Chapter 2

Physical Security

Contents

2.1 Physical Protections and Attacks

We live in a physical world. This is an obvious fact, of course, but it is surprisingly easy to overlook when discussing the security of digital information. Our natural tendency is to consider computer security strictly in a digital context, where computers are accessed only over a network or through a well-specified digital interface and are never accessed in person or with physical tools, like a hammer, screwdriver, or container of liquid nitrogen. Ultimately, however, digital information must reside somewhere physically, such as in the states of electrons, magnetic media, or optical devices, and accessing this information requires the use of an interface between the physical and digital worlds. Thus, the protection of digital information must include methods for physically protecting this interface.

Physical security is broadly defined as the use of physical measures to protect valuables, information, or access to restricted resources. In this chapter, we examine the physical dimensions of computer security and information assurance, focusing on the following aspects:

- ***Location protection***: the protection of the physical location where computer hardware resides, such as through the use of locks.
- ***Physical intrusion detection***: the detection of unauthorized access to the physical location where computer hardware resides.
- ***Hardware attacks***: methods that physically attack the hardware representations of information or computations, such as hard drives, network adapters, memory chips, and microprocessors.
- ***Eavesdropping***: attacks that monitor light, sound, radio, or other signals to detect communications or computations.
- ***Physical interface attacks***: attacks that penetrate a system's security by exploiting a weakness in its physical interface.

We discuss these physical aspects of computer security and information assurance and we give several examples of vulnerabilities in the physical aspects of some security solutions, including smart cards, automated teller machines (ATMs), radio-frequency identification (RFID) tags, biometric readers, and voting machines. An important theme that runs throughout this discussion is the way in which physical security directly impacts the integrity and protection of computer hardware and digital information.

2.2 Locks and Safes

The notion of using a mechanical locking device to protect access to a building, vehicle, or container has been in use since ancient times. Primitive tumbler locks, which are discussed below, have been found in the ruins of ancient Egyptian and Persian cities. Today, a wide variety of locks are in common usage, including those that require a key, a combination, or both, and these mechanisms are often used to protect the physical locations where computers and digital media are stored. This section covers commonly used lock types and techniques for attacking them without having the corresponding key or combination.

2.2.1 Lock Technology

Pin Tumbler Locks

The most commonly used type of keyed lock is the ***pin tumbler lock***, illustrated in Figure 2.1. In this design, a cylindrical ***plug*** is housed within an outer casing. The lock is opened when the plug rotates and releases a locking bolt, typically through a lever. When the lock is closed, the rotation of the plug is prevented by a series of ***pin stacks***, which are housed in holes that have been drilled vertically through the plug and the outer casing. A pin stack typically consists of two cylindrical pins. The top pins, called ***driver pins***, are spring loaded.

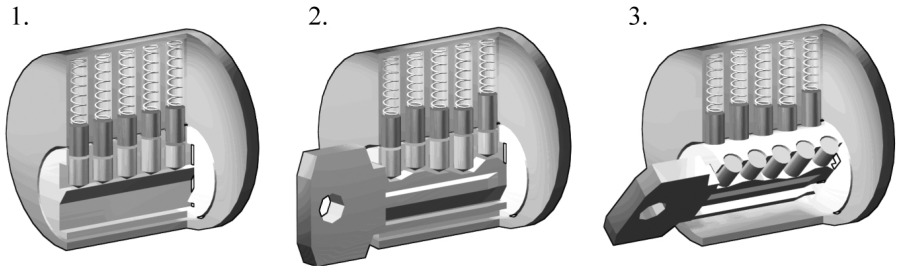


Figure 2.1: A pin tumbler lock: (1) When a key is not present, the pin stacks are pushed down by the springs so that the driver (top) pins span the plug and the outer casing, preventing the plug from rotating. Image included with permission [108]. (2) When the correct key is inserted, the ridges of the key push up the pin stacks so that the cuts of the pin stacks are aligned with the shear line. Image included with permission [75]. (3) The alignment of the cuts with the shear line allows the plug to be rotated. Image included with permission [76].

The bottom pins are called the ***key pins***, since they make contact with the key when the key is inserted. The heights of the respective driver and key pins can vary. When there is no key inserted, the springs force the pin stacks down so that the driver pins span the plug and the outer casing, preventing the plug from rotating. When the appropriate key is inserted, however, the ridges of the key push up the pin stacks so that the cut between each key pin and its driver pin is at the point where the plug meets the outer casing, known as the ***shear line***, allowing the plug to rotate and open the lock.

Tubular and Radial Locks

A variation on the classic pin tumbler design is known as the ***tubular lock***, or ***radial lock***, depicted in Figure 2.2. The premise is the same: several spring-loaded pin stacks prevent the rotation of the plug by obstructing the shear line. Rather than having the pins located on a line parallel to the axis of the plug, as in the traditional pin tumbler lock, the pins of a tubular lock are arranged in circle. As a result, keys are cylindrical in shape. These locks are most commonly used on laptops, vending machines, and bicycles.

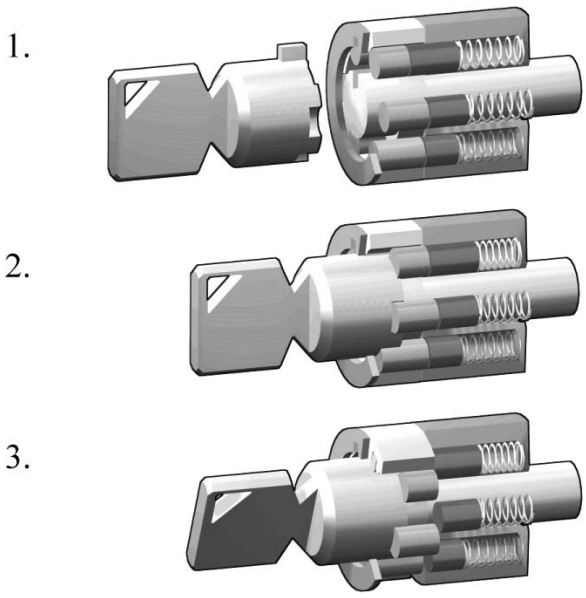


Figure 2.2: Opening a tubular lock: (1) Closed lock. Image included with permission [68]. (2) After inserting the key. Image included with permission [69]. (3) Open lock. Image included with permission [70].

Wafer Tumbler Locks

A third type of lock in common usage is known as the ***wafer tumbler lock***, depicted in Figure 2.3. Again, the general principle of the lock relies on preventing the rotation of a central plug. This time, the obstruction is a series of wafers that initially sit in a groove at the bottom of the outer casing. When the appropriate key is inserted, the wafers are raised out of this groove and allow the plug to rotate. Wafer tumbler locks are used in cars, filing cabinets, and other medium security applications.

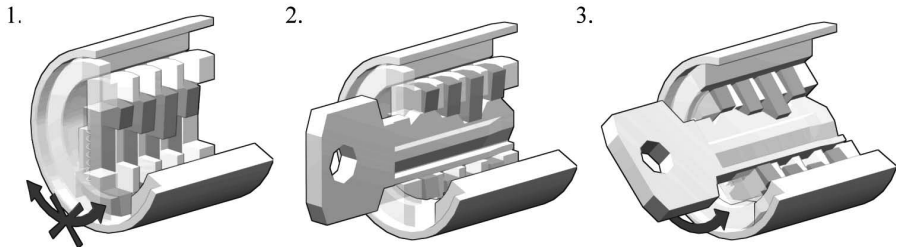


Figure 2.3: Opening a wafer tumbler lock: (1) Closed lock. Image included with permission [71]. (2) After inserting the key. Image included with permission [72]. (3) Open lock. Image included with permission [73].

Combination Locks

A ***combination lock*** is any lock that can be opened by entering a predetermined sequence of numbers. Combination locks typically come in one of three varieties, multiple dial, single dial, and electronic. Multiple-dial locks feature a sequence of notched disks around a toothed pin, as depicted in Figure 2.4.

Figure 2.4: Opening a multiple-dial combination lock. Image included with permission [74].

When the disks are rotated to the correct combination, the notches line up with the teeth of the pin, allowing it to be removed. Multiple-dial combination locks are often used in briefcases, bicycle locks, and other low-security applications, since it is often easy to quickly deduce the combination because of mechanical imperfections. Single-dial combination locks are generally considered more secure and are used in a wide variety of applications, including safes, which are discussed later in this section. Single-dial locks feature a series of disks attached to a numbered dial. When the correct combination is entered using the dial, these disks line up in such a way as to release a clasp or some other locking mechanism.

In an ***electronic combination lock***, an electronic mechanism is used to operate the lock using electromagnets or motors that are activated through an event that either turns on or turns off an electric current. The event that opens an electronic lock could be triggered by a number of different actions, including the following (which could be used in conjunction):

- An ***electronic combination***: the punching of an appropriate sequence of numbers on a keypad in a given amount of time
- A ***magnetic stripe card***: a plastic card with a magnetic stripe (Section 2.3.2) that contains an authorizing digital combination
- A ***smart card***: a small computational device contained in a card, as discussed in Section 2.3.3, that performs an authorizing computation to open the lock
- An ***RFID*** tag: a small radio frequency identification device that contains a computational element or memory, as discussed in Section 2.3.4, that either performs an authorizing computation or transmits an electronic combination
- A ***biometric***: a biological characteristic, as discussed in Section 2.3.5, that is read and matches a characteristic authorized to open the lock

One advantage of electronic locks is that it is relatively easy to change the combination or condition that opens such a lock—there is no need to change a physical plug or swap out pins. For instance, most hotels employ electronic lock systems for their guest rooms, allowing for easy changing of locks between consecutive guests staying in the same room.

Another advantage is that electronic locks can be fitted with digital storage devices or can be connected to a communication network to monitor and manage when the locks are opened and closed. The monitoring can even log who has entered and left through the doors that are fitted with the various locks in a building, by using different digital combinations or opening devices for different people. This type of monitoring was useful, for example, in determining who murdered a Yale graduate student in 2009. The monitoring system showed that the student had entered a secured

building, but never left, and it also allowed authorities to determine all the people who had entered the specific room where her body was found. Electronic locks are also frequently used in audit trails for regulatory compliance, especially in the health care and financial sectors.

Master and Control Keys

Many organizations require a key system that incorporates a hierarchy of access control. For example, some systems feature locks that have keys specific to each lock, known as ***change keys***, as well as a single ***master key*** that can open all of the locks in the system. Larger and more complex organizations may have several different master-keyed systems, with a single ***grandmaster key*** that can open any lock in the organization. Some locks also accept a ***control key***, which enables a locksmith to remove the entire core of the lock from its outer casing, allowing easy rekeying.

Locks designed to be opened by a master key have at least two keyings, one for the change key and one for the master key. Multiple keyings are created by inserting ***spacers***, or very short pins, between the driver and key pins. The height of the master key should be greater than that of the change key to prevent the owner of a change key from filing down their key to create a master key.

Master-keyed systems require the owner to incorporate access control policies and procedures for when a key is lost or stolen. If a master key is lost, it is necessary to rekey the entire system to prevent compromise. Handling the loss of a change key is left to the discretion of the organization, however. Some choose to merely rekey the specific lock that accompanies the lost key, while others rekey the entire system to ensure that the missing key does not allow an attacker to create a master key.

Safes

Valuables can be secured against theft by placing them in a ***safe***. Safes can range from small lockboxes in homes to large, high-security safes in banks. Safes can feature any of the locking mechanisms discussed in this chapter, but most high-security models employ a combination dial, with the possible addition of biometric authentication and electronic auditing.

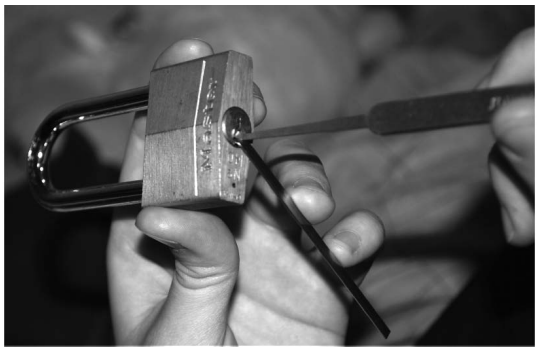
No safe is impenetrable. In fact, safes are rated by organizations such as ***Underwriters Laboratories*** (***UL***) in terms of how long it would take a well-equipped expert to compromise the safe. In their ratings, they consider both destructive and nondestructive approaches. Owners of safes are advised to ensure that the response time to alarms is less than the average time required to crack the safe.

2.2.2 Attacks on Locks and Safes

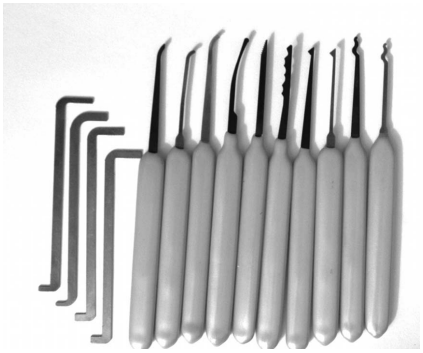
There are several ways to attack locks and safes, so as to open them without use of a key or a priori knowledge of a combination.

Lockpicking

The classic approach to bypassing locks is known as ***lockpicking***, which exploits mechanical imperfections in the locking mechanism that allow an attacker to replicate the effect of an authorized entry. (See Figure 2.5.)



(a)



(b)

Figure 2.5: Lockpicking: (a) A lock picker attempts to open a padlock by applying a rotational force with a tension wrench and picking the pins individually. Photo by Dan Rosenberg included with permission. (b) Lockpicking tools. Photo by Jennie Rogers included with permission.

