CHAPTER

15

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

15.1 Introduction

The catastrophic harm that could result from the unauthorized use of a nuclear weapon, or from the proliferation of nuclear technology, has led the US and other nuclear powers to spend colossal amounts of money protecting not just nuclear warheads but also the supporting infrastructure, industry and materials. Nuclear arms control is at the heart of international diplomacy: while North Korea now has the bomb, South Africa and Libya were persuaded to give it up, Iran's program has been stopped (by both diplomatic and cyber means) while Iraq and Syria have had their WMD programs terminated by force.

A surprising amount of nuclear security know-how has been published. In fact, there are 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, …) 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 forbore use of these weapons even when facing defeat at the hands of non-nuclear states. This is

backed by international agreements, such as the Nonproliferation Treaty and the Convention on the Physical Protection of Nuclear Material [951], enforced by the International Atomic Energy Agency (IAEA).

About ten 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 we have to guard the stuff, in ways that inspire international confidence – not just between governments but from an increasingly sceptical public[1].

A vast range of security technology has spun off from the nuclear program. The US 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. 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, and now used in India's Aadhaar identity system – was developed using US 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. After 9/11, the USA and its allies took many aggressive steps to control nuclear proliferation including:

1. the invasion of Iraq in March 2003, for which the casus belli was a claim that Iraq possessed weapons of mass destruction;
2. an agreement by Libya in December 2003 to abandon an undeclared weapons program;
3. the disclosure in 2004 that Abdul Qadeer Khan, a senior scientist with Pakistan's nuclear program, had helped a number of other countries including Syria, Libya, Iran and North Korea get hold of weapons technology, and the dismantling of his network;
4. the Israeli operation 'Outside the Box' where a suspected Syrian reactor near Deir-ez-Zor was bombed on September 6th, 2007;

[1]For example, the British government was seriously embarrassed in 2007 when the safety of its plutonium stockpile was criticised by eminent scientists [1629], and again in 2018 when parliament's public accounts committee criticised the weapons program's crumbling facilities, aging workforce, specialist staff shortages and endemic funding and practical problems [1563].

5. the 2015 Joint Comprehensive Plan of Action whereby Iran agreed with the USA, the UK, Russia, China, France, Germany and the EU to halt its weapons program.

Not all of the efforts were successful, the obvious case in point being North Korea, which had signed a treaty with the USA in 1994 to halt weapons development in return for oil shipments and help developing civil nuclear energy. This collapsed in 2003, after which Pyongyang withdrew from the Non-Proliferation Treaty and developed weapons. This history makes many people apprehensive of the possible long-term effects of the Trump administration's 2018 abandonment of the agreement with Iran (even though Iran was abiding by it). And then there's also its 2019 abandonment of the Intermediate-Range Nuclear Forces Treaty with Russia (even though that was the result of Russian cheating); and the fact that the New START treaty, signed in 2010 by Barack Obama, will run out in February 2021, unless America elects a president in November 2020 who agrees to renew it.

Nuclear controls apply to more than just warheads and the fissile materials required for their construction. Following 9/11, we learned that Al-Qaida had talked about a 'dirty bomb' – a device that would disperse radioactive material over a city block – which might not kill anyone but could lead to panic, and in a financial center could cause great economic damage. So in 2007, GAO investigators set up a bogus company and got a license from the Nuclear Regulatory Commission authorizing them to buy isotopes. 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, which could have been used in a dirty bomb [1114]. Thanks to the fear of terrorism, the control of nuclear materials has tightened and spread more widely in the economy.

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 US Air Force thought so too. Yet in October 2007, six US 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. All the missiles were supposed to be inspected by handlers in the storage area and checked 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 the glass portholes to see whether the warheads were real or dummy), by the driver calling in the identification numbers to a control centre (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 arrived to unload the missiles and discovered they were live [188, 549]. 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' checklist of their own. So how can you design systems that don't fail in this way?

In this chapter I describe the nuclear safety environment and some of the tricks that might find applications (or pose threats) elsewhere. It has been assembled from public sources – but even so there are useful lessons to be drawn.

