts or if you do not look human? Is this a first step toward building humanoids and cyborgs?

Current work on RFID nanotags (for example, body implants to be inserted into the body for identification) provokes thoughts of Big Brother and of other privacy issues. Morally, should we implant these identification tags into criminals or other people to track them, especially against their will? Do we

accept this technology and its applications or not? Different people will have different answers to that question; and the same people could have a different answer today from the answer they gave 5 years ago and a yet completely different answer 5 years in the future. As recently as 50 years ago, a person would have been considered totally insane to say that a human would walk on the Moon (jet planes were just becoming commonplace, and no rocket had ever been launched), but *Sputnik* in 1957 gave serious scientific pause, and the first footstep on the Moon on July 20, 1969 (only 12 years later!), proved that hypothetical person to be a visionary.

Genetic engineering, human cloning, tissue engineering, intelligent robotics, nanotechnology, suspended animation, regeneration, and species prolongation are but a few of the ideas that will revolutionize what it means to be human and what the ultimate fate of the species may be. Although not all of the implications of the new technologies can be foreseen at this time, it is critical to identify the candidates likely to disrupt our conventional thinking and investigate their social, behavioral, political, moral, and ethical implications. Although many critics would protest that some of the technologies cannot be accomplished in the next two to three decades, the issues are so profound that even longer time spans may be inadequate to prepare for the consequences. There are numerous implications and consequences of advanced technologies; however, the following examples address some of the known issues. These issues fall into several categories: scientific (is the science really safe?), social (what are the societal implications?), behavioral (how will individuals' behavior change?), political (how will the legal and regulatory systems react?), and philosophical (what fundamental moral and ethical precepts are challenged?).

In this short section, as well as elsewhere in this book, we have tried to raise a few of the potential moral and ethical dilemmas and questions facing society in these advanced technological times, without even attempting to answer any of them. In addition, there is an important question about whose definition of ethics and moral fabric of the society to use when we evaluate whether something is moral or ethical, and whose right is to make that decision on behalf of the rest of the humanity?

7.5 Other Developments in Auto-ID Systems

7.5.1 RuBee

RuBee is the commercial name for what is officially known as long-wavelength ID (LWID), as defined by the IEEE. The moniker was given to the technology by engineers at Miami-based Visible Assets, who coined the name for LWID technology after the hit 1967 Rolling Stones' song, "Ruby Tuesday."

RuBee networks operate at long wavelengths and accommodate low-cost radio tags at ranges to 100 feet. RuBee networks and tags are distinguished from most RFID tags in that they are unaffected by liquids and can be used underwater and underground. RuBee devices will be able to be used as implantable medical sensors having a 10- to 15-year battery life, depending on the number of reads and writes. The ability of RuBee tags to maintain performance around steel, meaning that they would work well when steel shelves are present, removes a key obstacle for low-cost deployment of RFID in retail, item-level tracking environments.

Tags based on the RuBee technology can be either active or passive, and all operate at a low frequency of 132 kHz, rather than the HF (13.56-MHz) or UHF (916-MHz) ranges used in the most widely deployed RFID systems today. Because low frequencies have significantly smaller bandwidth for data transfer relative to higher frequencies, only about 6 to 10 RuBee tags (active or passive) can be read per second, whereas several dozen passive HF tags and several hundred passive UHF tags can be interrogated in that same span of time. This is not seen as a problem because RuBee users are interested in tracking the locations of assets rather than the passage of fast-moving tagged goods through portals, for which passive UHF tags are optimized.

The new IEEE RuBee standard, called P1902.1, "RuBee Standard for Long Wavelength Network Protocol," will allow for networks encompassing thousands of radio tags operating below 450 kHz and represents candidate specification, which the standards group is using as a starting point for the future standardization of the technology.

IEEE P1902.1 will offer a real-time, tag-searchable protocol, using IPv4 addresses and subnet addresses linked to asset taxonomies that run at speeds of 300 to 9,600 baud. RuBee networks are managed by a low-cost Ethernet-enabled router. Individual tags and tag data may be viewed as a stand-alone Web server from anywhere in the world. Each RuBee tag, if properly enabled, can be discovered and monitored over the World Wide Web using popular search engines (e.g., Google).

7.5.2 Visible Light Tags

Visible light tags use visible light instead of radio waves for communication between tags and readers. Visible light tags can be used, for example, in a hospital where the use of radio waves is restricted, or underwater. The read range and scope can be modified by using optical devices, and people can visually see the communication range/scope, since it uses visible light. They are upwardly compatible with ISO15693, can replace part of the existing RFID infrastructure, and can be used in hospitals, underwater, and other places where radio communications do not work. Obviously, visible light tags are not restricted by radio

interference, so there is no need to obtain a license. Also, they can be installed in an environment that is already crowded with RFID readers.