As a simple example, let us examine the common pin tumbler lock. Recall from Section 2.2.1 that this type of lock features a cylindrical plug whose rotation is prevented by pins obstructing the shear line (where the plug meets the outer casing). To pick a tumbler lock using a common technique, an attacker first inserts a ***tension wrench*** into the keyhole of the lock and applies a slight rotational force. The plug will rotate a very small amount before being stopped by the pins. In particular, one of the pins will be directly in contact with the cylinder—this is known to lock pickers as a ***binding pin***. Only one pin comes in contact with the cylinder because of slight imperfections in the manufacturing process causing the pin stacks to not line up perfectly. Using a ***feeler pick***, the attacker first probes each pin, lifting it slightly to assess the amount of friction experienced. The binding pin will offer greater resistance to motion due to it making contact with the cylinder. The attacker then carefully raises the binding pin until the break

between the key pin and the driver pin is in line with the shear line. At this point, the plug will further rotate by a tiny amount until it is stopped by the next binding pin. The driver pin of the previous binding pin will now sit on a small ledge created by the additional rotation of the plug (it is "bound"). The attacker then repeats the process for the remaining pins, identifying and lifting them. When the last pin is picked, all of the pin breaks are aligned with the shear line, and the rotational force applied by the tension wrench will open the lock.

It takes a high degree of skill and intuition to determine when a pin has bound. Lock pickers must practice extensively to recognize the feeling of a pin binding in a locking mechanism. To make the process easier, lock pickers often employ a method known as ***raking*** or ***scrubbing***. In this technique, a pick designed to lift several pins simultaneously is run through the keyhole using a back and forth scrubbing motion in an attempt to bind more than one pin at the same time. Once several pins are bound, the remaining pins can be individually picked. Alternatively the attacker can make another pass with the rake to attempt to set more pins. Pickers typically use a snake or half-diamond rake for this purpose.

Another approach that can work on inexpensive locks is the use of a ***comb pick***. For simpler locks an attacker can lift all of the pins simultaneously above the shear line using a tool that resembles a hair comb. Once the pins have been pushed all the way into the lock housing then the plug is free to rotate. Well-made models combat this weakness by making sure that the pin stack is long enough to always extend past the shear line.

Lock Bumping

Lock bumping is a technique that received widespread media attention in 2006. The technique utilizes specially crafted ***bump keys***, which are unique to each particular brand of lock. (See Figure 2.6.)

Figure 2.6: Bump keys and hammer. Photo by Jennie Rogers included with permission.

Bump keys are made by taking an appropriate key blank (matching a specific lock brand) and filing down each of the pin ridges to the lowest setting, keeping the teeth in between each ridge. To "bump" a lock, the bump key is inserted into the keyhole, and then withdrawn a small amount so that each tooth rests immediately behind a pin. While applying a slight rotational force, the bump key is then reinserted by tapping it with a hammer or other object. This results in the ridges hitting the pin stacks. As a consequence, the key pins transfer kinetic energy to the driver pins, which jump above the shear line for a split second, allowing the plug to be rotated. Interestingly, more expensive locks are often more vulnerable to bumping because a reduction in mechanical imperfections allows the pins to move more freely when being bumped.

Professional locksmiths and law enforcement agents often employ the use of an electronic ***pick gun***, which operates on the same principle as lock bumping. A pick gun has a single pick that is vibrated rapidly and transfers energy to all of the pins simultaneously. During the split second after this energy transfer, which attempts to force the driver pins above the shear line, the pick gun applies a brief rotational force to attempt to open the lock.

Key Duplication and Impressioning

Several methods can be used to create a key for a given lock. A locksmith can easily create a key duplicate if the original is available, for instance. It is not always even necessary to have an original on-hand, however—a mere photograph of the key can be used if the locksmith is able to deduce the type of key and approximate its cut. Another technique used sometimes to "capture" a key consists of briefly taking it and pressing it into a soft, clay-like material that can later be hardened into a mold for key creation. A key does not need to be made of metal to be effective.

Another technique that is used to bypass locks is known as ***key impressioning***. An attacker begins with a key blank matched to a specific lock brand. The top of the blank is polished, and then the key is inserted into the target lock. A rotational force is applied to the key, which is then jiggled up and down. The blank is finally removed. Each pin that is not at the shear line will have left a small scrape on the polished blank. At each scraped location, the attacker files off a small amount of material. The process is repeated until no scrapes are created, resulting in a key that can open the lock. Key impressioning requires the use of keys blanks made of a soft metal, such as brass. Also, the attacker must use a very precise filing tool, such as a jeweler's file.

High Security Locks

A number of innovations have been developed to make bypassing locks more difficult.

One preventative measure is the incorporation of ***security pins***, such as ***mushroom head pins*** or ***spool pins***. In this design, the pins are narrower towards the middle, but flare outwards at the top and bottom. This design does not prevent normal operation of the lock, since a proper key moves the entire pin above the shear line. This technique makes picking more difficult, however, because the pin may bind in its midsection, preventing an attacker from binding the pin properly and from knowing whether or not the pin has bound in the correct place.

Another form of security pin is a serrated pin. This pin has a series of small ridges around it that make the pin feel like it is perpetually setting on the shear line to a lock picker working one pin at a time. Thus each time an attacker lifts a pin it gives the feeling that the cylinder is rotating slightly, despite it only moving up the ridges. The top and bottom pins may both have the ridges to further mislead an unauthorized user.

Security pins may defend against ordinary picking, but do little to stop techniques such as bumping. For this reason, lock manufacturers have developed high-security models that do not rely on the traditional pin tumbler design. Medeco developed a lock called the Biaxial that uses ***angular bitting***, which requires that each of the pins must be elevated and rotated by the angled cuts of the key.

As another example, Abloy manufactures a ***disc tumbler lock*** that utilizes a series of notched disks. This unique design makes traditional picking and bumping approaches impossible, but this lock may be vulnerable to other means of circumvention.

Higher security locks (including Medeco Biaxial and its variants) feature an internal sidebar, which prevents the cylinder from turning until all of the pins have been rotated and aligned, making picking extremely difficult. The lock is marketed as being bump-proof, but recent research suggests that highly specialized bump keys may still make bumping possible.

In addition, high security locks tend to be manufactured to tighter specifications, making it more challenging for a lock picker to identify the binding pins and feel out a lock. Also, to buy more time against a destructive attack, most higher security models also feature drill-resistant pins, which prevent an attacker from being able to use an off-the-shelf drill to attack the shear line of a lock.

High-value targets, such as nuclear facilities and banks, naturally require more security precautions for their locks. Typically, these requirements are mandated by either insurance underwriters or the government. There are two main standards for locks commonly used in the United States: Underwriters Laboratories (UL) 437 and ANSI/Building and Hardware Manufacturers Association (BHMA) 156.30.

These standards attempt to model how well a product will stand up to well-known attacks, including destructive and non-destructive approaches. In the case of UL, this means that they evaluate whether a lock can withstand picking, impressioning, drilling, pulling, driving, and salt spray corrosion. Each of these tests consists of a specified number of minutes that the lock must withstand the attack, where the picking and impressioning must be performed by a certified locksmith with at least five years of experience.

The reader my have noticed that bumping is not included in the list of UL tests. This attack was first widely published in 2006 and it can take many years to update standards for vulnerabilities as they are discovered. This is one of the weaknesses of the standards system. Criminals do not necessarily follow "well-known" methods of compromising locks and it behooves them not to share their techniques. A lock certified according to a standard may be still vulnerable to highly skilled attackers.

Compromising higher security locks often requires domain-specific knowledge and substantial research. A general specification may not encompass the necessary tests for all high security locks. For example, the attack by Tobias and Bluzmanis exploiting a vulnerability in the Medeco Biaxial system requires learning specialized codes to rotate the pins in the correct orientation.

The certification system also makes responsible disclosure for these locks considerably more complex. There is no common method to issue "patches" for locks in the field, nor retract a certification for a lock. Like many other aspects of security, high security lock management is a process that goes back and forth between security researchers and manufacturers.

Safe Cracking

There are many approaches to safe cracking. The classic approach, often portrayed in movies, involves a highly skilled expert manipulating the dial of the safe by feel and sound until they deduce the combination. To prevent this attack, many safe manufacturers include additional components inside the locking mechanism that are designed to prevent an attacker from correctly interpreting sounds and tactile clues. In addition, the wheels of the locking mechanism are often made with light materials such as nylon, reducing the noise and friction created by manipulating the lock.

An attacker may also attempt to drill through the front of the safe to see the lock's inner workings or to directly manipulate the lock mechanism. High-security safes incorporate composite ***hardplate***, which is an extremely rugged material designed to be resistant to drilling or other structural damage. Only highly specialized drilling equipment can be used to breach these materials. Brute-force techniques, such as explosives, may be employed, but often these approaches are impractical, because they risk damaging the contents of the safe. To further prevent drilling, many safes feature what is known as a ***glass relocker***, a thin glass plate that resides in the door of the safe. If this glass is broken, by drilling or some other force, spring-loaded bolts are released, permanently locking the safe.

Side Channel Attacks

Many of the principles observed in the design and circumvention of physical locks are analogous to essential principles of computer security. It is important to keep in mind that manipulating the mechanism of a lock is only one way to gain unauthorized access. For example, a door with a highly secure lock does little good if the door can be removed by unscrewing its hinges. Attacks such as these are referred to as ***side channel attacks***. (See Figure 2.7.)

Figure 2.7: A side channel attack vulnerability: the fire escape on the side of the building may lead to an entry point that is easier to attack than the front door. Photo by Jennie Rogers included with permission.

In a side channel attack, rather than attempting to directly bypass security measures, an attacker instead goes around them by exploiting other vulnerabilities not protected by the security mechanisms. Side channel attacks are sometimes surprisingly simple to perform.

A classic example of a side channel attack is door plunger manipulation. In doors that feature a plunger rather than a deadbolt, it may be possible to open the door by inserting a flat object in between the door and the doorframe (e.g., a slender screw driver) and manipulating the plunger until the door opens. This attack can be prevented by shielding the plunger or by using a deadbolt, but provides a good example of a situation where picking the locking mechanism may be difficult, but opening the door is still possible.

The concept of side channel attacks doesn't only apply to locks and safes. It can be applied to other aspects of computer security as well, since attackers often search for the simplest way of bypassing security, rather than the most obvious. Thus, the security of computer and information systems should be analyzed in a holistic way, looking at both physical and digital attacks, so as to identify the most vulnerable components in the system.

2.2.3 The Mathematics of Lock Security

The number of possible combinations or configurations for a set of objects, for which we are interested in finding one particular such object, is commonly known as a ***search space***. In computer security, the most common type of search space is the set of all possible keys used in a cryptographic function. A large search space reduces the possibility of a brute-force attack, where all possible combinations are tried. Therefore, anything that reduces the size of the search space would be extremely valuable for an attacker.

Protecting Against Brute-Force Attacks

The mathematics of search spaces also applies to lock security, of course. Traditional pin tumbler locks feature between 4 and 7 pin stacks, where the number of possible heights for the key pins is typically between 4 and 8. Higher quality locks have more pins stacks and a larger number of possible key pin heights. UL specifies that standard locks should have at least $1,000$ potential combinations, or ***differs***, and that security containers have $1,000,000$ or more differs. In addition, there are around 40 common varieties of key blanks. Collectively, this results in a search space where the

number of possible keys is no more than

$$40 \times 8^7 = 83{,}886{,}080.$$

If we know the specific key blank, the search space is much smaller. For example, for a given key blank with 6 pin stacks and 5 possible key pin heights, the number of possible keys is

$$5^6 = 15{,}625,$$

which is still large enough to prevent a brute-force attack. For this reason, attacks such as picking, key impressioning, and bumping are employed instead of brute-force techniques.

The mathematics of counting finite collections of objects is known as ***combinatorics***, by the way, and the above analysis is an example of a combinatorial argument that quantifies the security of a pin tumbler lock.

Reducing the Size of a Search Space

There are situations where effective use of combinatorics can allow for the bypassing of locks. For example, in standard lock picking, the attacker "solves" the lock one pin at a time, breaking the problem into two phases: finding the binding pin and then raising it slowly. If our attacker has P pin stacks and D possible pin heights this divide-and-conquer approach produces a search space of size $P \cdot D$ instead of P^D.

Privilege Escalation

Matt Blaze published a paper detailing how an attacker can use combinatorics to create a master key from an ordinary change key on a master-keyed lock system. This attack, which is an ***iterative master-key construction***, is analogous to ***privilege escalation*** in computer security, where an attacker leverages a low-privilege ability into a high-privilege ability.

The attack works as follows. Consider a lock having the same configuration of P and D as above. The attacker has a change key and wants to build a master key using a small number of keys blanks. The only basic operation used by the attacker is to test whether a given key opens the lock. For simplicity, we will use the term pin to denote a key pin.

Starting with the first pin stack, the attacker creates $D - 1$ keys, each keeping all pins but the first at the same height as in the change key, and trying all possible variations for the first pin height (except for the height of the first pin in the change key). The key that results in opening the lock reveals the height for the first pin of the master key. The process is then

repeated for each of the remaining pins, and a final master key is made using each pin's successful key. The above attack requires a total number of key blanks that is at most

$$P \cdot (D - 1).$$

Also, at most $P \cdot (D - 1)$ lock opening tests are performed.

In the case where a high-quality lock has 7 pin stacks and 8 possible key heights, this technique would require a maximum of 49 key blanks, which is within the realm of possibility. A lower quality lock with 5 pin stacks and 5 possible key heights would instead require no more than 20 key blanks. Alternatively, the attacker could file the test keys down on the fly, requiring only P key blanks.

