6.857 Computer and Network Security: Lecture Notes (Draft)

Part 2 : Physical Security

Version: September 15, 1998

Physical security is the basis of computer security. By studying it we gain insight into a good metaphor for other types of security which will study later on. We are going to discuss the two main uses of physical security: protection against natural risks and against malicious attacks.

In this part we focus on attacks where the attacker "needs to get close" to mount an attack, rather than on penetration by remote login. Here are some reasons why someone might want to attack: financial gain, disgruntled employee seeking revenge, you are the first available target, to gain fame, diversion, military advantage, etc... It is important to note the difference between a targeted attack and a generic attack: the former requires more robust security measures.

1 Natural Risks

Next we will examine which natural risks are the most common threats to computer physical security. To see specific examples of a variety of things that can go wrong, and how things can go wrong simultaneously read *Computer-Related Risks* by Peter Neumann.

- **Power interruption**: Power interruption may be hard to recover from. Sometimes it may result in permanent system damage, since electronic components are not designed to absorb sudden power fluctuations. UPS (uninterrupted power supply), surge protectors, alternative power generators are possible security measures although many times their natural state is untested and/or broken.
  example: Neumann p. 16. AT&T backup battery drain and ultimately causing a 3 airport shut downs.
- **Fire**: Use of chemical fire supressants (ie. halon), smoke detectors and sprinklers are possible security measures. Chemical fire supressants like halon can be tricky to use because we want to be able to evacuate people easily, but we also want to seal the room to keep the halon contained. When using water sprinklers to put out fires we have to make sure that the water from the sprinklers will not wet the systems, as this may cause electrical fire and/or water damage (see below.)
- **Storm/Flood:** Lightning during storms can be very damaging (surge protectors usually don't protect against surges of these magnitudes). The increase of waterfall during storms can critically overload pipes and cause them to rupture. There are two distinctions to be made between raising waters and falling waters. The first can happen in case of flood for example. Systems should be placed in high stands and possibly in higher floors of buildings (not in the

basement.) Falling waters may happen because of leakage during heavy storms or because of fire alarm sprinklers. The easiest form of protection in this case is lots of plastic sheets. Others advocate the use of foam fire extinguishers instead of water ones. However, the previous security measures together with a backup system kept offsite would be ideal.
example: Neumann p. 56. Passengers were stranded in Sydney's monorail due to water damage to control computer.

- **Earthquake/Meteor/War**: These risks may destroy a whole building. In order to quickly restart the computer system first of all we need up-to-date backups (see section on backups below). Of course the backups should be kept in a separate place possibly miles away. In some cases a *cold site* may be needed, i.e. a facility with power, cooling and network connections (i.e., at least phone lines), available to install a computer system. For more critical task *hot sites* may be required. A hot site is a facility with a computer system ready to run and be used. The cost of maintaining a hot site is high but companies usually share the services of a single hot site provider (given the very rarity of use) making the cost fairly affordable.

- **Communications interruption**: Lines may be cut by people or animals. To protect against accidental severance of communication lines, duplicate lines running through different routes are needed.
example: Neumann p. 15. Chicago area telephone cable.
example: Neumann p. 16. Virginia cable cut.
example: Neumann p. 17. Beaver chews on cables.

- **Operator error**: Errors such as coffee or coke spills and accidental file deletion are very common. Errors such as mounting the wrong tapes or accidentally hitting the wrong key can also occur.
example: Neumann p. 77. Chernobyl accident, 1986.

- **Software/Hardware error**
example: Neumann p. 68. Therac-25 radiation therapy machine killed and injured several patients due to software bugs (ie. hardware safelocks were removed and substituted by software ones) and operator errors.

Backups: The most effective solution to many of these natural risks is to create backups. Many types of backups are possible– daily, weekly, and monthly, as well as full, incremental, and differential. Backups can be used either for quick restart or restoration of damaged systems, as well as for record keeping and tracking purposes. Backups should be separately stored in a secure place away from the system. Backups should be regularly tested to make sure that the data on the backups are intact. Also backups may contain sensitive information so they should be protected since they provide an attacker with a potential entry point to private data other than the computer system itself.

