CHAPTER

11

Nuclear Command and Control

In Germany and Turkey they viewed scenes that were particularly distressing. On the runway stood a German (or Turkish) quick-reaction alert airplane loaded with nuclear weapons and with a foreign pilot in the cockpit. The airplane was ready to take off at the earliest warning, and the nuclear weapons were fully operational. The only evidence of U.S. control was a lonely 18-year-old sentry armed with a carbine and standing on the tarmac. When the sentry at the German airfield was asked how he intended to maintain control of the nuclear weapons should the pilot suddenly decide to scramble (either through personal caprice or through an order from the German command circumventing U.S. command), the sentry replied that he would shoot the pilot; Agnew directed him to shoot the bomb.

—JEROME WIESNER, PRESIDENTIAL SCIENCE ADVISOR, REPORTING TO PRESIDENT KENNEDY ON NUCLEAR COMMAND AND CONTROL AFTER THE CUBAN CRISIS

11.1 Introduction

The uniquely catasrophic harm that could result from the unauthorized use of a nuclear weapon, or from the proliferation of nuclear technology to unsuitable states or substate groups, has led the United States (and other nuclear states) to spend colossal amounts of money protecting not just nuclear warheads but also the supporting infrastructure, industry, and materials.

Quite a lot of nuclear security know-how has been published. In fact, there are severe limits on how much could be kept secret, even if this was thought desirable. Many countries are capable of producing nuclear weapons, but have decided not to (Japan, Australia, Switzerland, . . .) and so maintain controls on nuclear materials in a civilian

context. There are also international nonproliferation agreements, such as the Convention on the Physical Protection of Nuclear Material [409], enforced by the International Atomic Energy Agency (IAEA).

Eleven tons of plutonium are produced annually by civil reactors. So ways have to be found to guard the stuff, and these have to inspire international confidence—not just between governments but from an increasingly sceptical public.

A vast range of security technology has spun off from the nuclear program. The U.S. Department of Energy weapons laboratories—Sandia, Lawrence Livermore, and Los Alamos—have worked, with almost unlimited budgets, for two generations to make nuclear weapons and materials as safe as can be achieved. We've already seen some of their more pedestrian spin-offs, from the discovery that passwords of more than twelve digits were not usable under battlefield conditions to high-end burglar alarm systems. The trick of wrapping an optical fiber around the devices to be protected and using interference effects to detect a change in length of less than a micron is another of theirs. It was designed to loop around the warheads in an armory and alarm without fail if any of them are moved.

In later chapters, we'll see still more technology of nuclear origin. For example, iris recognition—the most accurate system known for biometric identification of individuals—was developed using U.S. Department of Energy funds with a view to controlling entry to the plutonium store; and much of the expertise in tamper-resistance and tamper-sensing technology originally evolved to prevent the abuse of stolen weapons or control devices.

In this chapter, I describe the environment in which these technologies were developed, and some of the tricks that might find applications—or pose threats—elsewhere. As I'm not an insider, I've assembled this chapter from public sources, and so may have missed important points (a proofreader with the relevant clearance and experience assures me that the material is indeed "accurate but incomplete"). Nevertheless, even from the available material, there are useful lessons to be learned.

11.2 The Kennedy Memorandum

Following the Cuban missile crisis, the U.S. government became concerned that a world war could start by accident. Hundreds of U.S. nuclear weapons were kept in allied countries such as Greece and Turkey, which were not particularly stable and occasionally fought with each other. These weapons were protected only by token U.S. custodial forces; there was no physical reason why the weapons couldn't be seized in time of crisis. There was also some concern about possible unauthorized use of nuclear weapons by U.S. commanders; for example, if a local commander under pressure felt that, "If only they knew in Washington how bad things were here, they would let us use the bomb." These worries were confirmed by three emergency studies carried out by presidential science adviser Jerome Wiesner. (The passage quoted at the beginning of this chapter can be found in [734].)

President Kennedy's response was National Security Action Memo number 160. This ordered that America's 7,000 nuclear weapons then in other countries should be got under positive control, or got out [705].

