

CYBER-WEAPONS

THOMAS RID AND PETER McBURNEY

What are cyber-weapons? Instruments of code-borne attack span a wide spectrum, from generic but low-potential tools to specific but high-potential weaponry. This distinction brings into relief a two-pronged hypothesis that stands in stark contrast to some of the received wisdom on cyber-security. Maximising the destructive potential of a cyber-weapon is likely to come with a double effect: it will significantly increase the resources, intelligence and time required for development and deployment – and more destructive potential is likely to decrease the number of targets, the risk of collateral damage and the political utility of cyber-weapons.

In the days and hours leading up to the afternoon of 19 March 2011, air force planners in France, Britain and several other NATO countries were frantically preparing an imminent bombing campaign against military targets in Libya. In Washington on that same March weekend an unusual discussion took place between the Department of Defense (DoD) and the White House. Should America deploy its cyber-arsenal against Libya's air defence system?¹ After the Pentagon's generals and geeks had briefed the president on the options, he decided that, No, the time was not ripe for cyber-weapons.

This behind-the-scenes episode is part of a much larger debate about offensive cyber-weapons. In September 2011, William J Lynn, the US deputy secretary of defense, warned, 'If a terrorist group does obtain destructive cyberweapons, it could strike with little hesitation.'² In January 2012, the Department of Defense announced its plans to equip America's armed forces for 'conducting a combined arms campaign across all domains – land, air, maritime, space, and cyberspace.'³ To counter a novel arms race, China and Russia, among others, have suggested discussing forms of 'cyber arms control' to restrict new forms of military conflict in cyberspace.⁴

But the debate and those trying to turn it into policy are getting ahead of themselves. Some fundamental questions

on the use of force in cyberspace are still unanswered. Worse, they are still unexplored: What are cyber 'weapons' in the first place? How is weaponised code different from physical weaponry? What are the differences between various cyber-attack tools? And do the same dynamics and norms that govern the use of weapons on the conventional battlefield apply in cyberspace?

Cyber-weapons span a wide spectrum. That spectrum, we argue, reaches from *generic but low-potential tools* to *specific but high-potential weaponry*. To illustrate this polarity, we use a didactically helpful comparison. Low-potential 'cyber-weapons' resemble paintball guns: they may be mistaken for real weapons, are easily and commercially available, used by many to 'play', and getting hit is highly visible – but at closer inspection these 'weapons' will lose some of their threatening character. High-potential cyber-weapons could be compared with sophisticated fire-and-forget weapon systems such as modern anti-radiation missiles: they require specific target intelligence that is programmed into the weapon system itself, major investments for R&D, significant lead-time, and they open up entirely new tactics but also novel limitations. This distinction brings into relief a two-pronged hypothesis that stands in stark contrast to some of the debate's received wisdoms. Maximising

the destructive potential of a cyber-weapon is likely to come with a double effect: it will significantly *increase* the resources, intelligence and time required to build and to deploy such weapons – and more destructive potential will significantly *decrease* the number of targets, the risk of collateral damage and the coercive utility of cyber-weapons.

The argument is presented in four steps. Firstly, we will outline conceptually what cyber-weapons are. Then we suggest a way to class cyber-attack tools by discussing the most important empirical cases on record. Thirdly, we explore why even some sophisticated and effective instruments of electronic attack *cannot* be sensibly called a cyber-weapon. Finally, we offer some conclusions.

What are Cyber 'Weapons'?

Weapons are, simply put, instruments of harm. Since the dawn of time, humans have used weapons to hunt prey and each other. Weapons range from the nuclear warhead to the bare human body trained in martial arts, their utility ranging from destroying an entire city to protecting one single person. Yet we often seem to take the meaning of the term 'weapon' for granted. Remarkably, even the US Department of Defense Dictionary of Military and Associated Terms, an authoritative 550-page compendium that defines anything from *abort* to *Zulu time*,



An operative in a US cyber-defence lab, Idaho Falls, Idaho. Photo courtesy of PA.

has no definition for *weapon*, let alone for cyber-weapon.⁵ For the purposes of this article, we understand a weapon as a tool that is used, or designed to be used, with the aim of threatening or causing physical, functional, or mental harm to structures, systems, or living things. This general definition is an essential building block for developing a more precise understanding of cyber-weapons, and by extension cyber-conflict.