2 Physical access control

Walls and **fences** define the security perimeter, while **guards** and **locks** enforce access control policy. Only authorized users may enter the perimeter. Possible ways of violating physical access control include:

- **Masquerading**: A person disguised as an authorized user. This can be done using forged IDs or pretending to be a repair man.
- **Piggy-backing**: A person who enters the security perimeter by following an authorized user.
- **Lock-picking**: Any locked can be picked. Sometimes there is substantial savings over combinatorial brute-force. However, many times the easiest way is simply to go through dropped ceilings or removing the hinges from door.
  example: *MIT Lockpicking Guide*, by "Ted the Tool", which you can find with a web search...
- **Visual/auditory access**: An unauthorized person can gain passwords by bouncing laser off a window (or better, a small plastic cup inside the window) to pick up vibrations caused by sound waves. They can also use binoculars, telephone bugs, etc.
  example: Russians spied on Americans by installing a telephone near coderoom. They got the secret key by hearing the Selectric balls on typewriters.

Possible solutions to these attacks include:

- **Safes, doors with special locks** Locks are useless unless the keys are hard to fake or to obtain (ie. bribery).
  example: Hotel doors with special access keys based on a sequence of random numbers.
- **Double doors** Piggy-backing can be prevented by having a small entry-way with two doors; one going in and one going out. The entry-way is only big enough for one person, and the doors are controlled so that at most one door can be open at a time.
- **Motion detectors**
  example: the movie *Sneakers*.

If physical access control is defeated or impossible to maintain, then other measures may be necessary to control what an intruder can do, once "inside." For example, every vacation weekend, one or two computers or notebooks are stolen from the labs in Technology Square. To prevent theft, computers can be secured by locking the machine to a heavy or bulky object (such as a desk), locking the media readers (making it difficult or impossible to use), or locking the computer itself to prevent operation (such as a keyboard or a power supply lock).

3 Protecting Physical Media

Classic attacks on physical media include:

- **Dumpster Diving:** Disposed documents and printouts may yield many secrets such as passwords and confidential information.
- **Magnetic Storage Media:** Deleted information can often be undeleted.

Possible counter-measures include:

- **Clearing:** Overwrite the media many times with random bits. Writing to magnetic media can be thought of as spray painting a wall. To cover up old information, you need to "paint" over it several times.
- **Purging:** Using ultraviolet lights to destroy EEPROMs or using a strong magnet on magnetic storage.
- **Destruction:** Burning or shredding paper documents, grinding magnetic coating off disks, melting tapes, etc.

The idea of protecting media is to build fences. Physical protection is never perfect; you want to build the barriers high enough so that your attacker will choose an easier target.

3.1 Protecting Software

There are various schemes that attempt to defeat software piracy using physical properties, such as:

- **Serial Number/Manual Text Check** A printed number or table of numbers must be consulted to start the program.
- **CPU ID orNetwork Address Check** The program checks that it is running on the "right" computer.
- **Harware Dongles:** A special hardware device that attaches to the serial port. It contains circuitry or data that is needed to run the software.
- **Holograms:** Special packaging that prevents foreign duplicators from making "official-looking" copies.
- **Watermarking:** Each software package has a unique feature or generates a unique output. This is more to *detect* fraud, rather than to *prevent* it. Adi Shamir suggested having "soft bits" bits that are read erroneously with high probability. *example for printed matter:* Diddling postscript files so that each line is one pixel higher or lower than normal, to encode recipient's name.

Often the most effective copy protection is no protection at all. By lowering the price of the software or by providing excellent support services, many users are willing to buy their own copy of the software.

Defeated software protection mechanisms are often detected through increased support calls, decreased sales, and by finding anonymous postings of binaries on web sites.

4 Communications

The phone system is the biggest communications system around. Securing this medium has proven to be a difficult task. Watching how security on the phones has changed, in adaptation to the different attacks on the system, is a good way to study the evolution of security, in general. Phone booths and pay phones have evolved into highly impenetrable fortresses. They are designed against physical attacks and attempts to make money off of them as well as being difficult to vandalize.

Some terminology was brought up: phone freaks, computer hackers, and phracks. A phone freak is someone who uses phones illegally, a computer hacker is someone who breaks into computer systems illegally; the combination of the two is called a phrack.