The Department of Energy was already working on safety devices for nuclear weapons. The basic principle was that one or more unique aspects of the environment had to be sensed before the weapon would arm. For example, missile warheads and some free-fall bombs had to experience zero gravity, while artillery shells had to experience an acceleration of thousands of G. There was one exception: atomic demolition munitions. These are designed to be taken to their targets by ground troops and detonated using time fuses. There appears to be no scope for a unique environmental sensor to prevent their accidental or malicious detonation.

The solution then under development was a secret arming code, which activated a solenoid safe lock buried deep in the plutonium pit at the heart of the weapon. The main engineering problem was maintenance. When the lock was exposed—for example, to replace the power supply—the code might become known. Clearly, it was not acceptable to have the same code in every weapon. Group codes had to be used—firing codes shared by only a small batch of warheads.

Following the Kennedy memo, it was proposed that all nuclear bombs should be protected using code locks, and that there should be a "universal unlock" action message that only the president or his legal successors could send. The problem was to find a way to translate this code securely to a large number of individual firing codes, each of which enabled a small batch of weapons. The problem became worse in the 1960s and 1970s, when the doctrine changed from massive retaliation to "measured response." Instead of arming all nuclear weapons or none, the president now had to be able to arm selected batches (such as "all nuclear artillery in Germany").

11.3 Unconditionally Secure Authentication Codes

This requirement led to the development of a theory of one-time authentication codes. These are similar in concept to the test keys invented to protect telegraphic money transfers, in that a keyed transformation is applied to the message to yield a short authentication code, also known as an *authenticator* or *tag.* As the keys are only used once, authentication codes can be made unconditionally secure. They do for authentication what the one-time pad does for confidentiality.

Recall from Chapter 5, "Cryptography," that while the perfect security provided by the one-time pad is independent of the computational resources available to the attacker, a computationally secure system could be broken by some known computation, and depends on this being infeasible.

There are differences, though, between authentication codes and the one-time pad. As the authentication code is of finite length, it's always possible for the opponent to guess it; and the probability of a successful guess might be different depending on whether the opponent is trying to guess a valid message from scratch (*impersonation*) or to modify an existing valid message to get another one (*substitution*).

An example should make this clear. Let's assume that a commander has agreed to an authentication scheme with a subordinate under which an instruction is to be encoded as a three-digit number from 000 to 999. The instruction may have two values: "Attack Russia" and "Attack China." One of these will be encoded as an even number, and the other by an odd number; which is which will be part of the secret key. The authenticity of the message will be vouched for by making its remainder, when divided by 337, equal to a secret number that is the second part of the key.

Now suppose the key is that:

- "Attack Russia" codes to even numbers, and "Attack China" to odd.
- An authentic message is one that has the remainder 12 when divided by 337.

Therefore, "Attack Russia" is 686 (or 12) and "Attack China" is 349.

An enemy who has taken over the communications channel between the commander and the subordinate, and who knows the scheme but not the key, has a probability of only 1 in 337 of successfully impersonating the commander. However, once he sees a valid message (say, 12 for "Attack Russia"), then he can easily change it to the other by adding 337. Then (provided he understood what it meant), he can send the missiles to the other country. Thus, the probability of a successful substitution attack in this case is 1.

As with computationally secure authentication, the unconditional variety can provide message secrecy or not: it might work like a block cipher, or like a MAC on a plaintext message. Similarly, it can use an arbitrator or not. One might even want multiple arbitrators, so that they don't have to be trusted individually. If the first arbitrator wrongfully finds against the cheated party, then his victim should be able to denounce him.

Schemes may combine unconditional with computational security. For example, an unconditional code without secrecy could have computationally secure secrecy added simply by enciphering the message and the authenticator using a conventional cipher system.

Authentication is, in some sense, the dual of coding. In the latter, given an incorrect message, we want to find the nearest correct message efficiently; in the former, we want finding a correct message to be impossible unless you've seen it already or are authorized to construct it. And just as the designer of an error-correcting code wants the shortest length of code for a given error recovery capability, so the designer of an authentication code wants to minimize the key length required to achieve a given bound on the deception probabilities.

The authentication terminology used in civil and military applications is slightly different [703]. More importantly, the threat models are different. Soldiers are, in general, more concerned about enemies than traitors and are not so worried about nonrepudiation (except when enforcing treaties with other countries, which might later repudiate a message claiming that the key had been leaked by a "defector"). In business, the majority of frauds are carried out by insiders, so shared control systems are the main issue when designing authentication mechanisms.