We use the term cyber-weapon in a much broader sense than cyber-war. Cyber-war is a highly problematic, even a dangerous, concept. An act of war must be instrumental, political and potentially lethal, whether in cyberspace or not.⁶ No stand-alone cyber-offence on record meets these criteria, so 'cyber-war' remains a metaphor for the time being. Not so in the case of cyber-weapons. Weapons, of course, are not just used in war. Arms are used for a wide range of purposes: to threaten others, to self-defend, to steal, to protect, to blackmail, to police, to break and enter, to enforce the law, to flee, to destroy things, even to train, to hunt and for sports and play. Weapons, of course, may also be used to make war, and some more complex weapons systems are exclusively

developed for that purpose – for instance, warships or anti-aircraft guns. But the majority of weaponry is neither designed for warfare nor used in wars. We argue that this is true also for cyber-weapons. Therefore, while it is counterproductive and distracting to speak about *cyber-war*, it can be productive and clarifying to speak about *cyber-weapons*. Yet conceptual precision remains a problem: 'There is currently no international consensus regarding the definition of a "cyber weapon"', lamented the Pentagon in November 2011, elegantly distracting from the problem that there is no consensus inside the DoD either.⁷ For the purposes of this article, a cyber-weapon is seen as a subset of weapons more generally: as computer code that is used, or designed to be used, with the aim of threatening or causing physical, functional, or mental harm to structures, systems, or living beings.

A psychological dimension is a crucial element in the use of any weapon, but especially so in the case of a cyber-weapon. This is the case in two ways.

The first psychological dimension is *the offender's intention to threaten harm or cause harm to a target*. An instrument may be expressively designed as a

weapon, like a rifle, or re-purposed for use as a weapon, as in using a hammer to threaten or hit somebody.⁸ Simple as well as complex products can be used both for peaceful purposes and as arms. Both in the case of sole-purpose weapon systems and in the case of re-purposed items, a tool is actually used as a weapon when an actor *is intending to use it as such*; whether harm is successfully inflicted or not is of secondary concern. A rifle, for instance, may be used to threaten; it may malfunction; or the bullet may miss the target. But in all cases the arm has been used because an attacker was intending to use it as such in a given situation.

The same logic applies to cyber-weapons. An illustration is the remarkable event at the Sayano-Shushenskaya hydroelectric plant in Russia. Keith Alexander, a general at the head of America's National Security Agency as well as of US Cyber Command, used the incident in a speech to highlight the *potential* risks of cyber-attacks.⁹ With a height of 245 m and a span of 1 km, the Shushenskaya dam is the largest in Russia, holding back the mighty Yenisei River in Khakassia in south-central Siberia.¹⁰ Shortly after midnight GMT on 17 August 2009, a 940-ton turbine,

one of ten 640 megawatt turbines at the plant, was ripped out of its seat by a so-called water hammer – a sudden surge in water pressure, which then caused a transformer explosion. The turbine's unusually high vibrations had eventually worn down the bolts that kept its cover in place. Seventy-five people died in the accident, energy prices in Russia rose, and rebuilding the plant will cost \$1.3 billion. The ill-fated Turbine 2 had been malfunctioning for some time and the plant's management was poor, but the key event that ultimately triggered the catastrophe seems to have been a fire at Bratsk power station, about 500 miles away. Because the energy supply from Bratsk dropped, authorities remotely increased the burden on the Shushenskaya plant. The sudden spike overwhelmed Turbine 2, which at twenty-nine years and ten months' age had nearly reached the end of its predicted lifecycle of thirty years.¹¹ The incident would have been a powerful example of the use of a cyber-weapon if intruders had intentionally caused the plant's crash through a remote command (although to plan such an attack they would have required remarkably detailed advance knowledge of the plant's long-standing maintenance deficiencies). But intention was absent. Intention may be the only line separating attack from accident.