We observed that phone system users depend on:

- Reliability

  Phone systems are expected to be working all the time. A January 15, 1990 incident was described where a bug in AT&T telephone switching software caused an outage for half of the United States. The bug was traced down to a case statement in some C code. According to *The Hacker Crackdown*, "Within the C software was a long 'do... while' construct. The 'do... while' construct contained a 'switch' statement. The 'switch' statement contained an 'if' clause. The 'if' clause contained a 'break.' The 'break' was supposed to 'break' the 'if clause.' Instead, the 'break' broke the 'switch' statement." This single bug caused all of the mayhem.

- Authenticity

  For instance, when one calls a bank phone number, one expects the bank to answer on the other side. But an example was given where phone phracks reprogrammed a Bell South telephone switch to divert the Palm Beach County Probation Department's incoming calls to a phone sex line.

  In general, one should not rely on the phone system for air-tight security or authentication.

Other examples of phone abuse were cited. In the mid 70's, devices called *blue-boxes* emulated the 2600Hz signal used by phone switches to make long distance calls. With this device, phone users could make free long distance calls. (You didn't have an excuse not to call home). The 2600Hz signal became the symbolistic name for the phreaks. The alt.2600 newsgroup and FAQ discuss vboxes, red-boxes, blue-boxes, and code for ethernet sniffers.

The cellular phone system is also vulnerable to attack. CNN reported that over 80,000 cell phone numbers were collected while unsuspecting motorists whizzed by on a busy Brooklyn highway. Thieves used a "black box" scanner to intercept numbers. See `http://cnn.com/US/9607/02/newsbriefs/index.html`.

Break-ins to company PBX's where also mentioned. Phracks were able to "bounce" off company PBX's for free long distance calls.

Items for warning of phone system abuse (eg, caller ID and callback modems) are also circumventable. Such devices are simply "higher fences" for intruders to jump over.

5 Tamper Resistance

The paper ("Tamper Resistance–a Cautionary Note") by Anderson and Kuhn documents the authors' experiences with breaking into "tamper-resistant" chips. In general, smartcards and other devices that rely on tamper-resistant hardware are vulnerable in many ways that are independent of the mathematics of cryptography. This becomes a real issue as smartcards gain acceptance, for example at a university or as part of a pay-TV service.

Cryptography requires secrets, such as the secret key held by a principal. If an attacker can discover the secret data by exploiting vulnerabilities in the implementation of a system, then as a rule the barriers imposed by hard-to-break cryptographic algorithms can be circumvented. Thus the maintenance of physical security over this secret data is an important part of a robust system.

Smartcards are a good example of a system that presents some interesting challenges in this area. In many cases, smart cards contain a secret encoded on a chip that implements a form of zero-knowledge proof. Smartcards are typically relatively readily available and are often distributed to a large user population. Typical applications include bank cards, debit cards, and access cards to buildings or pay-TV systems. If the embedded secret can be extracted, cards can be copied, allowing the holder of the duplicate to masquerade as a legitimate client, tenant, employee, etc. If the cards are distributed widely, it is impossible to maintain control over who has the opportunity to disassemble them, and "tamper-resistance" refers to the effort to make the secret difficult to recover from a card in the hands of an attacker.

What are the real risks? No security system is perfect, but in order to be cost-effective the security system must be secure enough to offset the potential costs, should the system be compromised. The credit card industry has performed this calculation in determining their policies and pricing structure. The level of security afforded by the mag stripe on a credit card is fairly low, but they have a number of other systems in place to identify fraudulent charges, and these mechanisms serve to keep their loss rate at an acceptable level.

The cell phone industry got into a great deal of trouble because their original protocol broadcasts plaintext ID information in the clear to the base station. It turned out to be very easy to copy an overheard ID number into another cell-phone, at which point the attacker's calls would be charged to the account associated with that ID. This practice is almost impossible to stop, and costs the cell-phone industry huge amounts of money in lost revenue and billing disputes. Newer protocols handle this ID information more securely.

The Anderson-Kuhn paper describes different ways of breaking into "tamper-resistant" chips. The overall message is that perfect tamper resistance is impossible. You must expect that someone will get past it, but a variety of techniques can make it more difficult and more expensive to extract the secret. The term "tamper-proof" seems to imply perfect tamper resistance, and is therefore not a good term.