7.5.3 RFID and Printable Electronics

The term *printable electronics* refers to circuitry created out of conductive inks using a wide variety of printing technologies, old and new [8]. Much of the buzz in printable electronics is about ink-jet printing, offering cost-effective device creation in very small volumes. But there are many other ways of creating printable electronics, including nanoimprint lithography (NIL), offset lithography, gravure, and flexographic printing. Different applications may be suited to different printing technologies, since each application has its own requirements and each printing technology has its own advantages and disadvantages. For example, ink-jet printing holds out the prospect of the economic creation of customized/small production run circuits, whereas conventional printing processes using masks are well suited to something closer to mass production. NIL, a technique that may or may not be considered part of the printed-electronics sector, is probably the only production approach considered today that might genuinely be considered to be capable of creating features at the nanoscale.

The ability to cost effectively create circuitry as part of other printed products is finding early use in greeting cards, but could become a big revenue generator through the success of printed RFIDs and smart packaging. It is possible to imagine some future smart-packaging product, entirely created by multiple inline printing processes, that encompasses high-quality graphics, RFIDs, sensors, and even a small display to show pricing as it varies.

7.5.4 RFID and Mobile Phone Integration

RFID can be combined with mobile phone technologies by inserting either a transponder or a reader into a smart phone. A smart phone featuring a transponder connects to the wireless network and communicates with RFID readers, while a smart phone featuring a reader also connects to a wireless network but instead retrieves RFID tag information. The smart phone transponders transfer their data to readers over the phone network. Communication between transponder and reader must be secured by cryptographic means because RFID is subject to the risks of virus and hacker attacks. As with RFID in general, communication between transponder and reader depends on the range. All information is stored in the wireless network's central database. Smart phones with RFID readers can be used for detecting RFID-tagged products' Web sites via the wireless network to get additional information on, for example, materials used, a product's origin, and so forth; for locating the coordinates of anything or anyone that has been RFID tagged; for retrieving information from tagged items

(computers, cars, furniture, ads); and occasionally for updating information via the RFID reader.

RFID tags can also be used to call people. Combining RFID with cellular technology enables instant access to information, which can lead to improved services, work efficiency, and performance. RFID provides direct, up-to-date data and diminishes manual typing of passwords, keys, and codes. Overall, RFID combined with mobile technology facilitates processes the same way RFID does in general.

7.6 Review Questions and Problems

1. Last week Edward B. Someone tried out a brand-new wireless LAN card on his laptop at work. He didn't expect anything to happen, because his organization's wireless LAN wasn't up and running yet. But to his surprise, he was able to connect without any trouble to the network of an office down the street. Oops! Discuss the fact that space around us is full of radio waves containing data and proprietary—as well as personal—information. Was Edward committing a crime by accessing someone else's network? Discuss different scenarios and outcomes.
2. Consumers have concerns about the privacy and use of their personal information. Many of the companies and organizations promoting the use of RFID have developed or are developing policies addressing consumer notice and privacy. However, it is important to remember that many applicable consumer protections are already written into law. For example, retailers are already restricted in the sale or distribution of consumer information, and secure computer systems and data encryption schemes are already in place for the electronic transfer of private data. Federal guidelines are already in place that address many of these concerns including the Privacy Act (1974), the Electronic Communications Privacy Act (1986), the Telecommunications Act (1996), the Health Insurance Portability and Accountability Act (1996), and the Financial Modernization Act (the Gramm-Leach- Bliley Act of 2000), among others. Do you think that we need new, RFID-specific legal protections?
3. Think about and discuss the following statement: *"How would you like it if, for instance, one day you realized your underwear was reporting on your whereabouts?"* (California State Democratic Senator Debra Bowen, at a 2003 hearing). What other famous (and erroneous)

statements in the history of science and technology can you compare this statement with? Two similar examples are given here:

- *"There is no reason anyone would want a computer in their home"* (Ken Olson, president, chairman, and founder of Digital Equipment Corp., 1977).
- *"640k memory ought to be enough for anybody"* (Bill Gates, Microsoft founder, 1981).

4. One concern of RFID technology is that UHF tags can be read wirelessly at distances of 30 feet (HF tags can be read at less distance, but still wirelessly). Tags can obviously be removed and destroyed after purchase, but that makes returning a product or recalls more difficult. The Gen 2 protocol offers a *kill command* to deactivate tags, but tags that are killed cannot be revived. One novel idea is to use so-called *clipped tag* that gives consumers the option of privacy protection by allowing them to tear off perforated sections that hold parts of the tags' antennas, reducing their read ranges to only a few inches. This provides a visible means of enhancing privacy protection and allows for later use of the tag for returns and recalls. Discuss this idea and list its pros and cons.
5. Does RFID enable tracking of people by satellite? Do you think that someone could covertly read the contents in a shopping bag?
6. AIM Global has instituted an image called the RFID emblem that acts as a visual indicator to consumers and retail workers to help them find and identify the presence (and type) of RFID tag in a label, tag, or item. Some think that this is a sufficient self-regulation of the industry. What is your opinion? What other measures, if any, would you suggest in order to further protect consumers' privacy?
7. Besides consumer goods, RFID technology can be beneficial in other applications, such as preventing drug counterfeiting and reducing tampering, alerting staff to wandering patients in retirement homes, locating lost children in public places, and finding stray animals. Do you believe these positive uses can counter skepticism about the privacy invasion of RFID? Explain your answer.
8. Here are a few examples of systems/programs targeting human subjects in the United States that were implemented without the consent of the individual(s) involved:
   - DoD's mandatory vaccinations against the biological germ weapon called anthrax;
   - Several mandatory vaccination programs throughout U.S. history;