Further Improvements

The search space in this case can be reduced even further, however, depending on the lock manufacturer. The ***maximum adjacent cut specification*** (MACS) of a lock defines the maximum vertical distance allowed between any two adjacent key cuts. If this distance is exceeded, the key will have a steep spike that will be breakable, cause the lock to jam, or prevent a pin from sitting properly. Removing all sequences that would violate the MACS of a particular lock from the search space results in a significant reduction in size. In addition, some master-keyed systems require that the master key pins are higher on the pin stack than the change keys, which further reduces the search space.

As another example of combinatorics at work, some brands of single-dial combination padlocks have mechanical dependencies as a result of the manufacturing process. As a result of these dependencies, it may be possible to drastically reduce the number of combinations for testing. On one brand, which has a dial ranging from 1 to 40 and requires a 3-number combination, it is possible to reduce the search space from 60,000 (40^3) to only 80, making a brute-force attack much more feasible.

It is important that single-dial combination locks have some sort of reset mechanism that is triggered whenever someone attempts to open that lock after trying a combination. If no such reset mechanism exists, the final digit of the combination is essentially rendered useless, since it requires a trivial amount of time to iterate through each final number, trying each one. This is an example of a measure that prevents a reduction of the search space.

2.3 Authentication Technologies

As mentioned in Chapter 1, the authentication of individuals can be derived from something they know, something they possess, and something they are. In this section, we discuss some physical means for achieving authentication through the use of something a person possesses or something they are (namely, a healthy human).

2.3.1 Barcodes

Printed labels called ***barcodes*** were developed around the middle of the 20th century as a way to improve efficiency in grocery checkout, and are now used universally in applications as diverse as identifying wildlife. First-generation barcodes represent data as a series of variable-width, vertical lines of ink, which is essentially a one-dimensional encoding scheme. (See Figure 2.8a.)

Some more recent barcodes are rendered as two-dimensional patterns using dots, squares, or other symbols that can be read by specialized optical scanners, which translate a specific type of barcode into its encoded information. Among the common uses of such barcodes are tracking postage, purchasing mass merchandise, and ticket confirmation for entertainment and sporting events. (See Figure 2.8b.)

(a) (b)

Figure 2.8: Examples of barcodes: (a) A one-dimensional barcode. (b) A two-dimensional barcode, which was used for postage.

Barcode Applications

Since 2005, the airline industry has been incorporating two-dimensional barcodes into boarding passes, which are created at flight check-in and scanned before boarding. In most cases, the barcode is encoded with an internal unique identifier that allows airport security to look up the corresponding passenger's record with that airline. Security staff then verifies that the boarding pass was in fact purchased in that person's name, and that the person can provide photo identification. The use of a private, external system prevents boarding passes from being forged, since it would require an additional security breach for an attacker to be able to assign an identifier to his or her own record with the airline.

In most other applications, however, barcodes provide convenience but not security. Since barcodes are simply ink on paper, they are extremely easy to duplicate. In addition, barcodes can be read from afar as long as the ink is within line of sight of the attacker. Finally, once a barcode is printed, it has no further ability to alter its encoded data. As a result, other mediums were developed that allowed writing data as well as reading it.

2.3.2 Magnetic Stripe Cards

Developed in the late 1960s, the ***magnetic stripe card*** is one of the most pervasive means of electronic access control. Currently, magnetic stripe cards are key components of many financial transactions, such as debit or credit card exchanges, and are the standard format for most forms of personal identification, including drivers' licenses. These cards are traditionally made of plastic and feature a stripe of magnetic tape contained in a plastic-like film. Most cards adhere to strict standards set by the ***International Organization for Standardization*** (***ISO***). These standards dictate the size of the card, the location of the stripe, and the data format of the information encoded into the stripe.

The magnetic stripe on standardized cards actually includes three tracks for storing information. The first track is encoded using a 7-bit scheme, featuring 6 bits of data and one ***parity*** bit per character, with a total of 79 characters. A parity bit is a bit whose value is a combinational function of the others, such as exclusive-or. Since magnetic stripes cards can potentially be worn down and subject to physical damage, the parity bit allows a stripe reader to read a card even if there is a small amount of data loss.

Magnetic Stripe Card Security

The first track of a magnetic stripe card contains the cardholder's full name in addition to an account number, format information, and other data at the discretion of the issuer. This first track is often used by airlines when securing reservations with a credit card.

The second track is encoded using a 5-bit scheme (4 bits of data and 1 parity bit per character), with a total of 40 characters. This track may contain the account number, expiration date, information about the issuing bank, data specifying the exact format of the track, and other discretionary data. It is most often used for financial transactions, such as credit card or debit card exchanges.

The third track is much less commonly used.

One vulnerability of the magnetic stripe medium is that it is easy to read and reproduce. Magnetic stripe readers can be purchased at relatively low cost, allowing attackers to read information off cards. When coupled with a magnetic stripe writer, which is only a little more expensive, an attacker can easily clone existing cards. Because of this risk, many card issuers embed a hologram into the card, which is harder to reproduce. Most credit cards also include space for a customer signature, verifying the authenticity of the card. Unfortunately, many vendors do not always check this signature. One effective deterrent against card fraud is a requirement for additional information known only to the owner, such as a ***personal identification number*** (***PIN***).

ISO standards do not permit vendors to store money on magnetic stripe cards. Account numbers can be stored instead, which can be used to access information in remote databases. Still, many organizations use cards that store contents of monetary value. For example, transportation tickets often store "money" that is only available for payment of transportation fees. So, vendors sometimes use proprietary technology that provides the convenience of storing data on a magnetic stripe in a format storing "points" or "credits" on the card that have monetary value.

Unfortunately, the use of a format that allows the cards to contain data that actually has a monetary value poses serious security risks. Because the money on the card is simply represented by data, attackers who know the format of the information on the stripe could create their own cards and provide themselves with free services. For this reason, it is essential that vendors protect the secrecy of their data format specifications and provide some means of validating data integrity, such as employing a cryptographic signature algorithm.

2.3.3 Smart Cards

Traditional magnetic stripe cards pose a number of security problems because they are relatively easy to duplicate and because there is no standardized mechanism for protecting the information contained on a card. Solutions to both of these problems are provided by ***smart cards***, which incorporate an integrated circuit, optionally with an on-board microprocessor. This microprocessor features reading and writing capabilities, allowing the data on the card to be both accessed and altered. Smart card technology can provide secure authentication mechanisms that protect the information of the owner and are extremely difficult to duplicate.

Smart cards do not suffer from the inherent weaknesses of the magnetic stripe medium. They are by design very difficult to physically disassemble, and an internal cryptoprocessor can provide data protection that simple stripes cannot. Most security problems in smart cards are a result of weaknesses in a specific implementation, not the basic technology itself.

First-generation smart cards require the integrated circuit to actually contact a reading device in order to access or alter information. This restricts the information on the card to those with physical access. A new generation of smart cards instead relies on radio frequency technology to allow contactless interaction of a smart card and a reading device. The introduction of this capability exposes smart cards to similar security risks as another popular technology, RFID, which is discussed in Section 2.3.4.

Smart Card Applications

Today, smart cards are used for a wide variety of applications. They are commonly employed by large companies and organizations as a means of strong authentication, often as a part of a single sign-on scheme. Some credit companies have begun embedding smart cards into their credit cards to provide more secure customer protection. In addition, many computer disk encryption technologies rely on smart cards for the storage of an external encryption key.

Smart cards may also be used as a sort of "electronic wallet," containing funds that can be used for a variety of services, including parking fees, public transport, and other small retail transactions. Current implementations of these types of smart cards provide no verification of ownership, so an attacker in possession of a stolen smart card can use it as if he were the owner. In all electronic cash systems where this is the case, however, the maximum amount of cash permitted on the card is low, to limit any possibility of serious fraud.

Smart Card Security

While most sophisticated smart cards feature a microprocessor that allows them to perform some computational work and alter the contents of the card, other less expensive versions simply contain a memory card, with no ability to alter the contents without an external writer. Many phone cards actually contain a monetary amount encoded on the card, for instance.

To prevent cloning and unauthorized alteration, most cards require that a secret authentication code be presented to a microcontroller reader/writer before memory can be written to a card. In addition, many phone cards authenticate themselves to the phone network using a unique serial number or PIN mechanism before any transfer of funds takes place, making them more difficult to clone.

Simple Attacks on Smart Cards

Unfortunately, if a phone card's secret code can be extracted, it may be possible to tamper with the monetary value on the card. Possible attacks include a social engineering approach (trying to recover the code from employees of the phone company) or eavesdropping on communications between a card and its reader.

Differential Power Analysis

In addition, even smart cards with secure cryptoprocessors can be subject to a side channel attack known as ***differential power analysis***. In this attack, a malicious party records the power consumption of a smart card's microprocessor while it is performing cryptographic operations. Because various operations on a processor require minutely different amounts of power consumption, it may be possible to statistically analyze this recorded information in order to reveal information about the cryptosystem or the underlying cryptographic key that is being processed. In some cases, this attack can be used to actually recover the secret key, breaking the cryptosystem.

Since power analysis attacks are passive in that they do not alter the operation of the analyzed processor, they are difficult to detect and prevent. As such, in order to prevent this type of attack, hardware designers must ensure that any information that could be gained by power analysis is insufficient to compromise the underlying cryptosystem. One way this is done is to include useless operations in conditional branches, so that the time and power consumed does not reveal much information about input values.

SIM Cards

Many mobile phones use a special smart card called a ***subscriber identity module card*** (***SIM card***), as show in Figure 2.9. A SIM card is issued by a network provider. It maintains personal and contact information for a user and allows the user to authenticate to the cellular network of the provider. Many phones allow the user to insert their own SIM card, making the process of switching phones simple and instantaneous. Standards for SIM cards are maintained by the ***Global System for Mobile Communications*** (***GSM***), which is a large international organization.

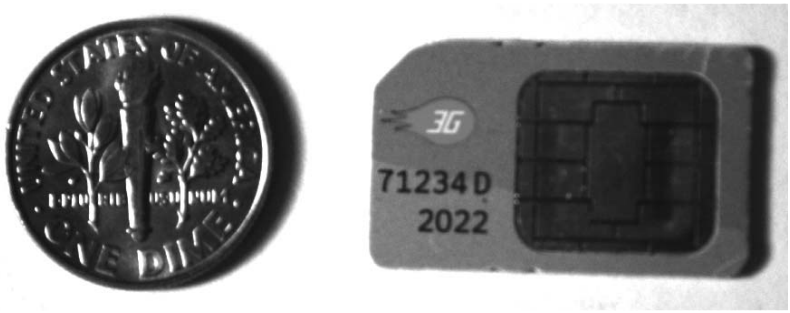


Figure 2.9: A SIM card used in a GSM cell phone, together with a dime to show size. Photo by Dan Rosenberg included with permission.

SIM Card Security

SIM cards contain several pieces of information that are used to identify the owner and authenticate to the appropriate cell network. Each SIM card corresponds to a record in the database of subscribers maintained by the network provider. A SIM card features an ***integrated circuit card ID*** (***IC-CID***), which is a unique 18-digit number used for hardware identification. Next, a SIM card contains a unique ***international mobile subscriber identity*** (***IMSI***), which identifies the owner's country, network, and personal identity. SIM cards also contain a 128-bit ***secret key***. This key is used for authenticating a phone to a mobile network, as discussed below. Finally, SIM cards may contain a contacts list.

As an additional security mechanism, many SIM cards require a PIN before allowing any access to information on the card. Most phones requiring the use of a PIN automatically lock after three incorrect password attempts. At this point, the phone can only be unlocked by providing an 8-digit ***personal unblocking key*** (***PUK***) stored on the SIM card. After ten incorrect PUK attempts, the SIM card is permanently locked and must be replaced.

GSM Challenge-Response Protocol

When a cellphone wishes to join a cellular network to make and receive calls, the cellphone connects to a local ***base station*** owned by the network provider and transmits its IMSI to declare its identity. If the IMSI corresponds to a subscriber's record in the network provider's database, the base station transmits a 128-bit random number to the cellphone. This random number is then encoded by the cellphone with the subscriber's secret key stored in the SIM card using a proprietary encryption algorithm known as ***A3***, resulting in a ciphertext block that is sent back to the base station. The base station then performs the same computation, using its stored value for the subscriber's secret key. If the two ciphertexts match, the cellphone is authenticated to the network and is allowed to make and receive calls. This type of authentication is known as a ***challenge-response*** protocol. (See Figure 2.10.)

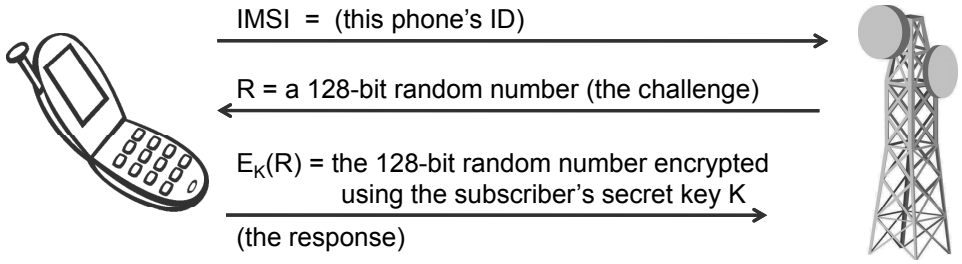


Figure 2.10: The challenge-response protocol between a cellphone (together with its SIM card) and a cell tower. The security of this protocol is derived from the fact that only the phone and the tower should know the subscriber's key.

After a SIM card has been authenticated to the network, the SIM card produces a 64-bit ciphering key by encoding the user's key and the previously sent random number, using another secret algorithm known as ***A8***. Finally, the phone is ready to make the call, and all communications are encrypted using the ciphering key with ***A5***, another proprietary algorithm.