11.4 Shared Control Schemes

The nuclear command and control business became even more complex with the concern, from the late 1970s, that a Soviet decapitation strike against the U.S. *national command authority* (i.e., the President and his lawful successors in office) might leave the arsenal intact but useless. There was also concern that, past a certain threshold of readiness, it wasn't sensible to assume that communications between authority and field commanders could be maintained, because of the possible effects of electromagnetic pulse and other attacks on communications. The solution was found in another branch of cryptomathematics known as *secret sharing*, whose development it helped to

inspire. The idea is that, in time of tension, a backup control system will be activated, whereby combinations of office holders or field commanders can jointly allow a weapon to be armed. Otherwise, the problems of maintaining detailed central control of a large number of weapons would likely become insoluble.

There is a simple and obvious way to do shared control: just give half of the authentication key to each of two people. This has the drawback that we need twice the length of key, assuming that the original security parameter must apply even if one of them is suborned. A better approach is to give each of them a number and have the two numbers add up to the key. This is how keys for automatic teller machines are managed.

However, this still may not be enough in command applications, as no one can be sure that the personnel operating the equipment will consent, without discussion or query, to unleash Armageddon.

A more general approach was invented independently by Blakley and Shamir in 1979 [111, 692]. Their basic idea is illustrated in Figure 11.1. Suppose the rule Britain wants to enforce, if the Prime Minister is assassinated, is that a weapon can be armed by any two cabinet ministers, or by any three generals, or by a cabinet minister and two generals. Let the point *C* on the *z* axis be the unlock code that has to be supplied to the weapon. We now draw a line at random through *C*, and give each cabinet minister a random point on the line. Now any two of them can together work out the coordinates of the line and find the point *C* where it meets the *z* axis. Similarly, we embed the line in a random plane and give each general a random point on the plane. Now any three generals, or two generals plus a minister, can reconstruct the plane and thence the firing code *C*.

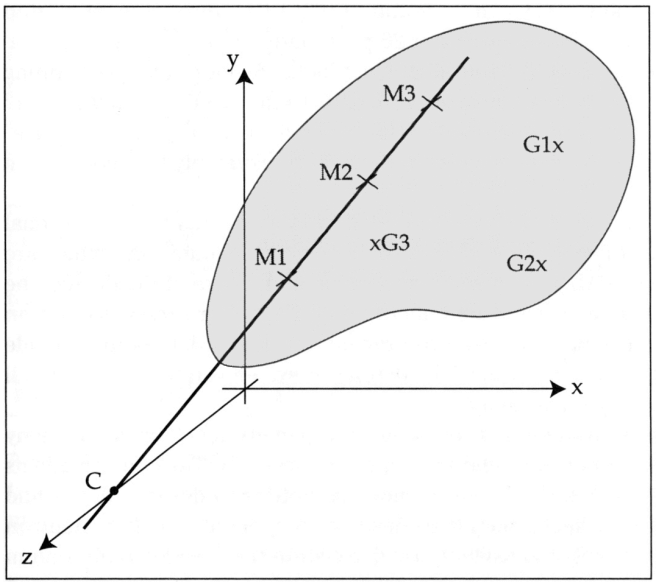

Figure 11.1 Shared control using geometry.

By generalizing this simple construction to geometries of *n* dimensions, or to general algebraic structures rather than lines and planes, this technique enables weapons, commanders, and options to be linked with a complexity limited only by the available bandwidth. (An introduction to secret sharing can be found in [738], and a more detailed exposition in [704].) Secret sharing also inspired the development of threshold signature schemes, which I described in Chapter 5, and can be used in products that enforce a rule such as, "Any two vice presidents of the company may sign a check."

As with authentication codes, there is a difference between civil and military views of shared secrets. In the typical military application, two-out-of-*n* control is used; *n* must be large enough that at least two of the keyholders will be ready and able to do the job, despite combat losses. Many details need attention. For example, the death of a commander shouldn't enable his deputy to use both halves of the key. So they typically have to be used simultaneously at consoles several yards apart.

In many civilian applications, however, many insiders may conspire to break a system. The classic example is pay-TV, where a pirate may buy several dozen subscriber cards and reverse engineer them for their secrets. Obviously, the pay-TV operator wants a system that's robust against multiple compromised subscribers.