15.2 The evolution of command and control

The first atomic bomb to be used in combat was the 'Little Boy' dropped on Hiroshima. Its safety was somewhat improvised. It came with three detonators, and the weapon 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 they used. The Enola Gay weapon officer, Navy Captain Deak Parsons, reckoned that if the plane crashed, the primer might explode, detonating the bomb and wiping out the island. So he spent the day before the raid practising 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 improvisation, and 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 been an evolutionary process. When weapons started being carried in single-seat tactical aircraft in the 1950s, and being slung under the wings rather than in a bomb bay, it was no longer possible to insert a bag of gunpowder manually. There was a move to combination locks: the pilot would arm the bomb after takeoff by entering a six-digit code into a special keypad with a wired-seal lid. This enabled some central control; the pilot might only get the code once airborne. But both the technical and procedural controls in the 1950s were primitive.

15.2.1 The Kennedy memorandum

The Cuban missile crisis changed all that. The Soviet B-59 was a Foxtrot-class diesel-electric submarine that came under attack on 27th October 1962 when a US battle group consisting of the aircraft carrier USS Randolph and 11

destroyers started dropping depth charges nearby. These were practice rounds, dropped in an attempt to force the submarine to the surface for identification; but the ship's captain, Valentin Savitsky, thought he was under attack, that war had started, and so he should fire a nuclear torpedo to destroy the carrier. But this could only be done if the three senior officers on board agreed, and luckily one of them, Vasily Arkhipov, refused. Eventually the submarine surfaced and returned to Russia.

This made the risk that a world war might start by accident salient to US policymakers, and President Kennedy ordered his science adviser Jerome Wiesner to investigate. He reported that hundreds of US 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 token US 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 unauthorized use of nuclear weapons by US officers – 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.' In [1828] we find the passage quoted at the head of this chapter.

Kennedy's response was National Security Action Memo no. 160 [218]. This ordered that America's 7,000 nuclear weapons then dispersed to NATO commands should be got under positive US control using technical means, whether they were in the custody of US or allied forces. Although this policy was sold to Congress as protecting US nuclear weapons from foreigners, the worries about a crazy 'Dr Strangelove' (or a real-life Captain Savitsky) were actually at the top of Wiesner's list.

The Department of Energy was already working on weapon safety devices. 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 starts to lead us to a system of some complexity, especially when we realise we need disarming codes too, for maintenance purposes, and some means of navigating the trade-offs between weapons safety and effective command.

15.2.2 Authorization, environment, intent

The deep question was the security policy that nuclear safety systems, and command systems, should enforce. What emerged in the USA 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, there must be a unique signal; 22 bits derived from a six-digit code are believed to be a good tradeoff between a number of factors from usability to minimising the risk of accidental arming [1351].

15.3 Unconditionally secure authentication

Nuclear command and control drove the development of a theory of one-time authentication codes. As I described in Chapter 5, "Cryptography", 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 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, in that the protection they give is independent of the computational resources available to the attacker. So they do for authentication what the one-time pad does for confidentiality.

Recall that we still have to choose the code length to bound the probability of a successful guess; this 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*). In the GCM mode of operation discussed in Chapter 5, these are set equal at 2^{128} but this need not be the case.

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 that 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 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 the commander's intent) 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. Schemes may also combine unconditional and computational security. For example, an unconditional code without secrecy could have computationally secure secrecy added by simply 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 one 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.

Quite a few 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 may 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. US 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.

One interesting spin-off from authentication research was the GCM mode of operation for block ciphers, described in the chapter on cryptography, which has become the most common mode of operation in modern ciphersuites.

15.4 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 US 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 likely damage from electromagnetic pulses (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. A particular case of this is in submarine-launched ballistic missiles. These exist to provide a second-strike capability – to take vengeance on a country that has destroyed

your country with a first strike. The UK government was concerned that, under the US doctrine, it is possible for the submarine commander to be left unable to arm his weapons if the USA is destroyed, and the President and his lawful successors in office are killed. So the British approach is for arming material to be kept in safes under the control of the boat's officers, along with a letter from the Prime Minister on the circumstances in which weapons are to be used. If the officers agree, then the missiles can be fired.

How can this be generalised? Well, you might just give half of the authentication key to each of two people, but then 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 may not be enough in command applications, as one cannot be sure that the people 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 [257, 1706]. Their basic idea is illustrated in Figure 15.1.