Initially, each of the algorithms used in protecting cellphone communication (A3, A5, and A8) were proprietary algorithms developed by GSM, and were closely kept secrets. These proprietary algorithms were chosen over other public options, such as 3DES or AES, because the newly developed algorithms were optimized for cell phone hardware at the time and had significantly better performance. In many phones, the A3 and A8 algorithms are implemented as a single algorithm, known as ***COMP128***.

GSM Vulnerabilities

Older SIM cards feature an implementation of COMP128 now known as COMP128-1, which was reverse-engineered and found to be cryptographically insecure. A weakness in the algorithm reveals information about the key, given a suitable input, allowing an attacker to recover a SIM card's key by rapidly issuing requests and examining the card's output over the course of several hours. This attack could be performed over the air, without the need for physical access to the phone.

If the internal key is recovered from a phone by breaking COMP128, cloning the SIM card is relatively simple, allowing an attacker to use a victim's account to place phone calls. Newer versions of COMP128, dubbed COMP128-2 and COMP128-3, have not been broken in this way, however, and as such are not vulnerable to this type of attack. Still, because the implementations of these algorithms are secret, there is little proof of security beyond GSM's assurances.

Security flaws have also been discovered in implementations of the A5 algorithm, which is used to encrypt the actual transmission of data and voice over cell phone networks. Several cryptographic weaknesses have been identified in A5/1 (the most common version of the A5 algorithm), which allow an attacker with extensive resources to break it. Compromise of the A5 algorithm could allow an attacker to eavesdrop on cell phone communications, a major security concern given the ubiquitous use of cell phones in our society. Another implementation, known as A5/2, was designed with heavy input by intelligence agencies to ensure breakability, and was deployed only in specific countries in Eastern Europe and Asia. Unfortunately, it was proven that the algorithm can be easily cracked. Historically, A5/1 and A5/2 were initially kept secret, but they have eventually become public knowledge due to reverse engineering.

The weaknesses in COMP128 and A5 demonstrate again the risks of ***security by obscurity***—the idea that a cryptographic algorithm is safe from being broken if it is kept secret, which contradicts the open design security principle (see Section 1.1.4). Indeed, this approach is dangerous, because algorithms can often be reverse-engineered. In addition, the people who work on programming such algorithms may leak their design, either deliberately or by accident. An algorithm is much more likely to leak out than cryptographic keys, for instance, since an algorithm is always fixed and determined, whereas keys are ever changing. Fortunately, because of the past reverse-engineered attacks, future cell phone encryption methods are more likely to be based on open standards. Because of the heavy public scrutiny placed on standard cryptographic algorithms, there is higher confidence in their security.

2.3.4 RFIDs

The use of ***radio frequency identification***, or ***RFID***, is a rapidly emerging technology that relies on small transponders to transmit identification information via radio waves. Like contactless smart cards, RFID chips feature an integrated circuit for storing information, and a coiled antenna to transmit and receive a radio signal. (See Figure 2.11.)

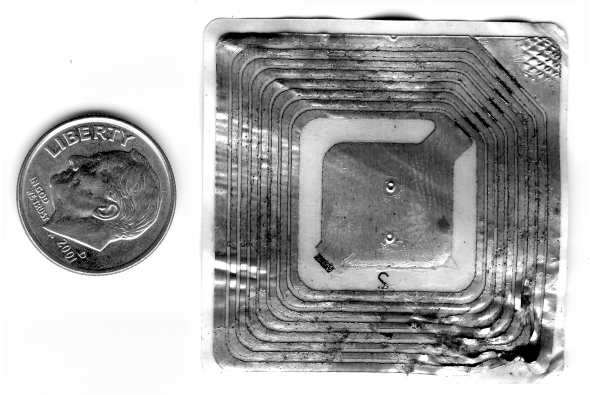


Figure 2.11: An RFID tag, taken from a DVD package, together with a dime to show size. Photo by Dan Rosenberg included with permission.

Like smart cards, RFID tags must be used in conjunction with a separate reader or writer. While some RFID tags require a battery, many are passive and do not. The effective range of RFID varies from a few centimeters to several meters, but in most cases, since data is transmitted via radio waves, it is not necessary for a tag to be in the line of sight of the reader.

This technology is being deployed in a wide variety of applications. Many vendors are incorporating RFID for consumer-product tracking, to either supplement or replace barcodes. Using RFID tags, a retailer can track which items are selling best as well as use the tags for theft detection, just by putting an RFID reader around the entrance to the shop. In addition, RFID provides an advantage over barcodes in that chips are much more difficult to replicate than simple ink on paper. Incidentally, RFID chips are also used to identify and track animals in the wild.

Because RFID chips operate using radio waves, they can release information without the need for direct physical contact. As such, it is crucial that some mechanism is employed to protect the information contained on RFID chips from unauthorized readers. If no such mechanism were used, a malicious party could easily steal personal information from a distance.

Hopping Codes and Remote Automobile Entry

Most modern vehicles feature a key fob that allows the owner to lock, unlock, or even start the engine of the vehicle from a distance. These fobs use RFID technology to communicate with a receiver in the car. Similar devices are commonly used to allow a driver to remotely open gates or garage doors. Several security measures are in place to prevent an attacker from eavesdropping on an RF transmission and recreating the signal, gaining access to the vehicle or property. The controller chips in the key fob and the receiver in the vehicle use what is known as a ***hopping code*** or ***rolling code*** to accomplish this. The controllers use the same pseudo-random number generator, so that each device produces the same sequence of unpredictable numbers.

The challenge is to keep the two sequences synchronized. When the owner presses a button, say, to unlock doors, the key fob transmits its hopping code—the next number in the key fob's sequence (along with a command instructing the car to open its doors). The receiver in the car stores a list of the next 256 hopping codes in its sequence, starting from the last time the key fob and the car synchronized their sequences. If the hopping code sent by the key fob matches one of these 256 codes, then the receiver accepts the command and performs the requested action. The receiver then updates its sequence to the next 256 numbers after the one just sent by the key fob. Once a number is used, it is never used again. Thus, even if an attacker can eavesdrop on the communication between the key fob and the car, he cannot reuse that number to open the car.

The receiver keeps a list of 256 numbers in case the fob and the receiver become desynchronized. For example, if the button on the key fob is pressed while it is out of range of the receiver, it uses up the next number in its sequence. In the event that a fob is used more than 256 times while out of range, the receiver will no longer accept its transmissions and the two will need to be resynchronized using a factory reset mechanism.

Because hopping codes are essentially random numbers, it is extremely unlikely that a key fob would be able to successfully execute a command on an unmatched receiver. Nevertheless, even though an eavesdropper cannot reuse a successful communication between a key fob and its receiver, an attack might be able to capture and replay a signal transmitted while the fob is out of range. In this case, the receiver will not have incremented its list of 256 acceptable hopping codes. To take advantage of this, some car thieves employ a technique where they jam the radio channel used by a key fob and simultaneously capture the transmission. This prevents the owner from using their key fob but allows the attacker to unlock or start the victim's car by replaying the captured signal.

KeeLoq

More recently, the actual algorithms that generates hopping codes have been subject to cryptographic attacks. The most common algorithm employed to generate the pseudo-random codes is known as ***KeeLoq***, a proprietary algorithm designed specifically for RFID hardware. The algorithm requires a 32-bit key, which is then used to encrypt an initialization vector, which in turn is incremented with each use.

Researchers have developed attacks on KeeLoq stemming from common key bits being used in certain car models. These attacks allowed to reconstruct a fob's encryption key, given a high number of captured transmissions and several days of computing time. Subsequently, a side-channel attack completely broke the KeeLoq system by measuring the power consumption of a key fob during the encryption process and using this information to recover the encryption key. Once the attacker had acquired this key, it was possible to clone a remote entry fob after intercepting two consecutive transmissions. It was also demonstrated that it was possible to use this attack to reset the internal counter of the receiver, effectively locking owners out of their cars or garages. These weaknesses in the algorithm have been addressed by increasing the size of the KeeLoq key to 60 bits, which prevents these attacks, but this change has yet to be implemented on a mass scale.

Digital Signature Transponder

Several automobile key fobs and tags for automatic payment systems at gas stations use an RFID device called ***Digital Signature Transponder*** (***DST***), which is manufactured by Texas Instruments. A DST stores a 40-bit secret key and incorporates a proprietary encryption algorithm called ***DST40***. The main use of a DST is to execute a simple challenge-response protocol, similar to the one for GSM phones (see Figure 2.10), where the reader asks the DST to encrypt a randomly generated challenge to demonstrate possession of the secret key.

Confirming once again the failure of "security by obscurity," the DST40 algorithm has been reverse engineered and an attack that recovers the secret key from two responses to arbitrary challenges has been demonstrated. This attack allows to create a new device that fully simulates a DST and can be used to spoof a reader (e.g., to charge gas purchases to the account of the DST owner).

Electronic Toll Systems

Electronic toll systems allow motor vehicle owners to place an RFID tag near their dashboards and automatically pay tolls at designated collection sites. These systems provide great convenience, since they remove the hassle of dealing with cash and allow drivers to be tolled without coming to a stop. Unfortunately, many implementations of electronic toll collection systems provide no encryption mechanism to protect the contents of the RFID tag.

Because the tag only contains a unique identifier that toll collection sites use to deduct money from an account, it is not possible to actually alter the money stored on a user's account. Still, many tags may be easily cloned, allowing a malicious party to impersonate a victim and charge tolls to their account. In addition, it may be possible to create a "digital alibi" in the event of a crime, if an attacker clones their own tag and places it on another person's automobile. If checked, the cloned tag may provide false evidence that the attacker was not at the scene of the crime.

A typical defense mechanism against cloning attacks is to install cameras to capture photographs of the license plates of vehicles that pass through toll collection sites. This approach also allows authorities to identify and impose fines on drivers with missing or expired tags.

Passports

As another example, modern passports of several countries, including the United States, feature an embedded RFID chip that contains information about the owner, including a digital facial photograph that allows airport officials to compare the passport's owner to the person who is carrying the passport. (See Figure 2.12.)

In order to protect the sensitive information on the passport, all RFID communications are encrypted with a secret key. In many instances, however, this secret key is merely the passport number, the holder's date of birth, and the expiration date, in that order. All of this information is printed on the card, either in text or using a barcode or other optical storage method. While this secret key is intended to be only accessible to those with physical access to the passport, an attacker with information on the owner, including when their passport was issued, may be able to easily reconstruct this key, especially since passport numbers are typically issued sequentially. In addition, even if an attacker cannot decrypt the contents of an embedded RFID chip, it may still be possible to uniquely identify passport holders and track them without their knowledge, since their secret key does not change.

Figure 2.12: An e-passport issued by the United States of America.

To prevent unauthorized parties from reading private information from afar without an owner's knowledge, the covers of some RFID passports contain materials that shield the passport from emitting radio waves while it is closed. Even so, these measures can be circumvented if the passport is open slightly. For example, if a passport's owner is keeping additional papers or money inside the passport, it may leak radio waves.

2.3.5 Biometrics

The term ***biometric*** in security refers to any measure used to uniquely identify a person based on biological or physiological traits. In general, biometric systems may be used to supplement other means of identification (***biometric verification***), or they may provide the sole means of authentication (***biometric identification***). Generally, biometric systems incorporate some sort of sensor or scanner to read in biometric information and then compare this information to stored templates of accepted users before granting access.

Requirements for Biometric Identification

There are several requirements that must be met for a characteristic to be consider usable as biometric identification:

- ***Universality***. Almost every person should have this characteristic. For example, the presence of a birthmark would not be acceptable biometric identification, because many people do not have birthmarks. Fingerprints, on the other hand, are universal.
- ***Distinctiveness***. Each person should have noticeable differences in the characteristic. For example, retinal images and DNA are distinctive, fingerprints are mostly distinctive, and the presence or absence of tonsils is not distinctive.
- ***Permanence***. The characteristic should not change significantly over time. For instance, fingerprints and DNA have permanence; hair color and weight do not (even though they are commonly reported on government-issued IDs).
- ***Collectability***. The characteristic should have the ability to be effectively determined and quantified.

Other considerations, which are desirable but not absolutely necessary, include performance (the accuracy and speed of recognition), acceptability (whether or not people are willing to accept the use of the biometric characteristic), and circumvention (the degree to which the characteristic can be forged or avoided). The ideal biometric system would satisfy all of these requirements, both the required and desired ones, but real-life systems tend to be lacking in some of these areas.

How Biometric Identification is Done

One of the most important aspects of any biometric system is the mechanism used to actually verify a match between a user and a stored biometric template. Systems may use several techniques to perform this sophisticated pattern matching. It would be unreasonable to expect a provided biometric sample to match up exactly with a stored template, due to slight changes in biometric features and small errors in the sample collection. Some level of flexibility must be achieved in order for the system to work at all. Nevertheless, the system must be precise enough that false matches do not occur, allowing unauthorized users to gain access to restricted resources.

Typically, this is accomplished by converting attributes of a biometric sample into a ***feature vector***—a set of data values corresponding to essential information about the sample, and comparing that vector to a stored ***reference vector***, which is a feature vector of a previous biometric sample that the system is trying to test against. (See Figure 2.13.)