- President Franklin Roosevelt's signing of the Selective Training and Service Act of 1940, which created the country's first peacetime draft.

These examples stirred a highly emotional debate on ethics among many people and a few lessons learned can be gathered from these mandatory programs.

Speculation continues to grow as to which groups of the human population will have microchips introduced into their bodies to test the feasibility of the concept and for further future implementation. Some have alluded to the fact that eventual human microchip implantation is coming and is possible. These people believe that it will first be on a volunteer basis and then the government will intervene making it mandatory in the penal system and military and later for the general public as well. Discuss the topic in detail.

9. It has long been predicted that RFID and sensors would be combined, whereby a sensor gathers environmental information that is stored on an accompanying RFID tag. In February 2007, the Web site http://www.NewScientist.com discovered a recently filed patent application by Kodak that outlines a new application for RFID ingestible tags that act as monitors for health characteristics within the human body. The idea is that the RFID tag antenna could be composed of organic material that would dissolve as a result of certain chemical reactions within the human body. Once dissolved, the tag antenna, and therefore the tag itself, would stop transmitting a signal, indicating that the targeted chemical reaction had occurred. Kodak calls them *fragile tags.*

   For example, imagine an RFID reader–equipped drug dispenser installed in the home bathroom of a patient. The patient is prescribed to take a pill every day, which is issued by the dispenser. Once the pill is issued, the dispenser's RFID reader activates and begins polling for the signal of a tag, which is physically attached to the dispensed pill. In this uningested state, the tag functions properly, responding to the RFID reader's interrogation, which in turn informs the dispenser that the day's dosage has not yet been taken. Once the patient ingests the pill/tag, the organic tag antenna is subjected to chemicals within the patient's stomach. The tag antenna was designed to rapidly dissolve in the presence of normal stomach chemicals, so after only a few minutes it does so, and the tag ceases to respond to the RFID reader signal, which the dispenser interprets as the patient having taken his/her daily medication.

   The concept could be applied to changes in mechanical states as

well. Discuss an application in which a tag may be affixed to an artificial or natural body part and when wear on the body part, for example, an artificial hip, has proceeded to a predetermined level, the tag is rendered useless, thus alerting the remote query that the body part has achieved an unsatisfactory level of wear.

10. Could RFID ever lead to massive layoffs of workers?
11. Analyze and discuss the following statement, taken from [9]:

> The Smart Human Environment will completely change the way we interact with our environment and with each other. People will communicate with their technological environment naturally, using a variety of modalities and devices. The environment will be aware and will understand the user's social, physical and situational context. The environment will be able to smartly assist the user in his tasks, based on this context awareness and knowledge of the user's behavioral profile as well as common sense knowledge. It will exhibit pro-active behavior for recurring tasks and provide personalized information services.

References

[1] Ohkubo, M., et al., "RFID Privacy Issues and Technical Challenges," *Communications of the ACM,* Vol. 48, No. 9, September 2005.

[2] Garfinkel, S., et al., *RFID Privacy: An Overview of Problems and Proposed Solutions,* Los Alamitos, CA: IEEE Computer Society, 2005.

[3] Federal Trade Commission Report, *Radio Frequency Identification Applications and Implications for Consumers,* March 2005.

[4] "Possible Health Risks to the General Public from the Use of Security and Similar Devices," International Commission on Non-Ionizing Radiation Protection, 2002.

[5] Lindqvist, P., "RFID Monitoring of Healthcare Routines and Processes in Hospital Environment," M.S. Thesis, Department of Electrical and Communications Engineering, Master's Program in Bioinformation Technology, Helsinki University of Technology, Helsinki, Finland, August 10, 2006.

[6] Satava, R. M., "Biomedical, Ethical, and Moral Issues Being Forced by Advanced Medical Technologies," *Proc. American Philosophical Soc.,* Vol. 147, No. 3, September 2003.

[7] Naam, R., *More Than Human,* 1st ed., New York: Broadway Books, 2005.

[8] NanoMarkets LC, "Printable Electronics: Roadmaps, Markets and Opportunities," Executive Summary, September 2005.