The paper begins by defining three categories of attacker:

- Clever outsider: not well funded with incomplete knowledge of the devices
- Knowledgable insider: has access to detailed information about the design of the hardware
- Funded organizations: has significant funds, manpower, and lab equipment. May include insiders, and may be outside of legal reach.

When assessing the weaknesses of a particular system we assume that attackers can get unsupervised access to the devices themselves and any associated equipment. For example, in the case of a pay-TV access card, the attacker has several expired access cards, a non-expired card, and the set-top box that accepts them.

The paper describes several categories of attack in detail: non-invasive attacks, physical attacks, and advanced attacks.

5.1 Non-invasive Attacks

The typical smartcard consists of an 8-bit microprocessor, ROM, RAM, and an EEPROM, in a single chip. The secret key is stored in the EEPROM. The card has several inputs: clock, power, ground, reset, and several data lines. Non-invasive techniques try to exploit sensitivity to external and input conditions.

- Supply voltage: under-voltage, over-voltage, spikes
- Clock line: vary speeds, short clock cycles, etc.
- Reset line: trigger a reset and repeat an experiment
- Temperature: many components are temperature-sensitive

Often the EEPROM is protected by a security bit that is clear when the card comes from the factory and is set after the EEPROM is programmed. Once the security bit is set, it prevents access to the data in the EEPROM. Some microcontrollers can be tricked into clearing the security bit by raising the power to 0.5V less than the programming voltage. Another chip has an analog random number generator that consistently generates 1 values when given a low voltage.

In order to protect the cards from these kinds of attacks, card manufacturers have begun installing sensors that detect efforts to fiddle with the values and react by forcing a reset. However, in many cases the variations in actual environments tend to trigger these alarms and result in poor reliability.

One interesting attack uses power and clock transients to affect the execution of specific instructions, by adjusting the parameters of the transients to affect specific transistors. By causing specific instructions to be executed at specific times, a loop that outputs a specific memory range can be extended to dump the entire contents of memory. A countermeasure to this attack is the use of internal clock generators.

5.2 Physical Attacks

These attacks exploit physical characteristics of the device. Chips are composed of several layers of SiO_2 and deposited metals that form wires. Physical attacks often involve opening up the chip to figure out how it works or to detect specific voltages on the wires, and as a result many times involve physically damaging the device to get at the internals.

One simple physical attack involves focusing UV light on the part of the EEPROM containing the security bit, thus clearing it. A countermeasure to this attack is to include light detectors in the chip that disable the chip when triggered.

In many cases, a physical attack involves disassembling a chip to the point where probes can be attached to the components. The following series of steps works for many cards:

1. The plastic is cut away from the back of the chip with a knife.
2. Fuming nitric acid is used to remove the epoxy resin from the chip.
3. The passivation layer can be removed by probing with a needle that transmits ultrasonic vibrations, or with a laser. Probes can then be poked through the holes in the insulation and real-time values on the wires can be recovered.

5.3 Advanced Attacks

Advanced attackers often need to reverse-engineer the device in order to acheive their results. A recently developed technique cleanly etches away a single layer of a chip. By subsequently depositing a layer of gold and scanning the surface with an electron beam, the doping of various layers can be detected. By repeating this process for each layer in turn the entire contents of the chip can be derived. IBM used this technique to reverse-engineer the '386 in about 2 weeks, and recently used the same technique to reverse-engineer the Pentium.

Once the exact layout of the chip is known, the chip can be easily observed in operation. The voltage level of a specific feature on the chip can be read at a frequency of 25 MHz by shining a UV laser on a crystal of lithium niobate that is placed over the feature. This can be used to read the values of keys directly from the EEPROM output amplifiers.

In response to this attack, chip manufacturers use a "comformal glue" to bond to the surface of the silicon and make it harder to remove the glue without damaging the surface of the chip. The Clipper chip implemented some additional ideas in tamper-resistance. A Clipper chip is manufactured using an unclassified mask that includes fusable links. These links are configured individually for each chip to install a classified encryption algorithm and secret key.

Another attack involves shining an IR laser through the back of a chip. Using this technique it is possible to probe the values of specific transistors. Focused Ion Beam machines, which are used by semiconductor manufacturers, can be used to cut wires, add wires, and change the doping of layers without physically taking the chip apart.