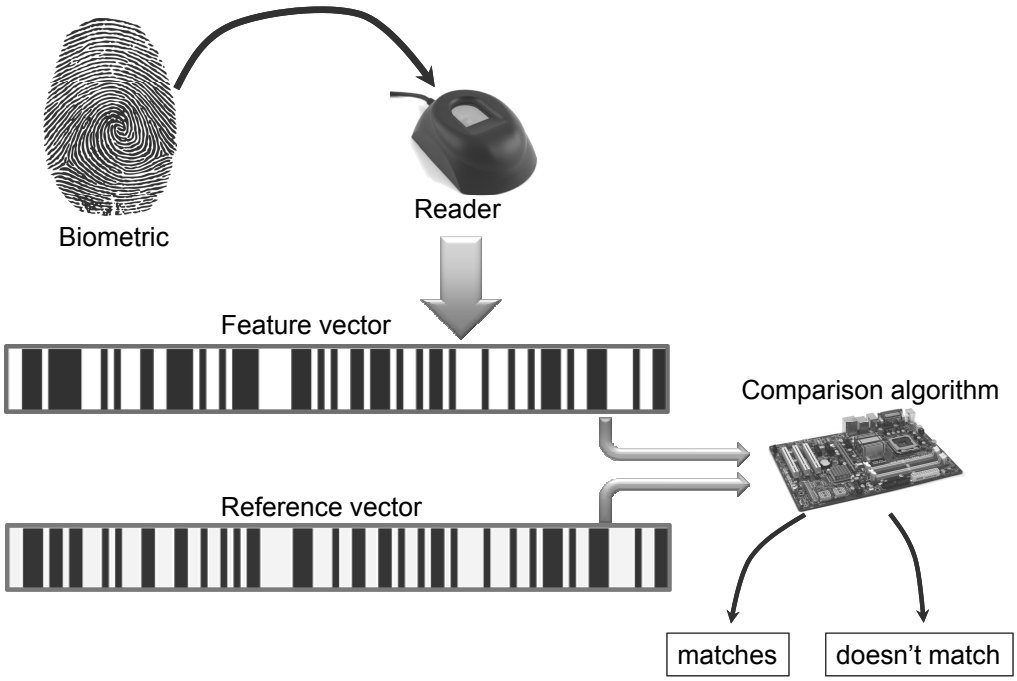


Figure 2.13: The verification process for a biometric sample. A biometric sample is converted into a feature vector and that vector is compared against a stored reference vector. If the similarity is good enough, then the biometric sample is accepted as being a match.

Generating Feature Vectors

Fingerprint pattern matching works by comparing the locations and orientations of key features in the fingerprint, such as ridge endings and bifurcations (splits in a line), while allowing for a small margin of error. Facial pattern matching is much more complex. Usually, the face is adjusted computationally so that it appears to be looking directly at the camera. Next, a feature vector is generated by calculating the location of distinct facial features, such as the ridge of the eyebrows, the edges of the mouth, the tip of the nose, and eyes. Using advanced techniques such as elastic graph theory or neural networks, this feature vector is compared to stored templates to assess the possibility of a match. Other types of biometric authentication may use different techniques to check for a match, but there is always the crucial step of generating a feature set that allows a reader to perform computational comparisons between biometric samples.

Candidate Biometric Characteristics

The most common biometric information is a person's signature, which is intended to be unique for each person. Even so, not everyone has a prepared signature, and such a signature may change over time, or may be easy to forge. Because of these limitations, signatures are not effective as a secure means of biometric authentication.

Fingerprints have been used in forensic work since the mid-19th century to identify criminals, but more recently, fingerprint scanners have been incorporated into electronic authentication systems as a means of granting access to specific users. Unlike signatures, fingerprints are universal except in rare cases, unique, easily collected and analyzed, and difficult to circumvent, making them an effective biometric characteristic. While fingerprints may change slightly over time, the degree to which they change does not affect a biometric system's ability to identify the owner.

Voice recognition does not score as well. While most people have a voice and are willing to use it as a means of authentication, it is often not distinctive enough to differentiate from another person's voice. In addition, the human voice changes significantly from year to year, and voice recognition systems can be easily circumvented using a sound recording of an authorized user.

Another common biometric system uses a person's eyes as a unique characteristic. These types of scans satisfy universality, distinctiveness, permanence, and collectability, and are very difficult to circumvent. Older systems employ retinal scanning, which involves illuminating the eye with a bright sensor and capturing an image of the blood vessels in the back of the eye. Many users find retinal scanning uncomfortable or invasive, and would prefer other means of authentication. Iris scanning systems are generally better received, providing equally strong authentication by taking a high-quality photograph of the surface of the eye.

Other biometric systems are more commonly used to identify people in public, rather than provide authentication for a select pool of users. For example, the United States government is funding research in technologies that can identify a person based on facial characteristics and gait (the unique way that a person walks), for use in applications such as airport security. The advantage of these techniques in a surveillance context is that they do not require a subject's cooperation, and can be conducted without a subject's knowledge.

Nevertheless, current implementations of these technologies are not very effective. Face recognition is not especially accurate, and does not perform well under many conditions, including poor lighting or situations where the subject's face is captured at an angle rather than straight-on. In

addition, wearing sunglasses, changing facial hair, or otherwise obstructing the face can foil facial recognition technology. Similarly, a person can defeat gait recognition by simply altering the way they walk. With further development, however, these surveillance techniques may become more accurate and difficult to circumvent.

Privacy Concerns for Biometric Data

The storage of biometric data for authentication purposes poses a number of security and privacy concerns. Access to stored biometric data may allow an attacker to circumvent a biometric system or recover private information about an individual. Since biometric data does not change over time, once a person's biometric data is compromised, it is compromised forever. As such, encryption should be used to protect the confidentiality of biometric data, both in storage and transmission. This security requirement poses a unique problem, however.

A biometric sample provided to a system by a user is not expected to match the stored template exactly—small discrepancies are expected, and allowing for these discrepancies is necessary for the system to function correctly. Thus, the comparison function between a fresh feature vector and a stored reference vector must be done to allow for slight differences, but it should also be ideally performed in a way that avoids a confidentiality breach.

The standard method of storing a cryptographic hash of a value to be kept private does not work for biometric applications. For example, suppose that we store a cryptographic hash of the reference vector obtained from the biometric template and we compare it with the cryptographic hash of the feature vector obtained from the biometric sample collected. The comparison will fail unless the sample and template are identical. Indeed standard cryptographic hash functions, such as SHA-256, are not distance preserving and are very sensitive to even small changes in the input.

Recently, various methods have been proposed that support efficient biometric authentication while preserving the privacy of the original biometric template of the user. One the approaches consists of extending the concept of a message authentication code (MAC) to that of an ***approximate message authentication code*** (***AMAC***), which has the following properties:

- Given the AMACs of two messages, it is possible to determine efficiently whether the distance between the original messages is below a certain preestablished threshold δ.

- Given the AMAC of a message, it is computationally hard to find any message within distance δ from it.

2.4 Direct Attacks Against Computers

Acquiring physical access to a computer system opens up many avenues for compromising that machine. Several of these techniques are difficult to prevent, since hardware manufacturers generally assume that the user is a trusted party. This vulnerability to physical, direct access to computers further emphasizes the need for secure access control measures that prevent physical access to sensitive computer systems. Likewise, the mere fact that computing equipment is ultimately physical implies a number of environmental considerations as well.

2.4.1 Environmental Attacks and Accidents

Computing equipment operates in a natural environment and if this environment is significantly altered, then the functionality of this computing equipment can be altered, sometimes significantly. The three main components of a computing environment are the following:

- ***Electricity***. Computing equipment requires electricity to function; hence, it is vital that such equipment has a steady uninterrupted power supply. Power failures and surges can be devastating for computers, which has motivated some data centers to be located next to highly reliable hydroelectric plants.
- ***Temperature***. Computer chips have a natural operating temperature and exceeding that temperature significantly can severely damage them. Thus, in addition to having redundant fire-protection devices, high-powered supercomputers typically operate in rooms with massive air conditioners. Indeed, the heating, ventilating, and air conditioning (HVAC) systems in such rooms can be so loud that it is difficult for people to hear one another without shouting.
- ***Limited conductance***. Because computing equipment is electronic, it relies on there being limited conductance in its environment. If random parts of a computer are connected electronically, then that equipment could be damaged by a short circuit. Thus, computing equipment should also be protected from floods.

For example, accidentally dropping one's cellphone into a pot of boiling spaghetti will likely damage it beyond repair. In general, the protection of computing equipment must include the protection of its natural environment from deliberate and accidental attacks, including natural disasters.

2.4.2 Eavesdropping

Eavesdropping is the process of secretly listening in on another person's conversation. Because of this threat, protection of sensitive information must go beyond computer security and extend to the environment in which this information is entered and read. Simple eavesdropping techniques include using social engineering to allow the attacker to read information over the victim's shoulder, installing small cameras to capture the information as it is being read, or using binoculars to view a victim's monitor through an open window. These direct observation techniques are commonly referred to as ***shoulder surfing***. Simple eavesdropping can be prevented by good environmental design, such as avoiding the placement of sensitive machines near open windows. Nonetheless, more complex techniques of eavesdropping have emerged that are more difficult to prevent.

Wiretapping

Given physical access to the cables of a network or computer, it may be possible for an attacker to eavesdrop on all communications through those cables. Many communication networks employ the use of inexpensive coaxial copper cables, where information is transmitted via electrical impulses that travel through the cables. Relatively inexpensive means exist that measure these impulses and can reconstruct the data being transferred through a tapped cable, allowing an attacker to eavesdrop on network traffic. These ***wiretapping*** attacks are passive, in that there is no alteration of the signal being transferred, making them extremely difficult to detect. (See Figure 2.14.)

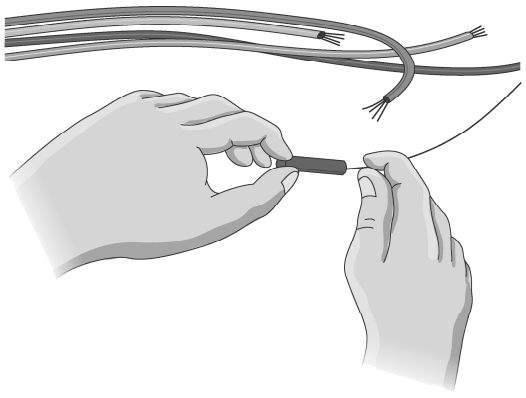


Figure 2.14: Wiretapping.

Defenses Against Wiretapping

Many networks, including much of the telephone network and most computer networks, use ***twisted pair*** cables, which feature two wires, usually copper, that are entwined to eliminate electromagnetic interference. Unshielded twisted pair (UTP) cable is inexpensive compared to coaxial or fiber optic cable. These cables are subject to the same types of signal leakage attacks as coaxial cable, without a loss of signal strength.

A more common and less expensive approach, however, is to briefly disconnect an Ethernet cable, insert a passive wiretapping device, and reconnect it. While this may go undetected by human users, many intrusion detection systems are triggered by the disconnection of network cables.

High-security networks often employ the use of fiber optic cable as a more secure alternative. Fiber optic cables transmit light rather than electricity, which prevents the signal leakage that occurs in coaxial cable. It is still sometimes possible to eavesdrop on communications transmitted over fiber optic cable, however. An attacker can place a fiber optic cable in a micro-bend clamping device, which holds the cable in a bent position, where it leaks a tiny amount of light. An attached photo sensor can transmit the information via an optical-electrical converter, where it can be reinterpreted by a computer. This attack results in a tiny drop in the signal being transmitted over the network, so it may be detected by fiber optic intrusion detection systems. More advanced attacks may employ means of reboosting the signal to make up for this signal drop.

Both of these attacks demonstrate the importance of protecting not only computer systems, but also the network cables over which sensitive information is transmitted. Attacks on fiber optic cables are expensive and may be detected, but are still a possibility. Many organizations use end-to-end encryption to protect data being transmitted over the network—eavesdropping is rendered useless if the contents are not readable by the attacker.

Radio Frequency Emissions

One of the earliest techniques of computer eavesdropping gained widespread attention through the 1985 publication of a paper by Dutch computer researcher Wim van Eck. Cathode Ray Tube (CRT) displays, used by older computer monitors, emit electromagnetic radiation in the Radio Frequency (RF) range.

Van Eck demonstrated that these emissions could be read from a distance and used to reconstruct the contents of a CRT screen. Since RF emissions can travel through many nonmetallic objects, computer monitors could be read regardless of whether they are within eye-shot of an attacker.

More recent research has extended this principle to modern Liquid Crystal Display (LCD) screens. Fortunately, preventative measures have been developed that utilize techniques to shield monitors and reduce these emissions, but they are rarely deployed outside of high-security government applications, due to a low prevalence of this type of attack and the expensive cost of equipment necessary to perform it. In an environment where other attacks are impossible due to security measures, however, this form of eavesdropping is certainly within the realm of possibility.

Optical Emissions

A more recent attack has emerged that allows eavesdropping using emissions in the range of visible light, rather than the RF range. The attack requires the use of a photosensor, which is relatively cheap compared to the expensive equipment needed for an RF eavesdropping attack. CRT displays work by using an electron beam that scans the surface of the screen, refreshing each pixel individually at incredibly fast speeds. At the moment the electron beam hits a pixel, there is a brief burst of brightness, which makes this attack possible. A photosensor can be trained on a wall in the room, and by analyzing changes in the light of the room and applying imaging technology to reduce "noise," it is possible to reconstruct an image of the contents of the screen. This attack has been proven to work from up to 50 meters away, requiring only that the attacker can train a photosensor on any wall in the room. Fortunately, since the attack relies on visible light, as long as a room is not in an attacker's line of sight, it is safe from this attack. Also, this attack does not work on LCD monitors, because of differences in how pixels are refreshed on the screen. Like RF eavesdropping, this attack is possible, but considered unlikely to occur in most contexts, and, especially with the advent of LCD screens, it is not expected to be a high-priority security concern in the future.

Acoustic Emissions

In addition to RF radiation and visible light, computer operations often result in another byproduct, sound. Recent research has proven it possible to use captured acoustic emissions to compromise computer security. These techniques are still in their infancy, so they are unlikely to occur outside a lab, but they may become security concerns later on.

