CHAPTER
13

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, reporting to President Kennedy on nuclear arms command and control after the cuban crisis

13.1 Introduction

The catastrophic 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 U.S. and other nuclear powers to spend colossal amounts of money protecting not just nuclear warheads but also the supporting infrastructure, industry and materials. The growing concern about global warming makes nuclear protection all the more critical: how do we build new nuclear power stations without greatly increasing the risk that bad people get hold of weapons or fissile materials?

A surprising amount of nuclear security know-how has been published. In fact, severe limits have been placed 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. Much of the real force of nonproliferation is cultural, built over the years through diplomacy and through the restraint of nuclear powers who since 1945 forebore use of these weapons even when facing defeat at the hands of non-nuclear states. The culture is backed by international nonproliferation agreements, such as the Convention on the Physical Protection of Nuclear Material [640], enforced by the International Atomic Energy Agency (IAEA).

Eleven tons of plutonium are produced by civil reactors each year, and if the human race is to rely on nuclear power long-term then we'll be burning it in reactors as well as just making it as a side-effect of burning uranium. 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[1].

So 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 for two generations to make nuclear weapons and materials as safe as can be achieved, using almost unlimited budgets. I've already mentioned 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 round the devices to be protected and using interference effects to detect a change in length of less than a micron, is also one of theirs — it was designed to loop round the warheads in an armoury 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 to control 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. The increased tension since 9/11 has led to further spread of controls, especially once it was realised that for terrorist purposes it isn't necessary to get fissile materials like plutonium or uranium-235. A 'dirty bomb' — a device that would disperse radioactive material over a city block — is also a real threat, and one that jihadists have talked about. It might not kill anyone but it could lead to panic, and in a financial center it could cause great economic damage. For example, in March 2007, GAO investigators set up a bogus company and got a license from the Nuclear Regulatory Commission authorizing them to buy isotopes with which they could have built such a radiological dispersion device. What's

[1]For example, the British government was seriously embarrassed in 2007 when its safety arrangements for its 100-ton plutonium stockpile were criticised by eniment scientists [1089].

more, the license was printed on ordinary paper; the investigators altered it to change the quantity of material they were allowed to buy, then used it to order dozens of moisture density gauges containing americium-241 and cesium-137 [757]. This incident suggests that materials control may spread quite widely in the economy, and it may involve the wider deployment of many of the technologies described in this book.

Nuclear safety continually teaches us lessons about the limits of assurance. For example, it's tempting to assume that if a certain action that you don't want to happen has a probability of 1 in 10 of happening through human error, then by getting five different people to check, you can reduce the probability to 1 in 100,000. The U.S. Air Force thought so too. Yet in October 2007, six U.S. hydrogen bombs went missing for 36 hours after a plane taking cruise missiles from Minot Air Force Base in North Dakota to Barksdale in Louisiana was mistakenly loaded with six missiles armed with live warheads. This was supposed to be prevented by the handlers inspecting all the missiles in the storage area and checking them against a schedule (which was out of date), by ground crew waiting for the inspection to finish before moving any missiles, (they didn't), by ground crew inspecting the missiles (they didn't look in little glass portholes to see whether the warheads were real or dummy), by the driver calling in the identification numbers to a control center (nobody there bothered to check), and finally by the navigator during his preflight check (he didn't look at the wing with the live missiles). The plane took off, flew to Louisiana, landed, and sat unguarded on the runway for nine hours before the ground crew there arrived to unload the missiles and discovered they were live [127, 380]. This illustrates one of the limits to shared control. People will rely on others and slack off — a lesson also known in the world of medical safety. Indeed, in the USAF case it turned out that the airmen had replaced the official procedures with an 'informal' schedule of their own. So how can you design systems that don't fail in this way?

In this chapter I'm going to describe the nuclear safety environment and some of the tricks that might still find applications (or pose threats) elsewhere. This chapter has been assembled from public sources — but even from the available material there are useful lessons to be drawn.

13.2 The Evolution of Command and Control

The first atomic bomb to be used in combat was the 'Little Boy' dropped on Hiroshima. It came with three detonators, and the weapons officer was supposed to replace green dummy ones with red live ones once the plane was airborne. However, a number of heavily loaded B-29s had crashed on takeoff from Tinian, the base that was used. The Enola Gay weapon officer, Navy Captain Deak Parsons, reckoned that if the Enola Gay, crashed, the primer

might explode, detonating the bomb and wiping out the island. So he spent the day before the raid practicing removing and reinstalling the primer — a gunpowder charge about the size of a loaf of bread — so he could install it after takeoff instead.