[9] Sipilä, M., (ed.), "Communications Technologies, The VTT Roadmaps," Helsinki, Finland: VTT Research Notes 2146, *ESPOO 2002,* 2002.

Appendix

Common Units Conversion

$$1\text{m} = 1.09361 \text{ yards} = 3.28084 \text{ feet} = 0.001 \text{ km} = 6.21371 \times 10^{-4} \text{ miles} = 39.3701 \text{ inches}$$

$$1 \text{ m}^2 = 1{,}550 \text{ inches}^2 = 10.7639 \text{ feet}^2 = 1.19599 \text{ yards}^2$$

$$1 \text{ m}^3 = 61{,}023.7 \text{ inches}^3 = 1.30795 \text{ yards}^3 = 35.3147 \text{ feet}^3$$

$$1 \text{ km/hour} = 0.277778 \text{ m/second} = 0.621371 \text{ mi/hour} = 3.28084 \text{ feet/second}$$

$$1 \text{ kWh} = 3.6 \times 106\text{J} = 859.845 \text{ kcal} = 3412.14 \text{ Btu}$$

$$1\text{W} = 0.238846 \text{ cal/second} = 3.41214 \text{ Btu/hour}$$

$$1 \text{ kcal} = 4186.8\text{J} = 0.745700 \text{ kWh}$$

dBm

Because bel and decibel are ratios between two power values, such as input and output power, another measure is needed to express power in terms of a fixed reference point. This measure is called dBm:

$$\text{Power [dBm]} = 10 \log_{10} (\text{power [mW]}/1 \text{ mW})$$

- dBm uses 1 mW or milliwatt as the standard or reference point;
- 1 mW = 0 dBm;
- 10 dBm means a signal that is 10 dB above 1 mW.

dBW

$$\text{Power [dBW]} = 10 \log_{10} (\text{power [W]}/1\text{W})$$

- dBW uses 1W as the standard or reference point;
- dBW = dBm −30 dB.

Free-Space Loss

Free-space path loss is signal energy lost in traversing a path in free space only, with no other obstructions or propagation issues:

$$\text{FSPL [dB]} = 96.6 + 20 \log_{10} d\,(\text{miles}) + 20 \log_{10} f\ [\text{GHz}]$$

$$\text{FSPL [dB]} = 92.4 + 20 \log_{10} d\,(\text{km}) + 20 \log_{10} f [\text{GHz}]$$

Glossary of Terms

Access control An RFID application in which RFID tag-equipped badges are used to provide secure access to a facility.

Active tag A type of RFID tag that contains an internal power source, and in some cases also a radio transceiver. These additional component(s) are used to enhance the effective read/write range and rate of data transfer characteristics of the RFID tag.

ADC Automated data collection.

AIDC Automatic identification and data collection.

AIM Automatic Identification Manufacturers.

AIM Global The worldwide trade association of components and systems providers for automatic identification, data collection, and data integration in management information systems. Its members are manufacturers or service providers of identification technologies such as RFID, barcode, smart card, and biometrics.

Amplitude modulation (AM) Representation of data or signal states by the amplitude of a fixed-frequency sinusoidal carrier wave. Where data is in binary form, the modulation involves two levels of amplitude and is referred to as amplitude shift keying (ASK).

Amplitude shift keying (ASK) Representation of binary data states, 0 and 1, by the amplitude of a fixed-frequency sinusoidal carrier wave. Where the amplitudes are determined by the carrier being switched on and off, the process is known as *on-off keying* (OOK).

Antenna A conductive structure specifically designed to couple or radiate electromagnetic energy. In a driven mode the structure is a transmitter antenna. In receiver mode the structure is a receiver antenna. Antenna structures, often encountered in RFID systems, can be used to both transmit and receive electromagnetic energy, particularly data-modulated electromagnetic energy.

Antenna gain The measure of the amount of signal the antenna radiates or receives. It is given as a decibel ratio, compared to a theoretical omnidirectional antenna called an isotropic antenna. All other things being equal, a high-gain antenna will transmit and receive weaker signals farther than a low-gain antenna. Omnidirectional antennas, such as dipole antennas, will have lower gain than directional antennas because they distribute their power over a wider area.

Anticollision (anticontention) A term describing a facility for avoiding contention at the reader/interrogator receiver for responses arising from transponders simultaneously present within the read or interrogation zone of an RFID system and competing for attention at the same time.

Anticollision capability An RFID technology characteristic that allows for multiple RFID tags to be identified while present in an RF portal.

Application identifier (AI) A metadata element used to define the meaning of the data that follows.

Auto-ID Labs Lab that is currently headquartered at the Massachusetts Institute of Technology (MIT) in Boston, United States, and further based at six other leading universities worldwide: University of Cambridge, United Kingdom; University of Adelaide, Australia; Keio University, Tokyo, Japan; Fudan University, Shanghai, China; University of St. Gallen, Switzerland; and Information and Communications University (ICU) in Daejeon, Republic of Korea.