5.4 Advanced Protection

Protection against sophisticated attackers is practically impossible. To acheive protection with high certainty it requires physical guarding of the device and self-destruct capabilites to prevent loss of a cryptographic key. While the security devices used to lock out nuclear weapons use explosive charges to protect the key, many commercial security devices rely on similar but less drastic measures. Charged capacitors are included in some chips that can erase the EEPROM when sensors detect tampering. These designs can sometimes be subverted by cooling the device so that the erase circuitry fails to work, or by chopping the chip up faster than the circuitry can completely erase the key.

5.5 The Dallas 5000 Microcontroller

The Dallas 5000 series secure microcontroller is a CPU designed to be used with an external bank of RAM. In order to maintain security, a secret key stored within the CPU is used to encrypt both the address and data lines using for encryption a function of both the key and the address data. Thus, when a memory address is written the data written there is encrypted, and data read in from RAM is decrypted before it is used by the processor.

However, an attack called "cipher-instruction search" can discover the encryption function using inexpensive hardware. The attack experiments with feeding different combinations of wrong data back to the CPU until a specific 3 byte sequence corresponding to `MOV 90h, #xxh` is discovered. This instruction causes the processor to output the decrypted value of the third byte to parallel port P1 (address 90h). Once such a sequence is discovered, the complete mapping for the address of the third byte can be discovered by trying each possible value at that location and executing the three byte sequence. A similar search technique, this time for the sequence `NOP MOV 90h #xx` can reveal the complete mapping for the next byte. After those two addresses are mapped, the rest of the addresses can be discovered without needing to guess a working sequence.

Fixing the Dallas chip can be done in a number of ways:

- use longer instructions
- add an onboard cache and perform fetches in 8 bytes blocks
- include MAC bytes for blocks of memory

5.6 Lessons

One of the foremost lessons to be learned from these attacks is to treat claims about smartcards with skepticism. The engineering that goes into a security system must take imperfect security into account by minimizing the damage caused by a single breach of security. Finally, hostile testing is very important to the development process.

6 Timing Attacks

Timing attacks are non-invasive attacks that rely on the variation in computation time required for the CPU on a smartcard to perform its secret calculation. Guarding against this attack is fairly easy; simply ensure that each response takes the same length of time. However, when this paper was released several smartcards were susceptible to this technique. See Kocher's paper.

6.1 Attacking Diffie-Hellman

As an example we will consider a card that implements a zero-knowledge proof based on Diffie-Hellman. In this case there is a secret value x, and the card computes $y^x \bmod n$, given y. Typically a smartcard uses the standard approach to modular exponentiation, as shown previously in class. Essentially, a result register is initialized to 1 and for each bit b_i of x, if $b_i = 1$ then the result register R is multiplied by $y_i = y^{2^i} \bmod n$, where y_i is computed incrementally as $y_i = {y_{i-1}}^2$.

The attack is simplest to understand if there are some known unusual cases in which the multiplication algorithm is very slow. Suppose the attacker has knowledge of bits b_0 through b_{i-1} of the secret exponent, the value of y, and has knowledge of certain values for which multiplication is slow. Then the attacker can compute the result of the exponentiation of the first $i-1$ bits of x. If the attacker now chooses y such that the multiplication of the $(i-1)^{st}$ result with y^{2^i} is very slow, then the i^{th} bit is set only if the total calculation time is much greater than the time required to compute the first $i-1$ bits.

However, wide variations in timing are not required for this attack to succeed. In fact, using many trials with random y values, even slight variations in timing can be exploited in this way to detect correlations that reveal the bits in the secret exponent. Because these correlations disappear after a single bit is chosen incorrectly, errors are quickly corrected. Extracting the bits of an exponent may require as few as 2000 trials.

6.2 Factoring RSA Keys

Timing attacks can also be used to factor RSA keys. If the Chinese Remainder Theorem is not used, a similar attack to the Diffie-Hellman attack given above can be used. If the CRT is used to perform the exponentiation, then a slightly different technique must be used.