Doctrine has rather moved away from improvization of weapon safety procedures in the field. If anything we're at the other extreme now, with mechanisms and procedures tested and drilled and exercised and analysed by multiple experts from different agencies. It has of course been an evolutionary process. When weapons started being carried in single-seat tactical aircraft in the 1950s, and also started being slung under the wings rather than in a bomb bay, it was no longer possible for someone to manually insert a bag of gunpowder. There was a move to combination locks: the pilot would arm the bomb after takeoff by entering a 6-digit code into a special keypad with a wired-seal lid. This enabled some measure of control; the pilot might only receive the code once airborne. However both the technical and procedural controls in early strategic systems were primitive.

13.2.1 The Kennedy Memorandum

The Cuban missile crisis changed all that. U.S. policymakers (and many others) suddenly became very concerned that a world war might 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 by only token U.S. custodial forces, so there was no physical reason why the weapons couldn't be seized in time of crisis. There was also some worry 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. In [1223] we find the passage quoted at the head of this chapter.

President Kennedy's response was National Security Action Memo no. 160 [153]. This ordered that America's 7,000 nuclear weapons then dispersed to NATO commands should be got under positive U.S. control using technical means, whether they were in the custody of U.S. or allied forces. Although this policy was sold to Congress as protecting U.S. nuclear weapons from foreigners, the worries about a psychotic 'Dr Strangelove' were also real: they were actually at the top of Wiesner's list, although of course they were downplayed politically.

The Department of Energy was already working on safety devices for nuclear weapons. The basic principle was that a unique aspect 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 accidental or malicious detonation.

The solution then under development was a secret arming code that 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. So it was not acceptable to have the same code in every weapon. Group codes were one possibility — 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 give. 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 needed to be able to arm selected batches (such as 'all nuclear artillery in Germany'). This clearly starts to lead us to a system of some complexity, especially when we realise that we need disarming codes too, for maintenance purposes, and that we need some means of navigating the trade-offs between weapons safety and effective command.

13.2.2 Authorization, Environment, Intent

So the deep question was the security policy that nuclear safety systems, and command systems, should enforce. What emerged was the rule of 'authorization, environment, intent'. For a warhead to detonate, three conditions must be met.

Authorization: the use of the weapon in question must have been authorized by the *national command authority* (i.e., the President and his lawful successors in office).

Environment: the weapon must have sensed the appropriate aspect of the environment. (With atomic demolition munitions, this requirement is replaced by the use of a special container.)

Intent: the officer commanding the aircraft, ship or other unit must unambiguously command the weapon's use.

In early systems, 'authorization' meant the entry into the device of a four-digit authorization code.

The means of signalling 'intent' depended on the platform. Aircraft typically use a six-digit arming or 'use control' code. The command consoles for

intercontinental ballistic missiles are operated by two officers, each of whom must enter and turn a key to launch the rocket. Whatever the implementation, the common concept is that there must be a unique signal; the effectively 22 bits derived from a six-digit code is believed to be a good tradeoff between a number of factors from usability to minimising the risk of accidental arming [908].

13.3 Unconditionally Secure Authentication

Nuclear command and control led to the development of a theory of one-time authentication codes. These are similar in concept to the test keys which were invented to protect telegraphic money transfers, in that a keyed transformation is applied to the message in order 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. So they do for authentication what the one-time pad does for confidentiality.

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

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 was trying to guess a valid message from scratch (*impersonation*) or modify an existing valid message so as to get another one (*substitution*).

An example should make this clear. Suppose a commander has agreed 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 which is the second part of the key.

Suppose the key is that:

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

So '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, and so (provided he understood what it meant) he can send the missiles to the other country. So 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 in favor of the cheated party, then a multi-arbitrator scheme lets his victim denounce him. Schemes may combine unconditional with computational security. For example, an unconditional code without secrecy could have computationally secure secrecy added by simple enciphering the message and the authenticator using a conventional cipher system.

Authentication is in some sense the dual of coding in that 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.

One application that's worth noting is the new GCM mode of operation for block ciphers, described briefly in Chapter 5, 'Cryptography'. In effect this uses the user-supplied key to generate an unconditionally-secure authentication code on the plaintext; it's just a polynomial function of the key and the plaintext. Combined with the counter-mode encryption of the plaintext, this gives an authenticated encryption mode that requires only one pass through the block cipher, rather than the two passes required for CBC plus MAC.

The authentication terminology used in civil and military applications is slightly different [1172]. More importantly, the threat models are different. Soldiers are in general not too worried about non-repudiation — 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.