Dmitri Asonov and Rakesh Agrawal published a paper in 2004 detailing how an attacker could use an audio recording of a user typing on a keyboard to reconstruct what was typed. (See Figure 2.15.) Each keystroke has minute differences in the sound it produces, and certain keys are known to be pressed more often than others. After training an advanced neural network to recognize individual keys, their software recognized an average 79% of all keystrokes.

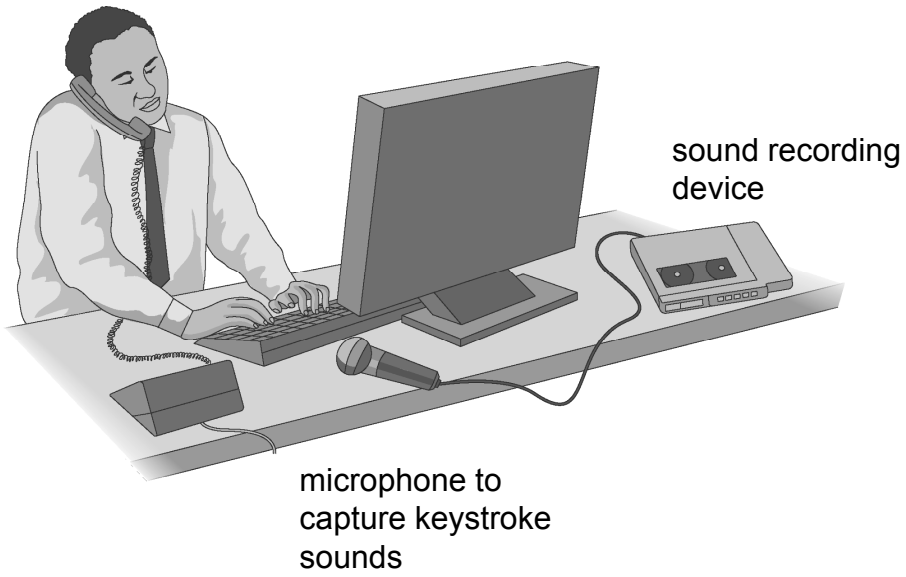


Figure 2.15: A schematic of how a keyboard acoustic recorder works.

Also in 2004, researchers Adi Shamir and Eran Tromer conducted an experiment that demonstrated the possibility of revealing a machine's CPU instructions by analyzing acoustic emissions from the processor. In theory, this may provide attackers with additional information about the inner workings of a computer, including exposing which routine or program is being executed to perform a certain task. In addition, information can be gathered to attack cryptographic functions, such as the algorithm used and the time required for each computation.

Hardware Keyloggers

A keylogger is any means of recording a victim's keystrokes, typically used to eavesdrop passwords or other sensitive information. There are many ways of implementing software keyloggers, which are discussed, along with other types of malware, in Chapter 4. A newer innovation,

however, is the hardware keylogger. Hardware keyloggers are typically small connectors that are installed between a keyboard and a computer. For example, a USB keylogger is a device containing male and female USB connectors, which allow it to be placed between a USB port on a computer and a USB cable coming from a keyboard. (See Figure 2.16.)

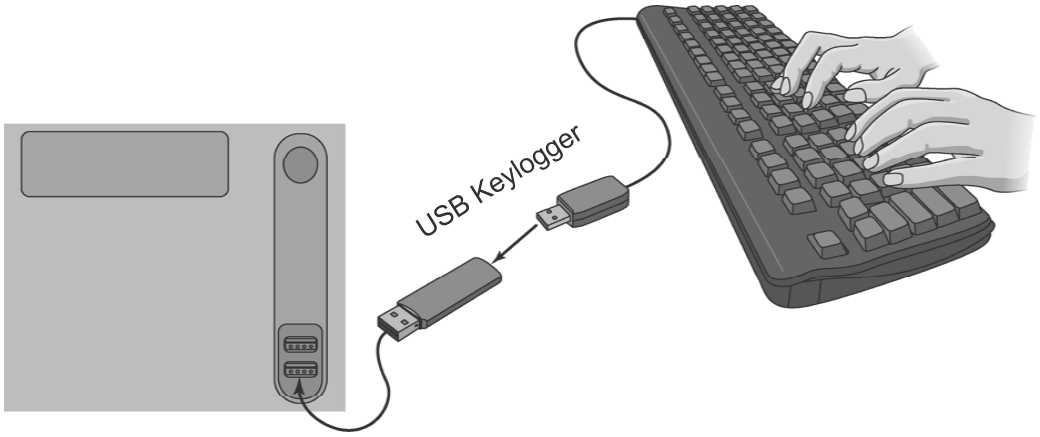


Figure 2.16: A schematic of how a USB keylogger works.

By including circuits that capture keystrokes and store them in a flash memory, a hardware keylogger can collect and store all the keystrokes coming from the keyboard over an extended period of time. An attacker can install a device like this in an Internet cafe, leave it to collect keystrokes for a week or more, and then come back to retrieve the device and download all the keystrokes. Thus, an attacker using such a device can hope to collect passwords and other personal information from the people who use the compromised keyboard.

While some advanced hardware keyloggers transmit captured text via wireless technology, most rely on the attacker's ability to retrieve them at a later date. After installing the device, it is completely undetectable by software, and since it operates at the hardware level, it can even record BIOS passwords entered before booting to the operating system. Because of this stealth, the best detection method is simple physical inspection, and the most effective preventative measure is employing strict access control to prevent physical access to sensitive computer systems.

2.4.3 TEMPEST

TEMPEST is a U.S. government code word for a set of standards for limiting information-carrying electromagnetic emanations from computing equipment. More broadly, the term "TEMPEST" has come to be used for the study, limitation, and protection of all types of information-carrying emanations that come from computing equipment and devices. In terms of a standard, TEMPEST establishes three zones or levels of protection:

1. An attacker has almost direct contact with the equipment, such as in an adjacent room or within a meter of the device in the same room.

2. An attacker can get no closer than 20 meters to the equipment or is blocked by a building to have an equivalent amount of attenuation.

3. An attacker can get no closer than 100 meters to the equipment or is blocked by a building to have an equivalent amount of attenuation.

To achieve the limits imposed by these three levels of protection, engineers can use emanation blockage and/or emanation modification.

Emanation Blockage

One approach to limiting the release of information-carrying emanations is to enclose the computing equipment in a way that blocks those emanations from escaping into the general environment. Some examples of this type of emanation limitation include the following:

- To block visible light emanations, we can enclose sensitive equipment in a windowless room.

- To block acoustic emanations, we can enclose sensitive equipment in a room lined with sound-dampening materials.

- To block electromagnetic emanations in the electrical cords and cables, we can make sure every such cord and cable is grounded, so as to dissipate any electric currents traveling in them that are generated from external (information-carrying) electromagnetic fields created by sensitive computing equipment.

- To block electromagnetic emanations in the air, we can surround sensitive equipment with metallic conductive shielding or a mesh of such material, where the holes in the mesh are smaller than the wavelengths of the electromagnetic radiation we wish to block. Such an enclosure is known as a ***Faraday cage***. (See Figure 2.17.)

Figure 2.17: An example Faraday cage. (Image by M. Junghans; licenced under the terms of the GNU Free Documentation License, Version 1.2.)

In order for these emanation blockage techniques to work, the sensitive computing equipment (including all its cables and junction boxes) have to be completely enclosed. Examples of such enclosures range from a classified office building, which is completely enclosed in a copper mesh and has two-pass copper doors for entering and exiting, to a metal-lined passport wallet, which encloses an RFID passport in a small Faraday cage so as to block unwanted reading of the RFID tag inside it.

Emanation Masking

Another technique for blocking information-carrying electromagnetic emanations is to mask such emanations by broadcasting similar electromagnetic signals that are full of random noise. Such emanations will interfere with the information-carrying ones and mask out the information in these signals by introducing so much noise that the information-carrying signal is lost in the cacophony.

2.4.4 Live CDs

A ***live CD*** is an operating system that can be booted from external media and resides in memory, without the need for installation. It can be stored on a CD, DVD, USB drive or any other removable drive from which the computer can boot. There are many legitimate uses for live CDs, including diagnostic and software repair purposes. Unfortunately, an attacker can boot from a live CD, mount the hard disk, and then read or write data, bypassing any operating system authentication mechanisms. A native operating system can do nothing to prevent this, because it is never loaded. Therefore, preventative measures must be built into hardware.

One effective means of preventing a live-CD attack is by installing a BIOS password. As discussed in Chapter 3, the BIOS is firmware code that is executed immediately when a machine is turned on and before loading the operating system. By protecting the BIOS, an attacker is unable to boot the computer without a password. Note, however, that this does nothing to prevent an attacker from removing the actual hard drive from the machine, mounting it in another machine off-site, and then booting to a live CD.

This vulnerability suggests the need for locking mechanisms preventing access to the interior of a sensitive computer system. Other prevention tactics include using a built-in hard drive password or utilizing hard disk encryption technology.

2.4.5 Computer Forensics

Computer forensics is the practice of obtaining information contained on an electronic medium, such as computer systems, hard drives, and optical disks, usually for gathering evidence to be used in legal proceedings. Unfortunately, many of the advanced techniques used by forensic investigators for legal proceedings can also be employed by attackers to uncover sensitive information. Forensic analysis typically involves the physical inspection of the components of a computer, sometimes at the microscopic level, but it can also involve electronic inspection of a computer's parts as well. (See Figure 2.18.)

An important principle of computer forensics is to establish, maintain, and document a ***chain of custody*** for the computer hardware being analyzed so that it can be shown that the items collected remains unaltered throughout the forensics analysis process.

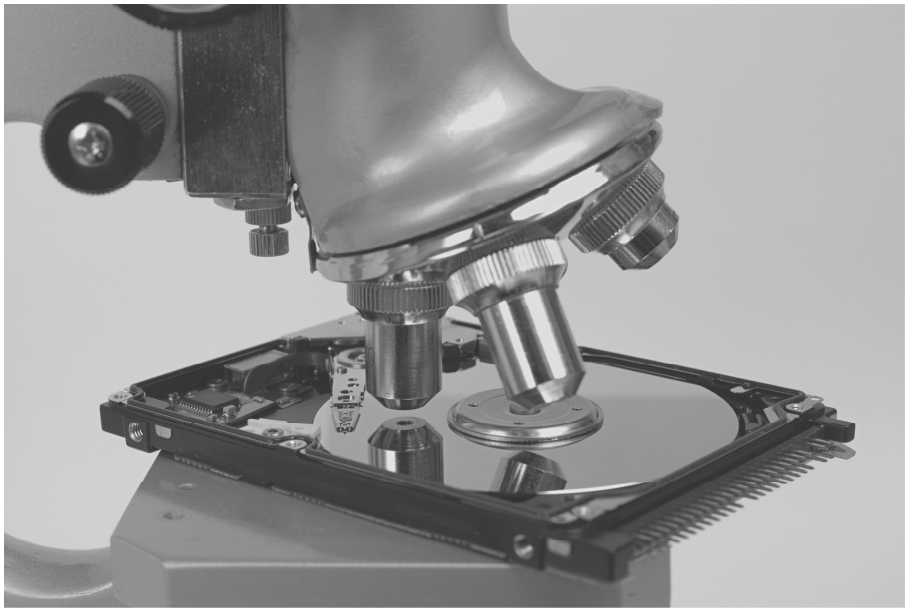


Figure 2.18: Microscopic inspection of a disk drive.

Security Concerns from Computer Forensics

Often, forensic analysis of a system while it is turned on can reveal information that would not be obtainable if it were powered off. For example, online analysis allows an investigator (or attacker) to use tools to examine or copy the contents of RAM, which is volatile and disappears when the computer is turned off. By examining RAM, an attacker could uncover recently entered passwords or other sensitive information that would be unavailable if the machine were off. In addition, online attacks can often reveal information about a machine's presence on a network.

Because computer forensics is designed to provide evidence that is suitable for use in court, most analysis is performed while the machine is turned off, in order to establish that its contents have not been altered in the process of the investigation. By mounting a hard drive in another machine, most investigators begin by making an exact copy of the entire hard disk and performing analysis on the copy.

Using forensic techniques, it may be possible to recover data that a user deleted. File operations on a computer, including reading, writing, and deleting files, are controlled by a portion of the operating system known as a filesystem. In the process of deleting a file, many filesystems only remove the file's metadata—information about the file including its size,

location on disk, and other properties—without actually overwriting the contents of the data on the disk. The space in which the file's data resides is freed, in that future file operations are allowed to overwrite it, but until it is overwritten, the deleted file's data will remain on the disk. Because of this, forensic investigators can use tools to analyze the contents of the disk to uncover "deleted" data.

The typical hard drive uses magnetic disks to retain data. A byproduct of this medium is that overwriting data may leave faint magnetic indicators of the state of the information bits before they were overwritten. It is possible that advanced hardware forensics techniques can be used to recover some overwritten data. With the increasing density of how information is stored on hard disks, this type of attack has become more difficult, since the probability of successfully recovering any usable amount of data by examining microscopic magnetic residue is prohibitively small. Nonetheless, United States government standards mandate that in order to safely delete classified information on magnetic media beyond all chance of recovery, it must be overwritten with multiple passes of random data or be physically destroyed. Note that flash media, which does not rely on magnetic disks or tape, is not susceptible to this type of attack—a single pass of overwriting is sufficient to remove data beyond chance of recovery in flash memory.