Backscatter modulation A process whereby a transponder responds to a reader/interrogation signal or field by modulating and reradiating or transmitting the response signal at the same carrier frequency.

BAPs Battery-assisted passive tags.

Batch reading The process or capability of an RFID reader/interrogator to read a number of transponders present within the system's interrogation zone at the same time. Alternative term for *multiple reading.*

Bandwidth The range or band of frequencies, defined within the electromagnetic spectrum, that a system is capable of receiving or delivering.

Baud A unit of signaling or transmission speed representing the number of signaling events per unit time. When the signal event is a single-bit, binary state representation, the baud is equivalent to the bit rate, expressed in bits per second (bps).

Biometrics The biological identification of a person, including characteristics of structure and of action such as iris and retinal patterns, hand geometry, fingerprints, voice responses to challenges, and the dynamics of handwritten signatures. Biometrics are a more secure form of authentication than using cards or typing passwords; however, some forms have relatively high failure rates. Biometric authentication is often a secondary mechanism in two-factor authentication. (Two-factor authentication is a system wherein two different methods are used to authenticate; using two factors as opposed to one implies a higher level of security.)

Bluetooth A radio technology developed by Ericsson and other companies built around a new chip that makes it possible to transmit signals over short distances between phones, computers, and other devices without the use of wires. More information is available at http://www.bluetooth.com.

Broadband A classification of the information capacity or bandwidth of a communication channel. Broadband is generally taken to mean a bandwidth higher than 2 Mbps.

Capture field, area, or zone Also *interrogation zone, area,* or *volume*. The region of the electromagnetic field, determined by the reader/interrogator antenna, in which the transponders are signaled to deliver a response.

CENELEC European Committee for Electrotechnical Standardization.

CEPT European Conference of Postal and Telecommunications Administrations.

Collision Term that denotes an event in which two or more data communication sources compete for attention at the same time and cause a clash of data,

inseparable without some means of anticollision or contention management. Also, see *Contention.*

Contention (clash) Term denoting simultaneous transponder responses capable of causing potential confusion, and misreading within a reader/interrogator system that is not equipped with anticontention facilities. See *Collision.*

Dense reader mode When the number of readers operating is large compared to the number of available channels, this is defined as a dense reader environment, for example, 20 readers operating on 20 available channels.

Detuning The change in the performance of transponders and readers caused by the presence of metal or ferromagnetic materials.

EAN European Article Numbering System, the international standard barcode for retail food packages.

EAS Electronic article surveillance. EAS is an RFID application in which the exit of items out of some physical environment is monitored electronically. Typically used for theft and loss prevention in the retail industry.

EC European Commission.

EIA Electronic Industries Association. EIA specifies electrical transmission standards, including those used in networking.

Electronic product code (EPC) The EPC numbering system uniquely identifies objects and facilitates tracking throughout a product's life cycle.

EMC Electromagnetic compatibility.

ETSI European Telecommunications Standards Institute. A body formed by the European Commission in 1988 that includes vendors and operators. ETSI's purpose is to define standards that will enable the European market for telecommunications to function as a single market.

EPCglobal, Inc. EPCGlobal, Inc., is a nonprofit corporation jointly established in September 2003 by European Article Numbers International, a European distribution standardizing organization, and the Uniform Code Council, Inc., an organization responsible for the distribution and management of barcodes. EPCglobal is leading the development of industry-driven standards for the electronic product code to support the use of RFID.

FCC Federal Communications Commission. The U.S. government agency responsible for the allocation of radio spectrum for communication services. Regulates interstate communications and licenses, rates, tariffs, standards, and limitations. In Canada, the same function is conducted by Industry Canada.

Field strength The strength of the electromagnetic signal at a specified distance from the transmitting antenna. The legal field strength limits vary with country. Units of measurement include milliamps per meter (mA/m), millivolts per meter (mV/m), decibel microvolts per meter (dBμV/m), decibel microamperes per meter (dBμA/m), microvolts per meter (μV/m), and microamps per meter (μA/m).

Frequency modulation (FM) Representation of data or signal states by using different transmission frequencies. Where data is in binary form, the modulation constitutes two transmission frequencies and is referred to as frequency shift keying (FSK).

Gen 2 Generation 2 standard from EPCglobal. It defines the coding and air interface for UHF tag operation.

Group selection Mode of operation whereby an interrogator can search for and identify unique tags within an RF portal or RF field of view.

International Standardization Organization (ISO) Global network of the national standards institutes of 150 countries, on the basis of one member per country. The organization has developed more than 13,000 international standards, including ISO9000, ISO14000, and ISO18000.

Interrogator Device that is used to read and/or write data to RFID tags; another name for an RFID reader.