Suppose the rule Britain wants to enforce is that if the Prime Minister is assassinated, then 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 together can 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 [1832] and a more detailed exposition in [1754]. 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 exchange may activate a cold bitcoin wallet'.

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

[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 may be OK.

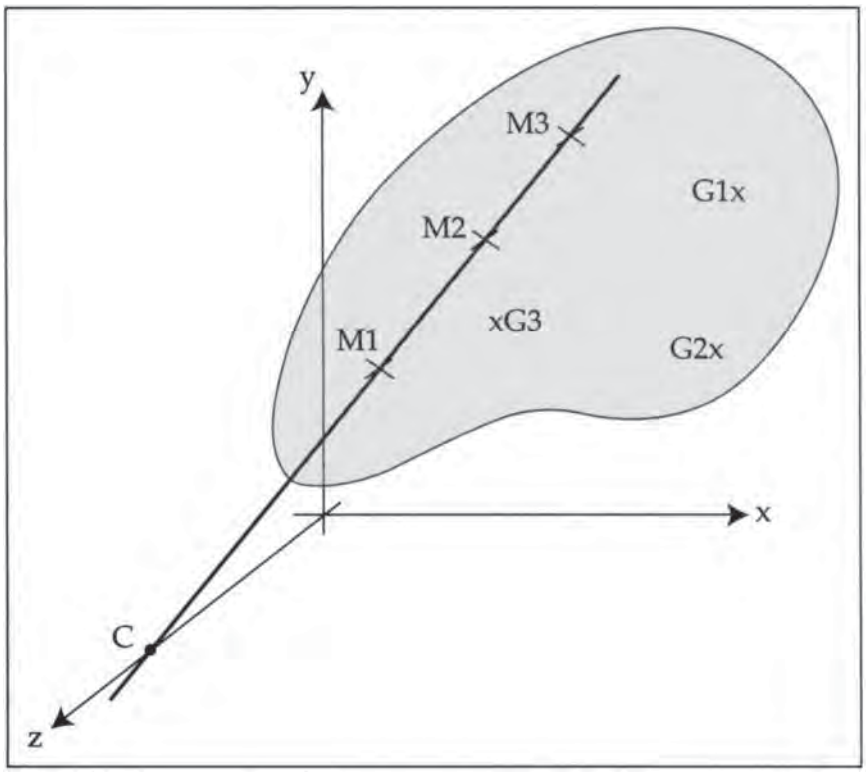

Figure 15.1: Shared control using geometry

there are all sorts of nitty-gritty issues such as who shoots whom when (on the same side). Banking is much the same; it may take two officers to release a large payment, and you need to take care that delegation rules don't allow both keys to fall into the one pair of hands.

In some civilian applications, a number of 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.

15.5 Tamper resistance and PALs

In modern weapons the solenoid safe locks have been superseded by *permissive action links* (PALs), which are used to protect most US nuclear devices. A summary of the published information about PALs can be found in [218]. PAL development started in about 1961, but deployment was slow. Even

twenty years later, about half the US 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 big and complex enough for the PAL to be made inaccessible, the weapon is also stored in tamper-sensing containers called *prescribed action protective system* (PAPS). 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. This self-destruct procedure 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 [1828]:

> *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

[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 [256]. Others say that this was just the use control code.

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 as thick as it is because of fear that shaped charges might be used to disable its protective mechanisms. 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. Finally, at all levels below the President, there is dual control as in banking; no unaccompanied person may approach a nuclear weapon.

I discuss tamper resistance in much more detail in its own chapter, as it's widely used in applications such as bank cards and phones. However, tamper resistance, secret sharing and one-time authenticators aren't the only technologies to have benefited from the nuclear industry's interest. There are more subtle system lessons too.

15.6 Treaty verification

A variety of verification systems are used to monitor compliance with nuclear nonproliferation treaties. For example, the IAEA and the US 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 [1751]. The goal in this application was to have sufficiently sensitive sensors 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 of the assumption of pervasive deceit. Because there is no 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 always disavow a signed message by saying that the official responsible for generating it had defected, and so the signature was forged. So the keys had to be generated within the seismic package itself once it had 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 the account of these treaty monitoring systems by Sandia's former crypto chief Gus Simmons in [1751].)

15.7 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.