Fundamental questions on the use of force in cyberspace are still unanswered

A second psychological dimension comes into play if a weapon is used as a threat, or if its use is announced or anticipated: the target's perception of the weapon's potential to actually cause harm. It is important to note that the attacker may use a weapon as a threat, which may achieve the objective without actually inflicting physical harm; or the attacker may use the weapon to harm instantly, without threatening to do so first. Furthermore, a target's estimate of a weapon's potential to harm is different from a target's estimate of an attacker's

intention to harm. To illustrate all this, a fictional scenario is useful: suppose an armed robber enters a bank and threatens the clerk with a paintball pistol; both the clerk *and the robber* assume that the paintball pistol is real and loaded with live bullets; money is handed over; the robber flees. Has a weapon been used? We argue that the answer is yes. This fictitious scenario is less anomalous than it may seem; it merely affords starker contrasts. The history of domestic as well as international armed confrontations offers plenty of examples where the aggressor's power to cause injury was vastly overestimated, both by the defender as well as by the aggressor himself.¹² The paintball-pistol scenario inevitably leads to a seeming paradox: suppose the bank clerk noticed that the robber's pistol could only shoot paintballs. Would it still be a weapon? The answer is no. The fake firearm would have lost its threatening character and thus ceased to be a weapon, even if the robber still believed it to be real. The conclusion: a weapon's utility may critically depend on the perception of the threatened party. In every real armed confrontation, both the victim and the aggressor hold crude theories of an arm's capability to inflict harm and their own ability to withstand or absorb this harm. These subjective estimates will necessarily vary in their accuracy when put to a violent test. The actual weapon may be more or less powerful than assumed. In the case of cyber-weapons, this discrepancy is especially large: all publicly known cyber-weapons have far less 'firepower' than is commonly assumed.

Weaponised Software

Cyber-weapons can be grouped along a spectrum: on the generic, low-potential end of the spectrum is malicious software – malware – that is able to influence a system from the outside but technically incapable of penetrating that system and creating direct harm – resembling the proverbial paintball pistol. On the specific, high-potential end of the spectrum is malware able to act as an intelligent agent – capable of penetrating even protected and physically isolated systems and *autonomously* influencing output processes in order to inflict direct

harm, thus resembling the proverbial fire-and-forget smart-bomb. In between are malicious intrusions that include generic system penetrations incapable of identifying and influencing a targeted process, but also targeted and specific intrusions capable of creating functional and even physical damage.

On the low-potential end of the spectrum a *paintball pistol effect*, as we call it, may be observed. Software used to generate traffic to overload a server, for instance, is not strictly speaking physically or functionally damaging a living being, a structure or a system; it is only temporarily slowing down or shutting down a system, without damaging it directly and immediately. Denial of service (DoS) attacks are easy to mount, relatively easy to defend against, but possibly highly visible¹³ – and for those who find themselves for the first time at the receiving end of an attack that is distributed for better effect across multiple attacking machines, the experience can be distressing and it may well create mental harm and even second-order damage: examples include a persistent high-volume distributed denial of service (DDoS) attack which may bring down a bank's website for an extended period of time; defaced websites which may seriously damage an organisation's reputation; and espionage or intellectual property theft that can cost real money put a company in a less advantageous market position. But these damages are second-order effects, not *direct* damage inflicted by a cyber-weapon.¹⁴ At closer inspection, the 'weapon' ceases to be a weapon.

An example was Estonia's reaction to a large DDoS attack in late April 2007. The small Baltic country was well-wired and technologically advanced, and therefore vulnerable to cyber-attacks. With indelicate timing, authorities in Tallinn decided to move the two-metre Bronze Soldier, a Russian Second World War memorial of the Unknown Soldier, from the centre of the capital to its outskirts. Both Estonia's Russian-speaking population and neighbouring Russia were aghast. On 26 and 27 April, Tallinn saw violent street riots, with 1,300 arrests, 100 injuries and one fatality. The cyber-attacks started in the late hours of Friday

27 April. Initially the attackers used rather crude, low-tech methods, such as ping floods and simple DoS attacks. Starting on 30 April, simple botnets were used to increase the volume of DDoS attacks. Estonia experienced what was then the worst-ever DDoS attack. The attacks came from an extremely large number of hijacked computers, up to 85,000, and continued for three weeks. The attacks reached a peak on 9 May, when Moscow celebrates Victory Day. Fifty-eight Estonian websites were down at once. The online services of Estonia's largest bank, Hansapank, were unavailable for ninety minutes on 9 May and for two hours a day later.¹⁵ The effect of these co-ordinated online protests on business, government and society was noticeable, but ultimately it remained minor. But at the time, Estonian officials and citizens were genuinely scared by the attack.