11.5 Tamper Resistance and PALs

In modern weapons, the solenoid safe locks have been superseded by *prescribed action links*, more recently renamed permissive action links (either way, PALs), which are used to protect most U.S. nuclear devices. (A summary of the open source information about PALs can be found in [92].) PAL development started in about 1961, but deployment was slow. Even 20 years later, about half the U.S. nuclear warheads in Europe still used the four-digit code locks. As more complex arming options were introduced, the codes increased in length from 4 to 6 and finally to 12 digits. Devices started to have multiple codes, with separate "enable" and "authorize" commands, and the capability to change codes in the field (presumably to recover from false alarms).

The PAL system is supplemented by various coded switch systems and operational procedures; and in the case of weapons such as atomic demolition munitions, which are not complex enough for the PAL to be made inaccessible in the core of the device, the weapon is also stored in tamper-sensing containers called PAPS (for *prescribed action protective system*). Other mechanisms used to prevent accidental detonation include the deliberate weakening of critical parts of the detonator system, so that they will fail if exposed to certain abnormal environments.

Whatever combination of systems is used, there are penalty mechanisms to deny a thief the ability to obtain a nuclear yield from a stolen weapon. These mechanisms vary from one weapon type to another, but include gas bottles to deform the pit and hydride the plutonium in it; shaped charges to destroy components, such as neutron generators and the tritium boost; and asymmetric detonation that results in plutonium dispersal rather than yield. It is always a priority to destroy the code. It is assumed that a renegade government prepared to deploy "terrorists" to steal a shipment of bombs would be prepared to sacrifice some of the bombs (and some technical personnel) to obtain a single serviceable weapon.

To perform authorized maintenance, the tamper protection must be disabled, and this requires a separate unlock code. The devices that hold the various unlock codes—for servicing and firing—are themselves protected in similar ways to the weapons.

The protection goal is summarized in [734]:

> *It is currently believed that even someone who gained possession of such a weapon, had a set of drawings, and enjoyed the technical capability of one of the national laboratories would be unable to successfully cause a detonation without knowing the code.*

Meeting such an ambitious goal requires a very substantial effort. There are several examples of the level of care needed:

- after tests showed that 1 mm chip fragments survived the protective detonation of a control device carried aboard airborne command posts, the software was rewritten so that all key material was stored as two separate components, which were kept at addresses more than 1 mm apart on the chip surface;
- the "football," the command device carried around behind the president, is said to be as thick as it is out of fear that shaped charges might be used to disable its protective mechanisms. (This may or may not be an urban myth.) Shaped charges can generate a plasma jet with a velocity of 8000 m/s, which could, in theory, be used to disable tamper-sensing circuitry. So some distance may be needed to give the alarm circuit enough time to zeroize the code memory.

This care must extend to many details of implementation and operation. The weapons-testing process includes not just independent verification and validation, but hostile "black hat" penetration attempts by competing laboratories or agencies. Even then, all practical measures are taken to prevent access by possible opponents. The devices (both munition and control) are defended in depth by armed forces; there are frequent zeronotice challenge inspections; and staff may be made to retake the relevant examinations at any time of the day or night.

I discuss tamper resistance in much more detail in a later chapter, as it's becoming rather widely used in applications from pay-TV to bank cards. However, tamper resistance, secret sharing, and one-time authenticators aren't the only technologies to have benefitted from the nuclear industry's interest. There are more subtle system lessons too.

11.6 Treaty Verification

A variety of verification systems are used to monitor compliance with nonproliferation treaties. For example, the IAEA and the U.S. Nuclear Regulatory Commission (NRC) monitor fissile materials in licensed civilian power reactors and other facilities.

An interesting example comes from the tamper-resistant seismic sensor devices designed to monitor the Comprehensive Test Ban Treaty [701]. The goal was to emplace sufficiently sensitive sensors in each signatory's test sites so that any violation of the treaty (such as by testing too large a device) could be detected with high probability. The tamper sensing here is fairly straightforward: the seismic sensors are fitted in a steel tube and inserted into a drill hole that is backfilled with concrete. The whole as-

sembly is so solid that the seismometers themselves can be relied upon to detect tampering events with a fairly high probability. This physical protection is reinforced by random challenge inspections.