15.7.1 Nuclear accidents

The main risk may be just an accident. We've already had two nuclear accidents rated at 7^4 on the International Nuclear and Radiological Event Scale, namely those at Chernobyl and Fukushima, and quite a few less serious ones. Britain's main waste reprocessing plant at Sellafield, which stores 160 tonnes of plutonium – the world's largest stockpile – has been plagued with scandals for decades. Waste documentation has been forged; radiation leaks have been covered up; workers altered entry passes so they could bring their cars into restricted areas; there have been reports of sabotage; and the nuclear police force only manage to clear up 10–20% of cases of theft or criminal damage [1133]. The task of cleaning it all up could take a century and cost over $100bn; meanwhile it has to be guarded [1870]. There are significant and pervasive problems elsewhere in the defence nuclear enterprise, including at the nuclear weapons factories and the submarine bases, ranging from dilapidated facilities, incompetent contractors, poor morale, project delays, spiralling costs, and 20 old submarines awaiting disposal – nine of which still contain fuel [1563]. The situation in Russia appears to be even worse. A survey

[4]The definition is 'Major release of radioactive material with widespread health and environmental effects requiring implementation of planned and extended countermeasures'

of nuclear safekeeping described how dilapidated their security mechanisms became following the collapse of the USSR, with fissile materials occasionally appearing on the black market and whistleblowers being prosecuted [955].

15.7.2 Interaction with cyberwar

A second, and growing, concern is that nuclear safety might be undermined by the possibility of cyber-attack. Even if the command and control channel itself has been made invulnerable to manipulation using the cryptographic and tamper-resistance mechanisms described here, it might be subject to service-denial attack; and in 2018, the Trump administration changed doctrine to allow the first use of nuclear weapons in response to such an attack. Another vital question is whether commanders can believe what they see on their screens. In 1983, a new Soviet early-warning system malfunctioned at a time of international tension, reporting that the USA had launched five Minuteman missiles at Russia. The commander in the Moscow bunker, lieutenant-colonel Stanislav Petrov, decided it was probably a false alarm, as launching only five missiles would have been illogical, and held fire until satellites confirmed it was indeed a false alarm. That was probably the closest that the world got to accidental nuclear war (there had also been a US false alarm three years previously). How would such a system failure play out today, now that we have much more complex systems, with AI creeping into the command chain in all sorts of places without our even realising it? And never mind failures – what about attacks on our intelligence, surveillance and reconnaissance (ISR) capability, including the satellites that watch for missile launches, detect nuclear detonations and pass on orders?

A 2018 report from the Nuclear Threat Initiative describes the concerns in some detail [1837]. It's not enough to protect the weapons themselves, as a cyber-attack on the planning, early-warning or communications systems could also have catastrophic consequences. The main risk is of use because of false warnings or miscalculation; there are also external dependencies, from networks to the electricity grid. Attacks on conventional command-and-control networks could be seen as strategic threats if these networks are also used for nuclear forces. Such issues have been acknowledged in the Trump administration's 2018 Nuclear Posture Review. Technical cybersecurity measures alone are unlikely to be enough, as there are significant soft issues, such as whether key people can be undermined by making them look incompetent.

There may also be fears that an opponent's capability at cyber operations may render one's own deterrent less effective or overconfidence that one's own capability might make attacking a rival less risky. I was personally told by a senior official in the signals intelligence agency of a non-NATO nuclear power that in a confrontation they 'had the drop on' a regional rival. Regardless

of whether this was actually true or not, such sentiments, when expressed in the corridors of power, can undermine deterrence and make nuclear conflict more likely. More recently, the US National Security Commission on Artificial Intelligence warned in 2019 that nuclear deterrence could be undermined if AI-equipped systems succeed in tracking and targeting previously invulnerable military assets [1417].

And it's not just the declared nuclear states. There are currently 22 countries with fissile materials in sufficient quantity and quality to be useful in weapons, and 44 with civil nuclear programs (45 once the UAE goes critical). Of these countries, 15 don't even have cybersecurity laws; energy companies generally won't invest in cybersecurity unless their regulators tell them to, while some companies (and countries) have no real capability.