At the opposite, high-potential end of the spectrum is the proverbial fire-and-forget missile. A useful military analogy is the AGM-88 High-speed Anti-Radiation Missile (HARM), initially produced by Texas Instruments, and one of the most widely deployed anti-radar weapons worldwide. The missile's critical innovation is a seeker that includes an intelligent, programmable video processor, designed to recognise characteristic pulse repetition frequencies of enemy radars. This means the weapon can be launched into a certain area where it then searches for suitable target radars, discriminating between friendly and hostile radar by band. Once an emitter is identified as hostile, the missile software's decision logic will allow it to select the highest-value target and home to impact. The missile can be seen as an intelligent agent.¹⁶ In computer science, intelligent agents are autonomous software entities able to assess the environment they find themselves in, and are capable of reacting autonomously in order to achieve a pre-defined goal. Such a quality is necessary to attack the most highly prized targets.

The proverbial HARM missile contrasts with proverbial paintball pistols in at least five important ways: firstly, its objective is not just interrupting traffic at a system's public-facing ports, but getting inside and penetrating a

system. Secondly, its objective is not just penetrating any system that happens to be vulnerable ('low-hanging fruit' in jargon) but specific systems of particular interest. Thirdly, these systems are likely to be protected. For any cyber-attacker with the goal of creating physical damage, the prime targets are likely to be industrial processes, public utilities and civilian as well as military telecommunication networks. The computerised control systems in the most critical installations tend to be well secured.¹⁷ Fourth, if the goal of a stand-alone cyber-attack is physical damage, not just enabling a conventional strike, then the target itself has to come equipped with a built-in potential for physical harm. Weaponised code, quite simply, does not come with an explosive charge. Potential physical damage will have to be created by the targeted system itself, by changing or stopping ongoing processes. Finally, an attack agent's objective is likely to be not just shutting down a penetrated system, but subtly influencing ongoing processes in order to achieve a specific malicious goal. Merely forcing the shutdown of one industrial control system may have the undesirable effect that a fail-safe mechanism or a backup system kicks in, or operators start looking for the bug. To work as an effective weapon, the attack software may have to influence an active process in a malicious way, and if the malicious activity extends over a certain period of time this should be done in a stealthy way. But stealthily or overtly influencing an active process is far more complicated than just hitting the virtual off-button. Three real-world examples of weaponised code illustrate this.

In a first (contested¹⁸) example, the CIA may have rigged the control system of a Soviet pipeline in order to cause a major explosion. The powerful 1982 Trans-Siberian Gas Pipeline explosion was not caused by a system shutdown, but by deliberately creating overpressure in the pipeline by manipulating pressure-control valves in an active control process.¹⁹ A second example is Israel's cyber-attack to blind the Syrian air defence system in September 2007. The goal was not just shutting down the entire air-defence radar station – this would have been

suspicious and could have triggered an alarm or investigation – but to trick the active system to display no approaching airplanes to its operators for a limited time. Thirdly, and most famously, the Stuxnet worm that sabotaged Iran's nuclear programme did not just shut down the centrifuges at Natanz. Before Stuxnet started sabotaging ongoing processes, it intercepted input values from sensors, for instance the state of a valve or operating temperatures, recorded these data, and then provided the legitimate controller code with pre-recorded fake input signals, while the actual processes in the hidden background were manipulated.

Stuxnet is noteworthy in several other respects. One observation concerns the high amount of intelligence programmed into the weapon itself. The attack vehicle was coded in a way that allowed its handlers to connect to the worm through a command-and-control server. But because the final target was not networked, 'all the functionality required to sabotage a system was embedded directly in the Stuxnet executable', Symantec observes in the W32.Stuxnet Dossier, an authoritative analysis of the worm's code.²⁰ Another observation is that it did not create notable collateral damage. Cyber-weapons with aggressive infection strategies built-in, a popular argument goes, are bound to create uncontrollable collateral damage.²¹ The underlying image is that of a virus escaping from the lab to cause an unwanted pandemic. But this comparison is misleading. Stuxnet infected more than 100,000 Windows hosts to increase the chances of reaching the targeted system – yet the worm did not create any damage on these computers. In the known cases of sophisticated cyber-weapons, collateral infections did not mean inadvertent collateral damage.