Cold Boot Attacks

In 2008, a team of Princeton researchers presented a technique that can be used to access the contents of memory after a computer has been shut down. ***Dynamic random-access memory*** (***DRAM***) is the most common type of computer memory. DRAM modules are volatile storage, which means that their contents decay quickly after a computer is turned off. Even so, the study showed that by cooling DRAM modules to very low temperatures, the rate of decay can be slowed to the point where the contents of memory can be reconstructed several minutes after the machine has powered off.

Using this technique, the researchers were able to bypass several popular drive encryption systems (see Section 9.7). Their ***cold boot attack*** consists of freezing the DRAM modules of a running computer by using a refrigerant (e.g., the liquid contained in canned-air dusters), powering off the computer, and booting it from a live CD equipped with a program that reconstructs the memory image and extract the disk encryption key (which was stored in unencrypted form in memory).

2.5 Special-Purpose Machines

There are certain types of computing machines that have a special purpose, that is, particular jobs that they are specialized to do. These jobs might involve sensitive information or tasks, of course, which presents particular security requirements. In this section, we study two such machines—automated teller machines and voting machines—and we discuss the particular types of risks that these machines have with respect to both their physical and digital security.

2.5.1 Automated Teller Machines

An ***automatic teller machine*** (***ATM***) is any device that allows customers of financial institutions to complete withdrawal and deposit transactions without human assistance. Typically, customers insert a magnetic stripe credit or debit card, enter a PIN, and then deposit or withdraw cash from their account. The ATM has an internal cryptographic processor that encrypts the entered PIN and compares it to an encrypted PIN stored on the card (only for older systems that are not connected to a network) or in a remote database. The PIN mechanism prevents an attacker with access to a stolen card from accessing account funds without additional information. Most financial institutions require a 4-digit numeric PIN, but many have upgraded to 6 digits. To prevent guessing attacks, many ATMs stop functioning after several failed PIN attempts. Some retain the previously inserted card, and require contacting a bank official in order to retrieve it.

ATM Physical Security

The ATM's role as a cash repository has made it a popular target for criminal activity. Several measures are commonly employed to prevent tampering, theft, and to protect sensitive customer information. Firstly, the vault, which contains any valuable items such as cash, must be secured. Vaults are often attached to the floor to prevent casual theft and include high-security locking mechanisms and sensors to prevent and detect intrusion.

While these measures are effective at preventing on-site removal of cash, they are ineffective at deterring more brazen criminals from using heavy construction equipment and a large vehicle to uproot and remove an entire ATM. In some instances, attackers go so far as to drive a vehicle through

the doors or windows of a financial institution to allow easy access to an ATM. This technique is known as ***ram-raiding***, and can be prevented by installing vehicular obstructions such as bollards. Other attacks include using carefully placed explosives to compromise the vault. To compensate for an inability to guarantee physical integrity in all situations, most modern ATMs rely on mechanisms that render their contents unusable in the event of a breach, such as dye markers that damage any cash inside.

ATM Encryption

To ensure the confidentiality of customer transactions, each ATM has a cryptographic processor that encrypts all incoming and outgoing information, starting the moment a customer enters their PIN. The current industry standard for ATM transactions is the ***Triple DES*** (***3DES***) cryptosystem, a legacy symmetric cryptosystem with up to 112 bits of security (See Figure 2.19.)

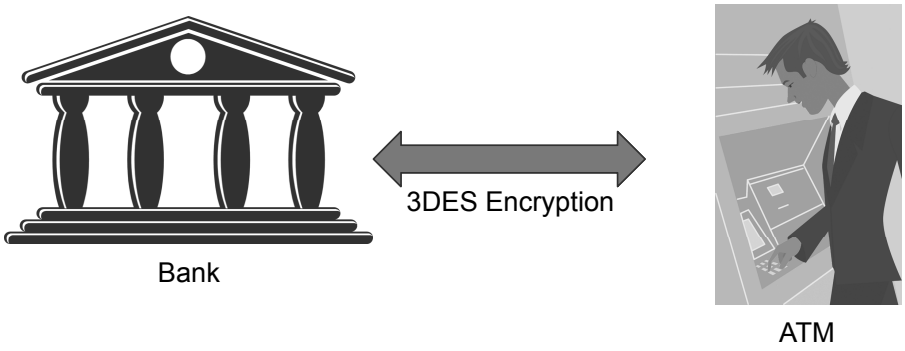


Figure 2.19: ATM communications are typically encrypted using the 3DES symmetric cryptosystem.

The 3DES secret keys installed on an ATM are either loaded on-site by technicians or downloaded remotely from the ATM vendor. Because the confidentiality of all transactions on an ATM relies on protecting the secrecy of the cryptographic keys, any attempts to access the cryptoprocessor will destroy the keys. It should be noted that since early ATM machines used the obsolete DES cryptosystem with 56-bit keys, the 3DES cryptosystem was chosen over the more secure AES cryptosystem because 3DES is backward compatible with DES and thus moving to 3DES was seen as a simple and inexpensive way to increase the key size. In addition, AES was not finalized as a standard until 2001, roughly three years after 3DES was standardized.

Attacks on ATMs

There are several techniques used to perpetrate ATM fraud. One popular attack involves the use of a thin sleeve of plastic or metal known as a ***Lebanese loop***. A perpetrator inserts this sleeve into the card slot of an ATM. When a customer attempts to make a transaction and inserts their credit card, it sits in the inconspicuous sleeve, out of sight from the customer, who thinks that the machine has malfunctioned. After the customer leaves, the perpetrator can then remove the sleeve with the victim's card.

Another technique makes use of a device known as a ***skimmer***, which reads and stores magnetic stripe information when a card is swiped. An attacker can install a skimmer over the card slot of an ATM and store customers' credit information without their knowledge. Later, this information can be retrieved and used to make duplicates of the original credit cards.

Finally, some scammers may even install fake ATMs in remote locations to capture both credit/debit cards and PINs at the same time. These fake ATMs typically respond with a fake error message after the cards and PINs have been captured, so as not to arouse the suspicions of the users.

In many cases, the card number or physical card is all that is necessary to make financial transactions, but if an attacker wishes to withdraw money from an ATM or make a debit transaction, a PIN is also required. Perpetrators may employ any number of eavesdropping techniques to acquire PINs, including installing cameras at ATM locations. Some attackers may install fake keypads that record customer PINs on entry. Collectively, these attacks stress the importance of close surveillance at ATM sites. Cameras and regular security checks are effective at deterring attacks as well as identifying culprits.

2.5.2 Voting Machines

Since the 1960s, electronic systems have been used around the world for another crucial function, voting. Electronic voting systems collect and tally votes for elections around the world, including the presidential election in the United States. Clearly, security is paramount—weaknesses could result in falsified elections and deprive citizens of their rights to voice their opinions on issues and leaders.

Types of Voting Machines

There are two general types of electronic voting, paper-based and direct-recording. In a paper-based system, voters submit their votes on a piece of paper or a punchcard, after which it is counted either by hand or by

an optical scanner designed to read and tally marked ballots. Paper-based systems have several advantages, including the fact that most people are familiar with how they work and they allow for hand recounts.

The other type of voting machine, which is used by many countries, is the direct-recording system, where votes are submitted and tallied electronically, using touch-screen technology, for example. These systems are faster, more environmentally friendly, more accessible to handicapped voters, and ostensibly more accurate, since they remove the additional step of tallying votes on paper. Nevertheless, these electronic voting systems are not as amenable to hand recounts, since they don't provide a paper audit trail.

Voting Machine Security

Both types of electronic voting systems introduce new potential avenues for electoral fraud. Coordinating an election across a region as large as the United States requires several steps. First, individual voting machines must accurately tally individual votes, and be tamper proof. Next, the transmission of vote totals to a centralized location must be done securely and in a way that prevents alteration of vote tallies. Finally, these centralized locations must calculate the final totals correctly in a tamper-proof way.

Most electronic voting machines in the United States are manufactured by Diebold, which is also the largest supplier of ATMs in the country. These voting machines are made with a closed-source platform, despite the demands of many information security experts, who claim that public scrutiny is the only way to verify the safety of electronic voting. Diebold publicizes that its voting machines use AES encryption to encrypt stored data, digitally signed memory cards, and Secure Socket Layer (SSL) encryption (see Section 7.1.2) to transmit vote data. Despite these measures, several researchers have demonstrated the possibility of tampering with these systems.

A group of Princeton researchers showed that by gaining physical access to a Diebold AccuVote-TS voting machine for one minute, an attacker could introduce malicious code into the machine that allowed the attacker to manipulate vote totals, delete specific votes, and perform other forms of voting fraud. Diebold issued a statement that the voting machine used in the study was obsolete, but the researchers insisted that newer machines are vulnerable to the same types of attacks. In any case, with an increased reliance on electronic voting during elections, extensive measures should be taken to assure the security of this important process.

2.6 Physical Intrusion Detection

Intrusion detection systems alert the owners of a facility, information, or other sensitive resources if that resource's security has been compromised. While visible intrusion detection equipment may act as a deterrent, these systems are primarily intended as a response measure rather than a preventative one. There are typically two parts to any intrusion detection system, detection and response.

2.6.1 Video Monitoring

Video monitoring systems are a standard means of intrusion detection. A network of video cameras remotely accessible via the Internet or a legacy ***closed-circuit television*** (***CCTV***) system, which uses a proprietary network, allow a centralized operator to monitor activity in many locations at once. (See Figure 2.20.) Most video monitoring systems are effective at providing evidence of wrongdoing, because videos can be recorded and archived. Of course, in order to be effective at intrusion detection, such systems require a human operator to successfully identify malicious activity.

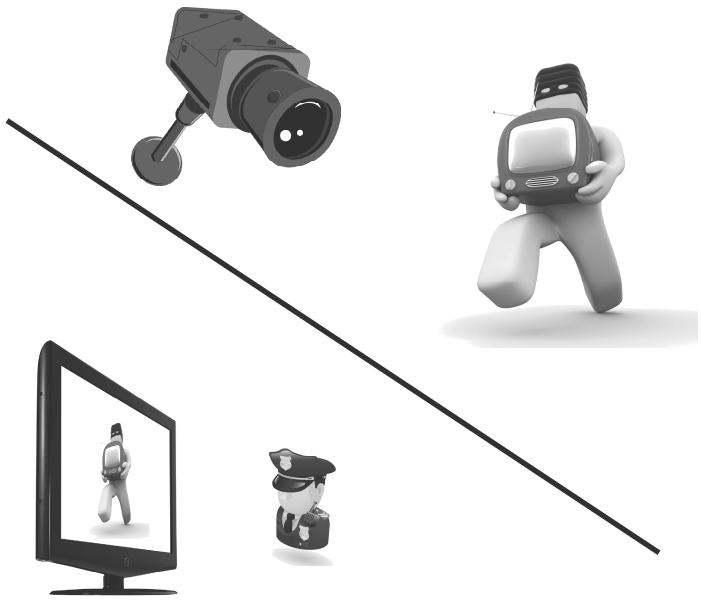


Figure 2.20: The components in a video monitoring security system.

More advanced video monitoring systems can automatically track movement across multiple camera zones, eliminating the need for a human operator. Systems are in development that can detect suspicious behavior in a crowded area by analyzing body movement patterns or particular types of clothing. Such methods of intrusion detection are designed to work automatically, without human assistance. Likewise, a motion sensor is a device that detects movement in a space using any number of mechanisms. For example, some sensors use infrared imaging and are triggered by changes in heat. Other sensors employ ultrasonic technology—the sensor emits an inaudible sound wave pulse and measures the reflection off objects in the room. Finally, other systems are triggered by changes in the sound of a room. In each case, triggered sensors may sound an alarm, either to deter attackers or alert security staff, activate a video monitoring system to record the intrusion, or activate additional locking mechanisms to protect resources or trap the intruder.

Several of the physical intrusion detection mechanisms mentioned above may be defeated. For example, a CCTV system may fail to provide crucial evidence if an intruder makes efforts to disguise his or her features, if cameras are dismantled or otherwise tampered with, or if an intruder is careful to stay out of sight. Infrared motion sensors may be defeated by placing a material that prevents the dissipation of body heat, such as a pane of glass or insulating suit, between the camera and the intruder. Ultrasonic sensors may be thwarted by using sound-dampening materials to prevent the pulse of the sensor from detecting the intruder. Finally, audio sensors can of course be defeated by remaining extremely quiet. Because of the relative ease of circumvention, most modern intrusion detection systems employ sensors that use a variety of technologies, making the system much more difficult to defeat.

Examining physical intrusion detection systems can provide some insights on what makes an effective network intrusion detection system, which is discussed in Chapter 6. Like physical intrusion detection, network intrusion detection can be used both as a preventative measure (where the response is intended to stop the intrusion) or as a means of providing important evidence after the breach (for example, keeping thorough log files). Also, the most effective network intrusion detection systems do not rely on a single mechanism to detect a breach, but rather employ a wide variety of techniques to prevent easy circumvention. Nevertheless, both types of systems often feature a critical component that cannot be overlooked, human involvement.

2.6.2 Human Factors and Social Engineering

Despite technological advances, using human guards is still one of the most common means of detecting intruders. In addition, most response measures to intrusion are dependent on fast human action. While each technology has its advantages, humans can adapt in ways that most computers can't, giving humans greater flexibility in security applications. Moreover, human perception can pick up details that computers miss.

Introducing people into a security model can result in a number of potential problems, however. For example, human-in-the-loop security solutions are vulnerable to ***social engineering*** attacks. (See Section 1.4.3.) Indeed, a major issue with the human element is reliability. Of course, computers are not perfect: software often has bugs, hardware occasionally fails, and sometimes systems seem to break without cause. Humans, on the other hand, may be unreliable for a whole slew of reasons, including improper training, physical ailment, ulterior motives, or simple lack of judgment. (See Figure 2.21.)