Quite a few more details have to be fixed before you have a fully-functioning command and control system. You have to work out ways to build the key control mechanisms into warheads in ways that will resist disarming or dismantling by people without disarming keys. You need mechanisms for generating keys and embedding them in weapons and control devices. You have to think of all the ways an attacker might social-engineer maintenance staff, and what you'll do to forestall this. And there is one element of

cryptographic complexity. How do you introduce an element of one-wayness, so that a maintenance man who disarms a bomb to change the battery doesn't end up knowing the universal unlock code? You need to be able to derive the code to unlock this one specific device from the universal unlock, but not vice-versa. What's more, you need serviceable mechanisms for recovery and re-keying in the event that a crisis causes you to authorize some weapons, that thankfully are stood down rather than used. U.S. systems now use public-key cryptography to implement this one-wayness, but you could also use one-way functions. In either case, you will end up with an interesting mix of unconditional and computational security.

13.4 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 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 the authority and field commanders could be maintained, because of the damage that electromagnetic pulse could do (and other possible 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 in which 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 was some precedent for this in submarine-launched ballistic missiles. These exist in part to provide a second-strike capability — that is, to take vengeance on a country that has destroyed your country with a first strike. In such circumstances it is impossible for the submarine commander to be left unable to arm his weapons unless he gets a code from the President. So arming material is kept in safes under the control of the boat's officers, along with orders from the command authority on the circumstances in which weapons are to be used.

Now there is an obvious way to do shared control — just give half of the authentication key to each of two people. The drawback is that you need twice the length of key, assuming that the original security parameter must apply even if one of them is suborned. An alternative approach is to give each of them a number and have the two of them add up to the key. This is how keys for automatic teller machines are managed[2]. But this still may not be

[2]Combining keys using addition or exclusive-or turns out to be a bad idea for ATMs as it opens up the system to attacks that I'll discuss later under the rubric of 'API security'. However in the context of unconditionally-secure authentication codes, addition is often OK.

enough in command applications, as one cannot be sure that the personnel operating the equipment will consent, without discussion or query, to unleash Armageddon. So a more general approach was invented independently by Blakley and Shamir in 1979 [181, 1146]. Their basic idea is illustrated in the following diagram (Figure 13.1).

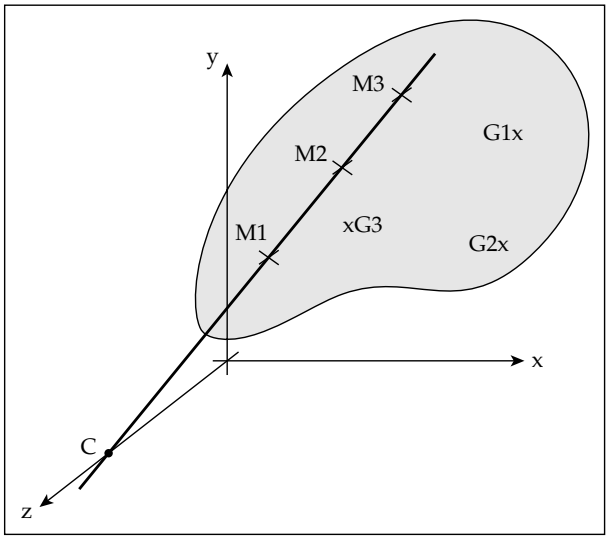

Figure 13.1: Shared control using geometry

Suppose the rule Britain wants to enforce if the Prime Minister is assassinated is that a weapon can be armed either by any two cabinet ministers, or by any three generals, or by a cabinet minister and two generals. To implement this, 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.

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 together with a complexity limited only by the available bandwidth. An introduction to secret sharing can be found in [1226] and a more detailed exposition in [1173]. This inspired the development of threshold signature schemes, as described in Chapter 5, 'Cryptography', 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 give his deputy both halves of the key, and there are all sorts of nitty-gritty issues such as who shoots whom when (on the same side).

In many civilian applications, however, many insiders may conspire to break your system. The classic example is pay-TV where a pirate may buy several dozen subscriber cards and reverse engineer them for their secrets. So the pay-TV operator wants a system that's robust against multiple compromised subscribers. I'll talk about this *traitor tracing* problem more in the chapter on copyright.

13.5 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 published information about PALs can be found in [153]. PAL development started in about 1961, but deployment was slow. Even twenty years later, about half the U.S. nuclear warheads in Europe still used four-digit code locks[3]. 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 also the ability to change codes in the field (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

[3]Bruce Blair says that Strategic Air Command resisted the new doctrine and kept Minuteman authorization codes at '00000000' until 1977, lying to a succession of Presidents and Defense Secretaries [180]. Other researchers have claimed this was not the authorization code but just the use control code.

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. Indeed most weapons have a self-destruct procedure that will render them permanently inoperative, without yield, if enemy capture is threatened. 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 assurance target is summarized in [1223]:

> 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 because 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 8000m/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 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 zero-notice challenge inspections; and staff may be made to re-sit the relevant examinations at any time of the day or night.

I'll 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.