Finally, Stuxnet is noteworthy for something it did not do. Stuxnet was an intelligent agent, but it was not a *learning* intelligent agent. The *New York Times* calls Stuxnet 'the most sophisticated cyberweapon ever deployed against another country's infrastructure.' One confidential study by America's national laboratories estimates that the worm set

back Iran's nuclear programme by one to two years. 'There were a lot of mistakes made the first time', one senior US official was quoted in the *New York Times*. 'This was a first-generation product. Think of Edison's initial lightbulbs, or the Apple II.'²² A next-generation product could be able to *learn*. Learning software agents and machine learning generally has been the focus of much research attention and funding in computer science of the past decade. The defence and intelligence establishments in the United States, Britain and Israel have traditionally been well ahead of general trends in computer science research, for instance in cryptography or distributed systems. It would be surprising if an intelligent coded weapon capable of learning had not been developed yet. A learning weapon could observe and evaluate the specifics of an isolated environment autonomously, analyse available courses of action and then take action.

Between the proverbial paintball pistol and the intelligent weapon is a large grey area: unauthorised intrusions. Intrusions can also be generic or specific. All can be highly costly, but the nature of the damage and the cost is different. In line with the logic of the spectrum of cyber-weapons outlined above, specific intrusions are far more dangerous than generic ones. Consider three of the most high-profile examples.

Perhaps the most costly generic intrusion to date was the ILOVEYOU worm. On 4 May 2000, a new malware rapidly spread by exploiting a generic scripting engine. A twenty-four-year-old undergraduate student in the Philippines, Onel De Guzman, had programmed the worm. Originating in Manila, it spread across the globe in one day, infecting around 45 million Windows PCs. The worm spread by sending e-mails to entire address books, thus pretending to be a love letter from a known and trusted person. The 'Love Bug', as the media called the worm, was capable of overwriting audio and picture files, replacing them with malicious code. In Britain, 30 per cent of all e-mail servers in private companies were brought down by the volume of requests. The estimated worldwide damages exceeded \$10 billion. Among the infected targets

were governments and defence establishments. Britain's House of Commons saw its internal communication system immobilised. The virus infected four classified internal systems in the Pentagon, according to Kenneth Bacon, then the DoD spokesperson,²³ and it was found on around a dozen CIA computers.²⁴

Cyber-war is a highly problematic concept

The vast majority of malware is entirely generic, not targeted. ILOVEYOU's generic intrusion stands in stark contrast to the highly specific and targeted high-profile intrusions into industrial control systems. So-called Supervisory Control and Data Acquisition (SCADA) systems are used to monitor and control processes in industrial facilities and public utilities, such as chemical plants, electric power plants, refineries, oil and gas pipelines, wastewater treatment and other installations. Large and complex SCADA installations, especially if they include a grid, can cover wide geographical areas. One of the most damaging breaches of a SCADA system happened in March and April 2000 in Maroochy Shire, in Queensland in Australia. After forty-six repeated wireless intrusions into a large wastewater plant over a period of three months, a lone attacker succeeded in spilling more than a million litres of raw sewage into local parks, rivers and even the grounds of a Hyatt Regency hotel. The author of the attack was forty-nine-year-old Vitek Boden. His motive was revenge; the Maroochy Shire Council had rejected his job application.²⁵ At the time Boden was an employee of the company that had installed the Maroochy plant's SCADA system. The Australian plant's system covered a wide geographical area and radio signals were used to communicate with remote field devices, which start pumps or close valves. And Boden had the software to control the management system on his laptop and the knowledge to operate the radio transmitting equipment. This allowed him to take control of 150 sewage pumping stations. Boden was

eventually arrested and jailed for two years.²⁶

Another illustrative demonstration of a cyber-weapon took place a few years later 'on range', that is, in a testing and training environment. In an experiment in 2006, the Idaho National Laboratory tested the so-called 'Aurora' vulnerability that meant some North American power stations were exposed to electronic attack. The test target was a \$1 million, 27-ton industrial diesel generator. The goal: to permanently disable the enormous machine in a controlled environment through an internet-based cyber-attack from 100 miles away. In the test, the generator started shuddering, shaking, and smoke came puffing out, ultimately disabling the machine. The lab allegedly came up with twenty-one lines of code that 'caused the generator to blow up.'²⁷ The malicious code caused the machine's circuit breakers to cycle on-and-off in rapid succession, causing permanent damage through vibration.²⁸

More research into the precise relationship of generic and targeted intrusions is needed. We still do not have a good understanding of the full potential of generic intrusions. This lack of knowledge arises from the complexity and uniqueness of most computer installations, with a bespoke mix of hardware types, networks and software systems, including in most cases software applications that can be many years old, so-called 'legacy systems'. Components of these large-scale systems may be updated and exchanged on a case-by-case basis, so that the larger system and its processes are continually changing. Different parts of such a complex system may be owned, designed, operated, maintained and administered by different organisations. This dynamic applies to modern commercial, governmental and military installations. In fact, the problem is so large that it has become a specific subject for research in Computer Science. The British government and other funders have sponsored research on large-scale complex IT systems. Industrial control systems fall into this category. In SCADA networks and their programmable field devices, attack vectors and configurations tend to be so specific that a purely generic attack seems to pose only limited risks.