The authentication process becomes somewhat more complex because one has to make an assumption of pervasive deceit. Because of the lack of a third party trusted by both sides, and because the quantity of seismic data being transmitted is of the order of 10^8 bits per day, a digital signature scheme (RSA) was used instead of one-time authentication tags. But this is only part of the answer. One party might, for example, disavow a signed message by saying that the official responsible for generating the key had defected, and so the signature was forged. So keys must be generated within the seismic package itself once it has been sealed by both sides. Also, if one side builds the equipment, the other will suspect it of having hidden functionality. Several protocols were proposed of the *cut-and-choose* variety, whereby one party would produce several devices of which the other party would dismantle a sample for inspection. A number of these issues have since resurfaced in electronic commerce. (Many e-commerce system builders should have paid more attention to the lessons in [701].)

11.7 What Goes Wrong

Despite the huge amounts of money invested in developing high-tech protection mechanisms, nuclear control and safety systems appear to suffer from the same kind of design bugs, implementation blunders, and careless operations as any others.

Recently, Britain's main waste reprocessing plant at Sellafield, which handles plutonium in multiple-ton quantities, has been plagued with a series of scandals. Waste documentation has been forged; radiation leaks have been covered up; workers altered entry passes so they could bring their cars into restricted areas; and there have been reports of sabotage. The nuclear police force only managed to clear up 17 out of 158 thefts, and 3 out of 20 cases of criminal damage [495]. It now looks as if the facility will be closed following loss of confidence by customers. The situation in the former Soviet Union appears to be very much worse. A recent survey of nuclear safekeeping describes how dilapidated security mechanisms have become in the decade following the collapse of the USSR, with fissile materials occasionally appearing on the black market and whistleblowers being prosecuted [401].

There are also a number of problems relating to the reliability of communications and other systems under attack. How can communication between the president and many sites around the world be assured? I'll discuss these problems in Chapter 16, "Electronic and Information Warfare."

There have also been a number of interesting high-tech security failures. One example is a possible attack, which led to the development of a new branch of cryptomathematics—the study of subliminal channels—which is relevant to later discussions on copyright marking and steganography.

The story of the invention of subliminal channels is told in [707]. During the Carter administration, the United States proposed a deal with the Soviet Union under which each side would cooperate with the other to verify the number of intercontinental ballistic missiles. At the same time, in order to protect U.S. Minuteman missiles against a possible Soviet first strike, it was proposed that 100 missiles be moved randomly around a field of 1,000 silos by giant trucks, which were designed so that observers couldn't determine whether they were moving a missile or not. The Soviets would have had to destroy all 1,000 silos to make a successful first strike; and in the context of the proposed arms controls this was thought impractical.

This raised the interesting problem of how to assure the Soviets that there were at most 100 missiles in the silo field, but without letting them find out which silos were occupied. The proposed solution was that the silos would have a Russian sensor package that would detect the presence or absence of a missile, sign this single bit of information, and send it via a U.S. monitoring facility to Moscow. The sensors would be packaged and randomly shuffled by the Americans before emplacement, so that the Russians could not correlate "full" or "empty" signals with particular silos. The catch was that only this single bit of information could be sent; if the Russians could smuggle any more information into the message, they could quickly locate the full silos—as it would take only 10 bits of address information to specify a single silo in the field. (There were many other security requirements to prevent either side cheating, or falsely accusing the other of cheating: for more details, see [706].)

To understand how subliminal channels work, consider the Digital Signature Algorithm described in Chapter 5. The systemwide values are a prime number p, a 160-bit prime number q dividing $p–1$, and a generator g of a subgroup of F^*_p of order q. The signature on the message M is r, s where $r = (g^k \text{ (modulo } p)) \text{ (modulo } q)$, and k is a random session key. The mapping from k to r is fairly random, so a signer who wishes to hide 10 bits of information in this signature for covert transmission to an accomplice can first agree how the bits will be hidden (such as "bits 72–81") and, second, try out one value of k after another until the resulting value r has the desired value in the agreed place.

This could have caused a disastrous failure of the security protocol, as there had been an agreement that the monitoring messages would be authenticated first with a Russian scheme, using Russian equipment, then by an American scheme using American equipment. Had the Russians specified a signature scheme like DSA, they could have leaked the location of the occupied silos and acquired the capability to make a first strike against the Minuteman force.