13.6 Treaty Verification

A variety of verification systems are used to monitor compliance with nuclear 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 [1170]. The goal in this application was to have sufficiently sensitive sensors emplaced in each signatory's test sites that any violation of the treaty (such as by testing too large a device) can 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 assembly 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 it had defected, and so the signature was forged. So it is necessary for keys to 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, in which 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 system designers since could have saved themselves a lot of grief if they'd read Gus Simmons' account of these treaty monitoring systems in [1170].)

13.7 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 just the same kind of design bugs, implementation blunders and careless operations as any others.

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 [776]. The situation in the former Soviet Union appears to be very much worse. A survey of nuclear safekeeping describes how dilapidated their security mechanisms have become following the collapse of the USSR, with fissile materials occasionally appearing on the black market and whistleblowers being prosecuted [644].

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 round the world be assured? I'll discuss these later in the chapter on 'Electronic and Information Warfare'.

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

The story is told in [1176]. During the Carter administration, the USA proposed a deal with the USSR under which each side would cooperate with the other to verify the number of intercontinental ballistic missiles. 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 1000 silos by giant trucks, which were designed so that observers couldn't determine whether they were moving a missile or not. So 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 USA 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 ten 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 [1175].)

To see how subliminal channels work, consider the Digital Signature Algorithm described in the chapter on cryptography. The system-wide values are a prime number p, a 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 $r = (g^k \pmod p) \pmod q$, and k is a random session key. The mapping from k to r is fairly random, so a signer who wishes to hide ten bits of information in this signature for covert transmission to an accomplice can firstly agree a convention about how the bits will be hidden (such as 'bits 72–81') and secondly, 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, and then by an American scheme using American equipment. Had the Russians specified a signature scheme like DSA then 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 (MRBM) statistical methods were used. The Russians could say 'we'd like to look at the following 20 silos' and they would be uncapped for the Soviet satellites to take a look. With the end of the Cold War, inspections have become much more intimate with inspection flights in manned aircraft, with 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 digital passport or identity card. Where this is unacceptable, and the card issuer isn't sufficiently trusted not to do it, then the remedy is to use a completely deterministic signature scheme such as RSA instead of one that uses a random session key like DSA.

13.8 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 U.S. 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 fifty buildings contains the alarm response force, gets marked *unclassified controlled nuclear information* (UCNI) adding yet another layer of complexity to the security policy model.

Yet the big picture is open (or so we're assured), and command and control technologies used to be explicitly offered to other states, including hostile ones like the USSR. The benefits of reducing the likelihood of an accidental war were considered to outweigh the possible benefits of secrecy. Post-9/11, it's clear that we'd rather have decent nuclear command and control systems in Pakistan rather than risk having one of their weapons used against us by some mid-level officer suffering from an attack of religious zealotry. This is a modern reincarnation of Kerckhoffs' doctrine, first put forward in the nineteenth century, that the security of a system must depend on its key, not on its design remaining obscure [713].

Indeed, the nuclear lessons should be learned more widely. Post-9/11, a number of governments (including those of the UK and the European Union) are talking up the possibility of terrorists using biological weapons, and imposing various controls on research and teaching in bacteriology, virology, toxicology and indeed medicine. My faculty colleagues in these disciplines are deeply unimpressed. 'You just shouldn't worry about anthrax', one of the UK's top virologists told me. 'The real nasties are the things Mother Nature dreams up like HIV and SARS and bird flu. If these policies mean that there aren't any capable public health people in Khartoum next time a virus comes down the Nile, we'll be sorry'.

13.9 Summary

The 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 huge contribution to the development of security technology.

The rational decision that weapons and fissile material 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 we've looked at in this chapter are authentication codes, shared control schemes and subliminal channels. There are other examples scattered through the rest of this book, from alarms to iris biometrics and from tamper-resistant electronic devices to seals.

Research Problems

The research problem I set at the end of this chapter in the first edition in 2001 was 'Find interesting applications for technologies developed in this area, such as authentication codes.' The recently standardised Galois Counter mode of operation is a pretty good response to that challenge. What else might there be?

Further Reading

As my own experience of this subject is rather indirect, being limited to working in the 1970s on the avionics of nuclear-capable aircraft, this chapter has been assembled from published sources. One of the best sources of public information on nuclear weapons is the Federation of American Scientists [460]. The rationale for the recent declassification of many nuclear arms technologies is

presented in detail at [460]. Declassification issues are discussed in [1361], and the publicly available material on PALs has been assembled by Bellovin [153].

Simmons was a pioneer of authentication codes, shared control schemes and subliminal channels. His book [1172] 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 Doug Stinson's textbook [1226].

Control failures in nuclear installations are documented in many places. The problems with Russian installations are discussed in [644]; U.S. nuclear safety is overseen by the Nuclear Regulatory Commission [976]; and shortcomings with UK installations are documented in the quarterly reports posted by the Health and Safety Executive [586].