Non-weapons

The line between what is a cyber-weapon and what is not a cyber-weapon is subtle. But drawing this line is important. For one, it has security consequences: if a tool has no potential to be used as a weapon and to do harm to one or many, it is simply less dangerous. Secondly, drawing this line has political consequences: an unarmed intrusion is politically less explosive than an armed one. Thirdly, the line has legal consequences: identifying something as a weapon means, at least in principle, that it may be outlawed and its development, possession, or use may be punishable. It follows that the line between weapon and non-weapon is conceptually significant: identifying something as *not a weapon* is an important first step towards properly understanding the problem at hand and to developing appropriate responses. The most common and probably the most costly form of cyber-attack aims to spy. But even a highly sophisticated piece of malware that is developed and used for the sole purpose of covertly exfiltrating data from a network or machine is *not a weapon*. A bug is no weapon either. Two recent high-profile examples illustrate this.

The first example is ‘Duqu’. In early October 2011, a computer security research lab in Hungary, Crysos Lab, discovered a new and exceptionally sophisticated malware threat, which created files with the prefix ‘~DQ’, and so the Hungarian engineers analysing it called it Duqu.²⁹ The threat was identified as a remote access tool, or RAT. Duqu’s mission was to gather intelligence from industrial control systems manufacturers, probably to enable a future cyber-attack against a third party using the control systems of interest. ‘The attackers’, Symantec speculates, ‘are looking for information such as design documents that could help them mount a future attack on an industrial control facility.’³⁰ Duqu was found in a number of unnamed companies in at least eight countries, predominantly in Europe.³¹ The attacks seem to have been launched by targeted e-mails – ‘spear phishing’ in security jargon – not by mass spam. In one of the first attacks, a ‘Mr. B. Jason’ sent two e-mails with an attached MS Word

document to the targeted company, specifically mentioning the company in the subject line as well as in the e-mail’s text. The first e-mail, sent on 17 April 2011 from a probably hijacked proxy in Seoul, Korea, was intercepted by the company’s spam filter. But the second e-mail, sent on 21 April with the same credentials, went through and the recipient opened the attachment. Duqu had a keystroke logger, was able to make screenshots, exfiltrate data and exploit a Windows kernel vulnerability – a highly valuable exploit. The threat did not self-replicate and although it was advanced, it did not have the capability to act autonomously. Instead, it had to be instructed by a command-and-control server. In one case, Duqu downloaded an ‘infostealer’ that was able to record keystrokes and to collect system data. These data were encrypted and sent back to the command-and-control server in the form of JPEG images, so as not to arouse the suspicion of network administrators. The command-and-control server could also instruct Duqu to spread locally via internal network resources.

All these attacks seemed to follow the same pattern. Duqu’s authors created a separate set of attack files for every single victim, including the compromised .doc file; they used a unique control server in each case; and the exploit was embedded in a fake font called ‘Dexter Regular’, including a prank copyright reference to ‘Showtime Inc’, the company that produces the popular *Dexter* sitcom about a crime scene investigator who is also a part-time serial killer.³² Symantec and Crysos Lab point out ‘striking similarities’ between Stuxnet and Duqu and surmise the two were written by the same authors: both were modular, used a similar injection mechanism, exploited a Windows kernel vulnerability, had a digitally signed driver, were connected to the Taiwanese hardware company JMicron, shared a similar design philosophy, and used highly target-specific intelligence.³³ One component of Duqu was also nearly identical to Stuxnet.³⁴ But in one crucial way the two threats were very different: Duqu, unlike Stuxnet, was not a weapon. It was neither intended nor used to harm anything,

only to gather information, albeit in a sophisticated way.