This has all been made highly salient to governments by the US/Israeli attack on Iran's uranium enrichment capability at Natanz using the Stuxnet virus. In 2009 their output of enriched uranium fell by 30%, and in 2010 the virus came to light. It had infected the centrifuge controllers, causing them to spin up and then slow down in such a way as to destroy about 1000 of Iran's fleet of 4,700. US government involvement was finally admitted in 2012 [1031].

15.7.3 Technical failures

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 work on copyright marking and steganography.

The story is told in [1757]. 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 US Minuteman missiles against a 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, which was thought impractical.

But how could the USA assure the Soviets that there were at most 100 missiles in the silo field, but without letting them find out where? 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 US monitoring facility to Moscow. 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 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 [1756].)

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 first agree a convention about 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 substring.

This could have caused a disastrous failure of the security protocol. But in the end, the "missile shell game", as it had become known in the press, wasn't used. Eventually the medium range ballistic missile treaty (MRBM) used statistical methods. 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. 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, the remedy is to use a completely deterministic signature scheme such as RSA instead of one that uses a random session key like DSA.

15.8 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 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 USA adopted a policy that atomic knowledge was born classified. That meant that if you were within US 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 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 US 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 through, a lot of the technology has been declassified and 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 advantage to an opponent of knowing broadly the system in use.

Many implementation details are kept secret, including information that could facilitate sabotage, such as which of a facility's fifty buildings contains the alarm response force. Yet the big picture is fairly open, with command and control technologies on offer at times to other states, including potentially hostile ones. The benefits of reducing the likelihood of an accidental war were considered to outweigh the possible benefits of secrecy. Post-9/11, we'd rather have decent command and control systems in Pakistan 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, the nineteenth-century maxim that the security of a system must depend on its key, not on its design remaining obscure [1044].

The nuclear lessons could be learned more widely. Post-9/11, a number of governments talked up the possibility of terrorists using biological weapons, and imposed controls on research and teaching in bacteriology, virology, toxicology and indeed medicine. My faculty colleagues in these disciplines were 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." Sadly, the events of 2020 confirm this wisdom.

15.9 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.

Yet even though we can protect the command and control channel that authorises the use of nuclear weapons, that is by no means the whole story. If cyber-attacks can undermine confidence in deterrence by targeting a country's intelligence, surveillance and reconnaissance capabilities, they can still be seriously destabilising. At a time of nuclear brinkmanship, each side could think they have an advantage because of an undeclared cyber capability. And given that US presidents have used nuclear threats about a dozen times since 1945 (Cuba, Vietnam and Iraq being merely the more obvious examples), we might expect several such crises each generation.

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.' By the second edition the Galois Counter mode of operation of block ciphers had been standardised, and by now it's pervasive. What else might there be?

The most serious research problem now might be the interaction between silicon and plutonium. The US/Israeli attack on Iran's uranium enrichment program in 2009–10 gave the world an example of cyber-attacks being used in the nuclear world. In what ways might the threat of such attacks increase the risk of nuclear conflict, and what can we do about it? Given that we can't harden everything the way we harden the command and control channel, what can we do to maintain trust in the supporting systems such as surveillance, or at least ensure that they degrade in ways that don't lead to lethal false alarms?

Further reading

As my own direct experience of nuclear weapons is rather dated – consisting of working in the 1970s on the avionics of nuclear-capable aircraft – this chapter has been assembled from published sources and conversations with insiders. One of the best sources of public information on nuclear weapons is the Federation of American Scientists, who discuss everything from bomb design to the rationale for the declassification of many nuclear arms technologies [672]. Declassification issues are also discussed in [2047], and the publicly available material on PALs has been assembled by Steve Bellovin [218].

Gus Simmons was the guy at Sandia who designed the football; he was a pioneer of authentication codes, shared control schemes and subliminal channels.

His book [1753] 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 [1832].

Control failures in nuclear installations are documented in many places. The problems with Russian installations are discussed in [955]; US nuclear safety is overseen by the Nuclear Regulatory Commission [1457]; and shortcomings with UK installations are documented in the quarterly reports posted by the Health and Safety Executive [876]. The best and most up-to-date survey of problems can be found in the Public Accounts Committee's 2018 report *'Ministry of Defence nuclear programme'* [1563]. As for the interaction 'between silicon and plutonium', there's a recent report on the subject from Chatham House [27].