In the end, the "missile shell game," as it had become known in the popular press, wasn't used. The cooling of relations following the 1980 election put things on hold. Eventually, with the Medium Range Ballistic Missile Treaty, statistical methods were used. The Russians could say, "We'd like to look at the following 20 silos," and they would be uncapped for their satellites to take a look. Since the end of the Cold War, inspections have become much more intimate with inspection flights in manned aircraft carrying observers from both sides, rather than satellites.

Still, the discovery of subliminal channels was significant. Ways in which they might be abused include putting HIV status, or the fact of a felony conviction, into a next-generation digital identity card. Where this is unacceptable, and the card issuer isn't sufficiently trusted not to do it, the remedy is to use a completely deterministic signature scheme such as RSA instead of one that uses a random session key like DSA.

11.8 Secrecy or Openness?

Finally, the nuclear industry provides a nice case history of secrecy. In the 1930s, physicists from many countries had freely shared the scientific ideas that led to the development of the bomb; but after the "atomic spies" (Fuchs, the Rosenbergs, and others) had leaked the designs of the Hiroshima and Nagasaki devices to the Soviet Union, things swung to the other extreme. The United States adopted a policy that atomic knowledge was *born classified*. That meant that if you were within U.S. jurisdiction and had an idea relevant to nuclear weapons, you had to keep it secret regardless of whether you held a security clearance or even worked in the nuclear industry. This was clearly in tension with the Constitution. Things have greatly relaxed since then, as the protection issues were thought through in detail.

"We've a database in New Mexico that records the physical and chemical properties of plutonium at very high temperatures and pressures," a former head of U.S. nuclear security once told me. "At what level should I classify that? Who's going to steal it, and will it do them any good? The Russians, they've got that data for themselves. The Israelis can figure it out. Gaddafi? What the hell will he do with it?"

As issues like this got worked though, a surprising amount of the technology has been declassified and sometimes published, at least in outline. Starting from early publication at scientific conferences of results on authentication codes and subliminal channels in the early 1980s, the benefits of public design review have been found to outweigh the possible advantage to an opponent of knowing broadly the system in use.

Many implementation details are kept secret, though; information that could facilitate sabotage, such as which of a facility's 50 buildings contains the alarm response force, gets marked *unclassified controlled nuclear information* (UCNI), adding yet another layer of complexity to the security policy model. There are also numerous nitty-gritty issues, such as who is authorized to shoot whom (on the same side) and under what circumstances.

Still, the big picture is open (or so we're assured); and even before the recent classification reviews, command and control technologies were explicitly offered to other states, including hostile ones like the former Soviet Union. Again, the benefits of reducing the likelihood of an accidental war were considered to outweigh the possible benefits of secrecy. This is a modern reincarnation of Kerckhoffs' doctrine, first put forward in the nineteenth century, that the security of a system should depend on its key, not on its design remaining obscure [454].

11.9 Summary

The command and control of nuclear weapons, and subsidiary activities—from protecting the integrity of the national command system through physical security of nuclear facilities to monitoring international arms control treaties—has made a disproportionate contribution to the development of security technology.

The quite rational decision that the relevant assets had to be protected almost regardless of the cost drove the development of a lot of mathematics and science that has found application elsewhere. The particular examples given in this chapter are authen-

tication codes, shared control schemes, and subliminal channels. We also started to discuss tamper-resistant devices, about which I'll have more to say later.

Research Problem

Find interesting applications for technologies developed in this area, such as authentication codes.

Further Reading

Simmons was a pioneer of authentication codes, shared control schemes, and subliminal channels. His book [703] remains the best reference for most of the technical material discussed in this chapter. A more concise introduction to both authentication and secret sharing can be found in [738].

One of the best open sources for public information on nuclear weapons is the Federation of American Scientists [286]. The rationale for the recent declassification of many nuclear arms technologies is presented in detail on their website [286]. Declassification issues are discussed in [812], and the publicly available material on PALs has been assembled by Steve Bellovin [92].

Control failures in nuclear installations are documented in a range of places. The problems with Russian installations are discussed in [401]; U.S. nuclear safety is overseen by the Nuclear Regulatory Commission [593]; and shortcomings with U.K. installations are documented in the quarterly reports posted by the Health and Safety Executive [375].