Figure 2.21: An example of a social engineering attack on a security guard: "Thanks for understanding about me leaving my ID card at home."

Human reliability also extends to computer security applications—equipment must be properly configured, monitored, and implemented in a way that is effective. Many examples of system compromise occur as a result of a single network administrator failing to install critical security patches or improperly monitoring server logs. These are mistakes that can be prevented by placing a high emphasis on training for all personnel, especially security and systems personnel.

2.7 Exercises

For help with exercises, please visit **securitybook.net**.

Reinforcement

R-2.1 Would increasing the number of pins in the design of a pin tumbler lock increase its security?

R-2.2 Would increasing the number of available pin heights in the design of a pin tumbler lock increase its security?

R-2.3 What do billiards and lock bumping have in common?

R-2.4 Given a change key for a certain type of lock, describe how to derive from it a bump key that works with all the locks of that type.

R-2.5 What is the full theoretical size of the search space for a pin tumbler lock that has 30 possible key blanks and 8 pins, each with 12 different distinct heights? What is the corresponding theoretical size of the search space for the corresponding iterative master-key construction?

R-2.6 Consider a pin tumbler lock with 5 pins and 8 pin heights. Explain why it is not actually possible to have 8^5 different change keys.

R-2.7 The Acme Combination is rated as a two-hour lock, meaning that it takes two hours to crack this lock by an experienced thief. The Smacme company has a half-hour lock that looks exactly the same as the Acme lock and is much cheaper to buy. The XYZ Company wanted to save money, so they bought one Acme lock and one Smacme lock. They put one on their front door and one on the back door of their building. Explain how an experienced thief should be able to break into the XYZ Company's building in about an hour or less.

R-2.8 Explain why storing secret encryption/decryption keys in a removable drive helps defend against cold boot attacks.

R-2.9 Among radio-frequency, optical, and radio emissions, which poses the most significant privacy threat for a user? Consider the cases of a home office, public library, and university department.

R-2.10 Explain why knowing in which language the user is typing helps perform an eavesdropping attack based on analyzing acoustic keyboard emissions.

R-2.11 Discuss whether barcodes are more or less secure than magnetic stripe cards.

R-2.12 Describe an application where smart cards provide sufficient security but magnetic stripe cards do not.

R-2.13 What are the main security vulnerabilities of SIM cards?

R-2.14 What happens if you accidentally press a car key fob 257 times while being far away from the car?

R-2.15 A salesperson at a high-end computer security firm wants to sell you a protective cover for your passport, which contains an RFID tag inside storing your sensitive information. The salesperson's solution costs "only" $79.99 and protects your passport from being read via radio waves while it is in your pocket. Explain how you can achieve the same thing for under $3.00.

R-2.16 How can you check if a public computer has a USB keylogger installed?

R-2.17 Describe which properties, such as universality, distinctiveness, etc., each of the following biometric identification characteristics do and do not possess: DNA, dental x-ray, fingernail length, and blood type.

Creativity

C-2.1 Describe a simple modification of the design of pin tumbler locks to defend against lock-bumping attacks.

C-2.2 For safety reasons, external locked doors on commercial buildings have mechanisms for people on the inside to escape without using a key or combination. One common mechanism uses an infrared motion detector to open an electronic lock for people moving towards a door from the inside. Explain how an air gap under such an external door could be exploited to open that door from the outside?

C-2.3 A group of n pirates has a treasure chest and one unique lock and key for each pirate. Using hardware that is probably already lying around their ship, they want to protect the chest so that any single pirate can open the chest using his lock and key. How do they set this up?

C-2.4 A group of n red pirates and a group of n blue pirates have a shared treasure chest and one unique lock and key for each pirate. Using hardware that is probably already lying around their two ships, they want to protect the chest so that any pair of pirates, one red

and one blue, can open the chest using their two locks and keys, but no group of red or blue pirates can open the chest without having at least one pirate from the other group. How do they set this up?

C-2.5 A group of four pirates has a treasure chest and one unique lock and key for each pirate. Using hardware that is probably already lying around their ship, they want to protect the chest so that any subset of three of these pirates can open the chest using their respective locks and keys, but no two pirates can. How do they set this up?

C-2.6 A thief walks up to an electronic lock with a 10-digit keypad and he notices that all but three of the keys are covered in dust while the 2, 4, 6, and 8 keys show considerable wear. He thus can safely assume that the 4-digit code that opens the door must be made up of these numbers in some order. What is the worst case number of combinations he must now test to try to open this lock using a brute-force attack?

C-2.7 You want to plant a bug in Company X's office to acquire business intelligence because they are a competitor. The package needs to get into their server room and get hooked up to sensitive hardware. You know the complex hires several guards from a private security company that regularly patrol and check for authentication by using well-known badges. You know that they regularly outsource several functions including janitorial staff, pest control, and purchasing IT equipment (think Staples delivery trucks). These jobs have a high turnover rate, but require authentication in order to get access to the premises in the form of a work order for IT supplies and pest control. The janitorial staff is a recurring service, but with a lower turnover rate. They are also periodically inspected by officials like the city or OSHA (Occupational Safety and Health Administration, an agency of the United States Department of Labor), but are usually provided with advanced notice of their arrival. What is your high-level plan of action? A guard challenges you when you enter, how do you continue your mission? What is your legend? What is your story? Why is this a good plan? What are your options for acquiring access to sensitive areas? You realize you are a target to this attack. How will you defend against it?

C-2.8 You are planning an urban exploration journey into the abandoned train tunnel of Providence. It has two ends, one of which is in a place you vaguely know, in the woods off the road, and the other is near a moderately populated street corner. Each end is secured with a simple padlock. The doors are clearly marked "no trespassing." Which end do you select and why? How do

you justify being at the end of the tunnel if you are observed and questioned? What are some of the dangers of this operation? What time of day do you go on this trip? Weekday or weekend?

C-2.9 A variation of the following biometric authentication protocol was experimentally tested several years ago at immigration checkpoints in major U.S. airports. A user registers in person by showing her credentials (e.g., passport and visa) to the ***registration authority*** and giving her fingerprint (a "palmprint" was actually used). The registration authority then issues to the user a tamper-resistant ***smartcard*** that stores the reference fingerprint vector and can execute the matching algorithm. The checkpoint is equipped with a tamper resistant ***admission device*** that contains a fingerprint reader and a smartcard reader. The user inserts her smartcard and provides her fingerprint to the device, which forwards it to the smartcard. The smartcard executes the comparison algorithms and outputs the result ("match" or "no match") to the device, which admits or rejects the user accordingly. Clearly, an attacker can defeat this scheme by programming a smartcard that always outputs "match." Show how to modify the scheme to make it more secure. Namely, the admission device needs to make sure that it is interacting with a valid smartcard issued by the registration authority. You can assume that the smartcard can perform cryptographic computations and that the admission device knows the public key of the registration authority. The attacker can program smartcards and is allowed to have an input-output interaction with a valid smartcard but cannot obtain the data stored inside it.

C-2.10 To save on the cost of production and distribution of magnetic stripe cards, a bank decides to replace ATM cards with printed two-dimensional barcodes, which customers can download securely from the bank web site, and to equip ATM machines with barcode scanners. Assume that the barcode contains the same information previously written to the magnetic stripe of the ATM card. Discuss whether this system is more or less secure than traditional ATM cards.

C-2.11 A bank wants to store the account number of its customers (an 8-digit number) in encrypted form on magnetic stripe ATM cards. Discuss the security of the following methods for storing the account number against an attacker who can read the magnetic stripe: (1) store a cryptographic hash of the account number; (2) store the ciphertext of the account number encrypted with the bank's public key using a public-key cryptosystem; (3) store the

ciphertext of the account number encrypted with the bank's secret key using a symmetric cryptosystem.

C-2.12 Consider the following security measures for airline travel. A list of names of people who are not allowed to fly is maintained by the government and given to the airlines; people whose names are on the list are not allowed to make flight reservations. Before entering the departure area of the airport, passengers go through a security check where they have to present a government-issued ID and a boarding pass. Before boarding a flight, passengers must present a boarding pass, which is scanned to verify the reservation. Show how someone who is on the no-fly list can manage to fly provided boarding passes can be printed online. Which additional security measures should be implemented in order to eliminate this vulnerability?

C-2.13 Develop a multiuser car-entry system based on RFID fobs. The system should support up to four distinct key fobs.

C-2.14 Consider the following simple protocol intended to allow an RFID reader to authenticate an RFID tag. The protocol assumes that the tag can store a 32-bit secret key, s, shared with the reader, perform XOR operations, and receive and transmit via radio 32-bit values. The reader generates a random 32-bit challenge x and transmits $y = x \oplus s$ to the tag. The tag computes $z = y \oplus s$ and sends z to the reader. The reader authenticates the tag if $z = x$. Show that a passive eavesdropper that observes a single execution of the protocol can recover key s and impersonate the tag. What if the tag and reader share two secret keys s_1 and s_2, the reader sends $x \oplus s_1$ and the tag responds with $x \oplus s_2$ after recovering x?

C-2.15 Passports are printed on special paper and have various anti-counterfeiting physical features. Develop a print-your-own passport pilot program where a passport is a digitally signed document that can be printed by the passport holder on standard paper. You can assume that border control checkpoints have the following hardware and software: two-dimensional barcode scanner, color monitor, cryptographic software, and the public keys of the passport-issuing authorities of all the countries participating in the pilot program. Describe the technology and analyze its security and usability. Is your system more or less secure than traditional passports?

C-2.16 Unlike passwords, biometric templates cannot be stored in hashed form, since the biometric reading does not have to match the template exactly. A ***fuzzy commitment*** method for a biometric

template can be developed from an error correcting code and a cryptographic hash function. Let f be the decoding function, h be the hash function, and w be a random codeword. A fuzzy commitment for template t is the pair $(h(w), \delta)$, where $\delta = t - w$. A reading t' is accepted as matching template t if $h(w') = h(w)$, where $w' = f(t' - \delta)$. Analyze the security and privacy properties of the scheme. In particular, show how this scheme protects the privacy of the template and accepts only readings close to the template (according to the error-correcting code).

Projects

P-2.1 Write a detailed comparison of the features of two high-security locks, Medeco M3 and Abloy. Discuss whether they are resilient to the attacks described in this chapter.

Java

P-2.2 Using the Java Card Development Kit, implement a vending card application that supports the following operations: add value to the card, pay for a purchase, and display the available balance. The vending card should authenticate and distinguish between two types of readers, those that can add value and those that can decrease value. Both readers can obtain balance information.

P-2.3 Design and implement a program simulating the main security functions of an ATM machine. In particular, your system should authenticate users based on a PIN and should transmit data to the bank in encrypted form. You should reduce the sensitive information stored by the ATM machine in between transactions to a minimum.

P-2.4 Write a term paper that discusses the different kinds of RFIDs, including both self-powered and not. Address privacy concerns raised by wide-spread RFID use, such as in e-passports. Use research articles available on the Internet as source material.

P-2.5 Using a conductive material, such as aluminum foil, construct a Faraday cage. Make it big enough to hold a cellphone or portable FM radio receiver and confirm that it blocks RF signals to such devices. Next, experiment with the size of holes in the exterior, to find the size of holes that allows RF signals to reach the device inside. Write a report documenting your construction and experiments. Include photographs if possible.

Chapter Notes

Jennie Rogers contributed material to Section 2.2 (Locks and Safes). Basics lock picking techniques are described in course notes by Matt Blaze [8] and in the" Guide to Lock Picking" by Ted Tool [102]. For more information on safe-cracking, refer to the paper "Safe-cracking for the Computer Scientist" by Matt Blaze [9]. The attack on the Medeco Biaxial system for locks is due to Tobias and Bluzmanis [101] The iterative master-key construction attack is presented by Matt Blaze [7]. The differential power analysis technique is described by Kocher, Jaffe and Jun [49]. Messerges, Dabbish and Sloan have done pioneering work on side-channels attacks on smart cards, showing that the RSA cryptosystem is vulnerable to differential power analysis [59]. An overview of GSM's encryption technology is provided in Jeremy Quirke's article "Security in the GSM System" [80]. Cloning techniques for GSM SIM cards based on side-channel attacks are presented by Rao, Rohatgi, Scherzer and Tinguely [81]. Several attacks have been demonstrated that completely compromise the KeeLoq and DST algorithms used in popular RFID devices [11, 19, 42, 67]. Jain, Ross and Prabhakar provide an overview of the subject of biometric recognition [43]. The collection of articles edited by Tuyls, Skoric and Kevenaar provides an advanced coverage of the subject of privacy protection for biometric authentication [104]. Di Crescenzo, Graveman, Ge and Arce propose a formal model and efficient constructions of approximate message authentication codes, with applications to private biometric authentication [25]. Wim van Eck pioneered the technique of eavesdropping on CRT displays by analyzing their radio frequency emissions [105]. Markus Kuhn has done notable work on eavesdropping techniques using radio frequency and optical emissions [50, 51, 52]. Adi Shamir and Eran Tromer have investigated acoustic cryptanalysis of CPUs [90]. Acoustic eavesdropping attacks on keyboards are discussed by Asonov and Agrawal [2] and by Zhuang, Zhou and Tygar [111]. A survey of results on acoustic eavesdropping is written by Adi Purwono [79]. Wright, Kleiman, and Shyaam debunk the myth of the possibility of data recovery after more than one pass of overwriting [110]. The cold boot attack to recover cryptographic keys from the RAM of a computer is due to Halderman et al. [38]. Electronic voting technologies and their risks are discussed in the book "Brave New Ballot" by Avi Rubin [85]. A security study by Feldman, Halderman and Felten found significant vulnerabilities in a Diebold voting machine [29].