ISO 14443 A four-part international standard for contact-less smart cards operating at 13.56 MHz in proximity to a reader antenna with a read range distance of up to 10 cm. The advantage products utilizing ISO 14443 have over those utilizing ISO 15693 is that the transaction speed is faster, making security and transaction speed superior for large packets of information, such as biometric templates. 14443A has grown to be the leading standard for access control and transportation and 14443B for banking.

ISO 15693 International standard regulating contactless, vicinity technology, typically representing a distance of more than 10 cm. The advantage ISO 15693

has over ISO 14443 is greater convenience due to longer read ranges and reduced power consumption.

Item management Those processes for the identification, tracking, and tracing of goods or items that are being manufactured, stored, transported, or discarded.

Label Sometimes called *inlay*; a tag that is thin and flexible.

LBT Listen-before-talk. Under European regulations, a reader must first listen on a particular channel for other signals before it is allowed to transmit; otherwise, it must select another channel. In addition, every 4 seconds the reader must release the channel for 0.1 seconds to allow other readers a chance to occupy that channel.

Manchester coding A biphase code format in which each bit in the source-encoded form is represented by 2 bits in the derived or channel encoded form. The transformation rule ascribes 01 to represent 0 and 10 to represent 1.

Manufacturer's Tag ID (MfrTagID) A reference number that uniquely identifies a tag.

Modulation The methods used to modulate or alter a signal between reader and transponder in order to carry the encoded information are quite varied. In some cases, the modulation can be different between the reader and the transponder and between the transponder and the reader. Some of the methods used are amplitude modulation (AM), phase modulation (PM), frequency modulation (FM), pulse-width modulation (PWM), and continuous-wave modulation (CW).

MRP Manufacturing resource planning.

Multitechnology Reader A reader utilizing two or more technologies, such as proximity (125 kHz) and contactless (13.56 MHz).

Near field Operating specification for an RFID tag to be near or in proximity to an interrogator's antenna. Near-field-capable interrogators and corresponding RFID tags typically have a read/write range of 4 to 6 inches.

NFC Near-field communication. A standards-based, short-range wireless connectivity technology that enables simple and safe two-way interactions among

electronic devices, developed in conjunction between Philips and Sony to compete with Bluetooth wireless communication.

Passive tag Type of RFID tag that does not contain an internal power source. This type of tag design is not complex and is usually of a single- or dual-chip design. It is said to be beam powered using the electromagnetic energy of an interrogator.

POS Point-of-sale.

Pulse oximetry Simple, noninvasive method of monitoring the percentage of hemoglobin (Hb), which is saturated with oxygen. Pulse oximeters are now a standard part of perioperative monitoring; they give the operator a noninvasive indication of the patient's cardiorespiratory status.

PCD Proximity coupling device (reader/writer).

PICC Proximity integrated circuit card.

PUPI Pseudo unique PICC identifier.

Radiation resistance Portion of an antenna's impedance that results in power radiated into space (i.e., the effective resistance that is related to the power radiated by the antenna). Radiation resistance varies with antenna length. Resistance increases as the wavelength increases.

RF absorption Radio phenomenon that occurs when transmitted RF signal energy is consumed or rapidly dispersed by some material in the pathway of the RF transmission.

RF cancellation Radio phenomenon that occurs when a transmitted RF signal is neutralized by competing RF interference.

RF portal Defined physical area of RF signal saturation. Also known as an *RF depth of field* or a *physical RF field of view.*

RF reflection Radio phenomenon that occurs when a transmitted RF signal is echoed off of another RF radiator placed within the pathway of the RF transmission.

RFDC Radio-frequency data collection. An implementation of automated data collection whereby portable ADC reader devices are connected to a host computer via RF so that interactive data transfers can occur.

RFID Radio-frequency identification. A method of storing and retrieving data via electromagnetic transmission to a radio-frequency-compatible integrated circuit.

RFID carrier frequency Defined radio-frequency for transmitting and receiving data. RFID frequencies include 2.45 GHz, 915 MHz, 13.56 MHz, and 125 kHz.

RFID site survey Comprehensive analysis aimed at determining or confirming that a proposed RFID solution meets the intended application requirements and technology specifications of use. It also defines the equipment needed to implement a proposed RFID system, and outlines the responsibilities of each party involved with the system implementation.

RS232 A common physical interface standard specified by the EIA for the interconnection of devices. The standard allows for a single device to be connected (point-to-point) at baud rates up to 9,600 bps, at distances up to 15m. More recent implementations of the standard may allow higher baud rates and greater distances.

RS422 Balanced interface standard similar to RS232, but using differential voltages across twisted-pair cables. More noise immune than RS232 and can be used to connect single or multiple devices to a master unit, at distances up to 3,000m.

RS485 Enhanced version of RS422 that permits multiple devices (commonly 32) to be attached to a two-wire bus at distances of more than 1 km (close to a mile).