Another example is the German government’s so-called Bundestrojaner (‘federal trojan’). On 8 October 2011, the Chaos Computer Club (CCC) caused a political uproar in Berlin. Germany’s famous hacker club made the news by publishing a report that accused the federal government of using a backdoor trojan to spy on criminal suspects. The software was able to take screenshots of browser windows and Skype, to record voice over IP conversations, and even to download more functional modules that were yet undefined.³⁵ The CCC accused the federal government of ‘state voyeurism’ and, because the trojan’s security precautions were allegedly faulty, of enabling third parties to abuse the software. In the following days several German states admitted to using the spyware, although, officials insisted, under strict legal limitations. Noteworthy for spyware that was ordered by the German government is the home address of the command-and-control server: the commercial internet service provider, Web Intellectuals based in Columbus, Ohio.³⁶

Like Duqu, the Bundestrojaner was a relatively sophisticated intelligence tool used by a state to gather information. It was used domestically, by a law-enforcement agency – that is federal or state police – and was designed to enforce the laws and to maintain the state’s legitimate monopoly of force through the use of arms. But also like Duqu, and like almost all sophisticated cyber-spying operations, this state-sponsored software was not a weapon; it was neither intended nor able to create physical harm, only to gather information, albeit in a sophisticated way.

Conclusions

A thorough conceptual analysis and a detailed examination of the empirical record corroborates our hypothesis: developing and deploying potentially destructive cyber-weapons against hardened targets will require significant resources, hard-to-get and highly specific target intelligence, and time to prepare, launch and execute an attack. Attacking secured targets would probably require the resources or the support of a state

actor; terrorists are unlikely culprits of an equally unlikely cyber-9/11. The scant empirical record also suggests that the greatest benefit of cyber-weapons may be using them in conjunction with conventional or covert military strikes, as Israel did when it blinded the Syrian air defence in 2007. This leads to a second conclusion: the cost-benefit payoff of weaponised instruments of cyber-conflict may be far more questionable than generally assumed: target configurations are likely to be so specific that a powerful cyber-weapon may only be capable of hitting and acting on one single target, or very few targets at best. The equivalent would be a HARM missile that can only destroy one unique emitter, not a set of targets emitting at the same frequency. But in contrast to the missile – where only the seeker needs to be specifically reprogrammed and the general aviation and propulsion systems remain functional – the majority of modular components of a potent cyber-weapon, generic and specific, would have a rather short shelf-life after discovery.

Two findings contravene the debate's received wisdom. One insight concerns the dominance of the offence. Most weapons may be used defensively and offensively. But the information age, the argument goes since at least 1996, has 'offence-dominant attributes'.³⁷ A 2011 Pentagon report on cyberspace again stressed 'the advantage currently enjoyed by the offense in cyberwarfare'.³⁸

But when it comes to cyber-weapons, the offence has higher costs, a shorter shelf-life than the defence, and a very limited target set.³⁹ All this drastically reduces the coercive utility of cyber-attacks. Any threat relies on the offender's credibility to attack, or to repeat a successful attack. Even if a potent cyber-weapon could be launched successfully once, it would be highly questionable if an attack, or even a salvo, could be repeated in order to achieve a political goal. At closer inspection cyber-weapons do not seem to favour the offence.

A second insight concerns the risk of electronic arms markets. One concern is that sophisticated malicious actors could resort to asymmetric methods, such as employing the services of criminal groups, rousing patriotic hackers, and potentially redeploying generic elements of known attack tools. Worse, more complex malware is likely to be structured in a modular fashion. Modular design could open up new business models for malware developers. In the car industry, for instance,⁴⁰ modularity translates into a possibility of a more sophisticated division of labour. Competitors can work simultaneously on different parts of a more complex system. Modules could be sold on underground markets. But if our analysis is correct, potential arms markets pose a more limited risk: the highly specific target information and programming design needed for potent weapons is unlikely to be traded

generically. To go back to our imperfect analogy: paintball pistols will continue to be commercially available, but probably not pre-programmed warheads of smart missiles.

The use of this weapon analogy points to a larger and dangerous problem: the militarisation of cyber-security. William J Lynn, the Pentagon's number two, responded to critics⁴¹ by pointing out that the Department of Defense would not 'militarise' cyberspace. 'Indeed,' Lynn wrote, 'establishing robust cyberdefenses no more militarizes cyberspace than having a navy militarizes the ocean.'⁴² Lynn may be right that the Pentagon is not militarising cyberspace – but his agency is unwittingly militarising the ideas and concepts to analyse security in cyberspace. We hope that this article, by focusing not on war but on weapons, will help bring into relief the narrow limits and the distractive quality of most martial analogies. ■

Thomas Rid is a Reader in War Studies at King's College, London. Peter McBurney is a Professor in the Agents and Intelligent Systems Group of the Department of Informatics at King's College, London.