In the RSA cryptosystem, using CRT, the public key n is formed from two secret primes p and q, and the card performs the exponentiation $C^d \bmod n$, where C is a cyphertext chosen by the attacker. Using the CRT, the card performs the following calculations:

$$\begin{aligned}
&C_1 = C \bmod p\\
&C_2 = C \bmod q\\
&M_1 = {C_1}^{d_1} \bmod p\\
&M_2 = {C_2}^{d_2} \bmod q\\
&M = M_1 a_1 + M_2 a_2 \bmod n
\end{aligned}$$

Because the modular reduction operations performed in the first two steps do not run in constant time, the attacker can tell if a chosen C value is greater or less than p or q. If $C < \min(p, q)$, then those operations will take no time, but otherwise p or q will need to be subtracted from C at least once.

7 Active Attack on RSA Smart Card

The timing attacks above show that there are ways of completely circumventing theoretical difficulties to break cryptographic systems. While the timing attacks were passive, here we see an active attack on some RSA Smart Cards.

In practice, RSA public exponents are chosen to be small (3, 17, etc...), which makes encryption and signature verification reasonably fast. Private exponents, however, should never be small for obvious security reasons. This makes decryption slow. One way to speed things up is to split up calculations modulo p and modulo q using the Chinese Remainder Theorem. For smart cards which usually have limited computing power, this is a very useful technique.

The steps of the CRT computation are as follows:

$$c = M^e \pmod{n}$$

$$\begin{array}{ll} c_1 = c \pmod{p} & c_2 = c \pmod{q} \\ d_1 = d \pmod{p-1} & d_2 = d \pmod{q-1} \\ m_1 = c_1^{d_1} \pmod{p} & m_2 = c_2^{d_2} \pmod{q} \\ a_1 = q^{-1} \pmod{p} & a_2 = p^{-1} \pmod{q} \end{array}$$

$$m = a_1 * q * m_1 + a_2 * p * m_2 \pmod{n}$$

The danger that the card exposes here is precisely in performing separate calculations for p and q. An active attack consisting simply in a power spike during the calculation of one of the two m's is enough to reveal n's factorization.

Suppose that a power spike or other disturbance causes a change in the calculation of m_2, such that it now becomes m_2'.

The attacker can then take a random plaintext m, encrypt it to obtain c, then pass c through the the smart card, providing the disturbance to obtain:

$$m' = a_1 * q * m_1 + a_2 * p * m_2' \pmod{n}$$

$$m - m' = a_2 * p * (m_2 - m_2') \pmod{n}$$

Thus,

$$gcd(m - m', n) = p$$

and the attacker obtains n's factorization, which easily yields the RSA secret exponent d.

One possible fix for this is to simply have the smart card check that the output it is about to produce still encrypts to the proper ciphertext it was fed in the beginning. If there is a difference, the card can simply refuse to output the result, thereby securing its secret key. Another fix is simply to perform the operation twice, and yielding an output when the two results are the same.

8 Differential Power Analysis

Closely related to timing attacks, and very effective. See the paper by Paul Kocher.

9 Electromagnetic Radiations – TEMPEST

It is possible, with simple equipment, to read what is being displayed on a fairly distant CRT. The paper by Kuhn and Anderson gives an excellent discussion, including specialized design of screen fonts to minimize intelligible emanations.

10 References

Anderson, Ross, and Markus Kuhn, "Tamper Resistance–a Cautionary Note," Proceedings of the Second USENIX Workshop on Electronic Commerce (1996), 1–11. Also available online at `http://www.cl.cam.ac.uk/users/rja14/`.

Kocher, Paul. *Differential Power Analysis.* `http://www.cryptography.com/dpa/index.html`.

Kocher, Paul. *Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems.* `http://www.cryptography.com/timingattack/paper.html`.

Kuhn, Markus G. and Ross J. Anderson, *Soft Tempest: Hidden Data Transmission Using Electromagnetic Emanations*, Second Workshop on Information Hiding, 1998. Also available online at `http://www.cl.cam.ac.uk/users/rja14/`.

Cowcher, Randle. *Physical Security.* Chapter 24 of *Computer Security Reference Book*, edited by K.M. Jackson, J. Hruska, and Donn B. Parker (Butterworth Heinemann, 1992).

Neumann, Peter G. *Computer Related Risks* (ACM Press, 1995).