Savant Servers that act as local repositories for data and associated information, supporting sophisticated, flexible middleware for serving database/XML queries.

SAW Surface acoustic wave. Technology used for automatic identification in which low-power microwave RF signals are converted to ultrasonic acoustic signals by a piezoelectric crystalline material in the transponder. Variations in phase shift in the reflected signal can be used to provide a unique identity.

Smart label Passive RFID data carrier structured into a flexible label-like form that allows overprinting with text, graphics, or data carrier symbols, such as linear barcodes, multirow barcodes, or matrix code symbols. Alternatively, it is used to describe a passive chip or chipless RFID devices that are used to emulate and extend a printed label function.

Spurious emissions Unwanted harmonic outputs. Type approval testing includes measurement of the harmonics of the reader to ensure that they are within required limits.

Supply chain Grouping of at least four distinct management business processes that define the planning of a product, the sourcing of a product's components, the making of a product, and the delivery of a product.

Synchronization A mechanism that allows multiple readers to operate in proximity by synchronization of their transmissions.

Transmitter (exciter) The electronics that drive an antenna. Together with the antenna and a receiver, they are called a reader or scanner.

Transponder Type of integrated circuit designed to store data and respond to RF transmissions of a given frequency. A transponder is another name for an RFID tag.

Transportation management Term used to reference several RFID applications within the transportation industry. These include electronic toll and traffic management, rail and intermodal tracking, fleet management, and vehicle parking/security access control.

UID Unique identifier. Number that uniquely identifies a transponder; is used for addressing each transponder individually.

U-NII Unlicensed National Information Infrastructure. The 5-GHz microwave band that does not require licensing (at least, not in the United States).

Wiegand format The most common data format in an access control system consisting of 26 bits of information

Write The transfer of data to a tag. The tag's internal operation may include reading the data in order to verify the operation.

Write broadcast capability RFID technology characteristic that allows data to be written to multiple tags while those tags are within an RF portal.

Write-once, read-many (WORM) An RF transponder that can be partly or totally programmed once by the user and thereafter only read.

Write rate The rate at which data can be transferred to a tag, written into the tag's memory and verified as correct. It is measured in bits (or bytes) per second.

WSN Wireless sensor networks.

About the Author

Harvey Lehpamer completed his primary education and technical high school education with a specialization in electronics in Zagreb, Croatia. He graduated from the School of Electrical Engineering (Electrotechnical Faculty) of the University of Zagreb, Croatia, Radiocommunications and Professional Electronics Department and received his master's degree from the same institution.

Mr. Lehpamer has more than 20 years of experience in the planning, design, and deployment of wireless and wireline networks, including microwave, fiber-optic, and other transmission (transport) systems in Europe, Canada, the United States, Africa, Mexico, and other parts of the world. He also has experience in project management, proposals and sales support, teaching, conducting seminars, electronic circuit design, and manufacturing and testing, and he is also author of the books *Transmission Systems Design Handbook for Wireless Networks* (Artech House, 2002) and *Microwave Transmission Systems: Planning, Design, and Deployment* (McGraw-Hill, 2004).

Mr. Lehpamer has worked for Ericsson Wireless Communications, Inc., San Diego, California; Qualcomm, Inc., San Diego; Clearnet, Inc., Toronto, Canada; Ontario Hydro, Canada; Lucas Aerospace, Inc., Microwave Technologies Division, Canada; and Electroproject, Zagreb, Croatia; and others.

He is a licensed professional engineer in the Province of Ontario, Canada. Today, he is the owner and principal engineer of HL Telecom Consulting, a consulting company in San Diego, California, (http://www.HLTelecomConsulting.com), and can be contacted at hl_2@hotmail.com or HarveyLehpamer@HLTelecomConsulting.com.

Index

Recent Titles in the Artech House Microwave Library

Active Filters for Integrated-Circuit Applications, Fred H. Irons

Advanced Techniques in RF Power Amplifier Design, Steve C. Cripps

Automated Smith Chart, Version 4.0: Software and User's Manual, Leonard M. Schwab

Behavioral Modeling of Nonlinear RF and Microwave Devices, Thomas R. Turlington

Broadband Microwave Amplifiers, Bal S. Virdee, Avtar S. Virdee, and Ben Y. Banyamin

CMOS RFIC Design Principles, Robert Caverly

Computer-Aided Analysis of Nonlinear Microwave Circuits, Paulo J. C. Rodrigues

Design of FET Frequency Multipliers and Harmonic Oscillators, Edmar Camargo

Design of Linear RF Outphasing Power Amplifiers, Xuejun Zhang, Lawrence E. Larson, and Peter M. Asbeck

Design of RF and Microwave Amplifiers and Oscillators, Pieter L. D. Abrie

Digital Filter Design Solutions, Jolyon M. De Freitas

Distortion in RF Power Amplifiers, Joel Vuolevi and Timo Rahkonen

EMPLAN: Electromagnetic Analysis of Printed Structures in Planarly Layered Media, Software and User's Manual, Noyan Kinayman and M. I. Aksun

Essentials of RF and Microwave Grounding, Eric Holzman

FAST: Fast Amplifier Synthesis Tool—Software and User's Guide, Dale D. Henkes

Feedforward Linear Power Amplifiers, Nick Pothecary

Foundations of Oscillator Circuit Design, Guillermo Gonzalez

Fundamentals of Nonlinear Behavioral Modeling for RF and Microwave Design, John Wood and David E. Root, editors

Generalized Filter Design by Computer Optimization, Djuradj Budimir

High-Linearity RF Amplifier Design, Peter B. Kenington

High-Speed Circuit Board Signal Integrity, Stephen C. Thierauf

Intermodulation Distortion in Microwave and Wireless Circuits, José Carlos Pedro and Nuno Borges Carvalho

Introduction to Modeling HBTs, Matthias Rudolph

Lumped Elements for RF and Microwave Circuits, Inder Bahl

Lumped Element Quadrature Hybrids, David Andrews

Microwave Circuit Modeling Using Electromagnetic Field Simulation, Daniel G. Swanson, Jr. and Wolfgang J. R. Hoefer

Microwave Component Mechanics, Harri Eskelinen and Pekka Eskelinen

Microwave Differential Circuit Design Using Mixed-Mode S-Parameters, William R. Eisenstadt, Robert Stengel, and Bruce M. Thompson

Microwave Engineers' Handbook, Two Volumes, Theodore Saad, editor

Microwave Filters, Impedance-Matching Networks, and Coupling Structures, George L. Matthaei, Leo Young, and E.M.T. Jones

Microwave Materials and Fabrication Techniques, Second Edition, Thomas S. Laverghetta

Microwave Mixers, Second Edition, Stephen A. Maas

Microwave Radio Transmission Design Guide, Trevor Manning

Microwaves and Wireless Simplified, Third Edition, Thomas S. Laverghetta

Modern Microwave Circuits, Noyan Kinayman and M. I. Aksun

Modern Microwave Measurements and Techniques, Second Edition, Thomas S. Laverghetta

Neural Networks for RF and Microwave Design, Q. J. Zhang and K. C. Gupta

Noise in Linear and Nonlinear Circuits, Stephen A. Maas

Nonlinear Microwave and RF Circuits, Second Edition, Stephen A. Maas

QMATCH: Lumped-Element Impedance Matching, Software and User's Guide, Pieter L. D. Abrie

Phase-Locked Loop Engineering Handbook for Integrated Circuits, Stanley Goldman

Practical Analog and Digital Filter Design, Les Thede

Practical Microstrip Design and Applications, Günter Kompa

Practical RF Circuit Design for Modern Wireless Systems, Volume I: Passive Circuits and Systems, Les Besser and Rowan Gilmore

Practical RF Circuit Design for Modern Wireless Systems, Volume II: Active Circuits and Systems, Rowan Gilmore and Les Besser

Production Testing of RF and System-on-a-Chip Devices for Wireless Communications, Keith B. Schaub and Joe Kelly

Radio Frequency Integrated Circuit Design, John Rogers and Calvin Plett

RF Design Guide: Systems, Circuits, and Equations, Peter Vizmuller

RF Measurements of Die and Packages, Scott A. Wartenberg

The RF and Microwave Circuit Design Handbook, Stephen A. Maas

RF and Microwave Coupled-Line Circuits, Second Edition, R. K. Mongia, I. J. Bahl, P. Bhartia, and J. Hong

RF and Microwave Oscillator Design, Michal Odyniec, editor

RF Power Amplifiers for Wireless Communications, Second Edition, Steve C. Cripps

RF Systems, Components, and Circuits Handbook, Ferril A. Losee

RFID Design Principles, Harvey Lehpamer

Stability Analysis of Nonlinear Microwave Circuits, Almudena Suárez and Raymond Quéré

System-in-Package RF Design and Applications, Michael P. Gaynor

TRAVIS 2.0: Transmission Line Visualization Software and User's Guide, Version 2.0, Robert G. Kaires and Barton T. Hickman

Understanding Microwave Heating Cavities, Tse V. Chow Ting Chan and Howard C. Reader

For further information on these and other Artech House titles, including previously considered out-of-print books now available through our In-Print-Forever® (IPF®) program, contact:

Artech House
685 Canton Street
Norwood, MA 02062
Phone: 781-769-9750
Fax: 781-769-6334
e-mail: artech@artechhouse.com

Artech House
46 Gillingham Street
London SW1V 1AH UK
Phone: +44 (0)20 7596-8750
Fax: +44 (0)20 7630 0166
e-mail: artech-uk@artechhouse.com

Find us on the World Wide Web at: www.artechhouse.com