The authors would like to thank David Betz, Clement Guitton, Pavan Katkar, Tim Miller, Chris Mottershead, Richard Overill, Tim Stevens, John Stone and Gareth Tyson for their constructive comments.

Notes

- 1 Ellen Nakashima, 'U.S. Cyberweapons Had Been Considered to Disrupt Gaddafi's Air Defenses', *Washington Post*, 18 October 2011.
- 2 William J Lynn, 'The Pentagon's Cyberstrategy, One Year Later', *Foreign Affairs*, 28 September 2011.
- 3 Department of Defense, 'Sustaining U.S. Global Leadership: Priorities for 21st Century Defense', Washington, DC, January 2012, p. 4.

- 4 For an overview of such proposals, see Paul Meyer, 'Diplomatic Alternatives to Cyber Warfare: A Near-Term Agenda', *RUSI Journal* (Vol. 157, No. 1, February/March 2012).
- 5 Department of Defense, 'Dictionary of Military and Associated Terms, as amended through 15 September 2011', Joint Publication 1-02, p. 365.
- 6 Thomas Rid, 'Cyber War Will Not Take Place', *Journal of Strategic Studies* (Vol. 35, No. 1, February 2012).

- 7 Department of Defense, 'Cyberspace Policy Report', November 2011, p. 2.
- 8 For a related distinction combined with a rather wide definition, see Lior Tabansky, 'Basic Concepts in Cyber Warfare', *Military and Strategic Affairs* (Vol. 3, No. 1, May 2011), pp. 75–92, 80.
- 9 Keith Alexander, 'Cybersecurity Symposium Keynote Address', University of Rhode Island, 11 April 2011, available at <<http://bitly.com/tNyDmX+>>.

- 10 US Department of Energy, 'Russian Hydroelectric Plant Accident: Lessons to be Learned', Office of Health, Safety and Security, 4 August 2011, available at <<http://bitly.com/t4LDzy>>.
- 11 Steve Gutterman, 'Negligence Factor in Russian Power Plant Accident', *Associated Press*, 3 October 2009.
- 12 For a detailed discussion, see Geoffrey Blainey, *The Causes of War* (New York: Free Press, 1973), pp. 35–56.
- 13 For a detailed look at recent trends in DDoS attacks, see Arbor Networks, 'Worldwide Infrastructure Security Report 2011', Vol. 7, 7 February 2012. Even politically motivated application-layer attacks would hardly count as a weapon.
- 14 An example of a definition that is too narrow is: 'A cyber weapon is an information technology-based system that is designed to *damage* the structure or operations of some other information technology-based system.' Peeter Lorents and Rain Ottis, 'Knowledge Based Framework for Cyber Weapons and Conflict' in Christian Czosseck and Karlis Podins (eds), *Conference on Cyber Conflict Proceedings* (Tallinn: NATO Cooperative Cyber Defence Centre of Excellence, 2010), pp. 129–42.
- 15 These disruptions were the worst of the entire cyber-attack, according to *ibid.*, p. 20.
- 16 M Luck, P McBurney, S Willmott and O Shehory, 'The AgentLink III Agent Technology Roadmap', AgentLink III, the European Coordination Action for Agent-Based Computing (Southampton: 2005).
- 17 A large percentage of industrial control systems connected to the Internet, perhaps as many as 80 per cent, seem to be insufficiently secured. See Éireann P Leverett, 'Quantitatively Assessing and Visualising Industrial System Attack Surfaces', University of Cambridge, June 2011. For a detailed presentation of a volunteer-led security audit of leading programmable logic controllers (PLCs), see the remarkable presentation by Reid Wightman, 'Project Basecamp: Hacking and Exploiting PLC's', S4 Conference, Miami Beach, 19 January 2012.
- 18 Anatoly Medetsky, 'KGB Veteran Denies CIA Caused '82 Blast', *Moscow Times*, 18 March 2004.
- 19 For a more detailed description, see Rid, *op. cit.*
- 20 Nicolas Falliere, Liam O Murchu and Eric Chien, 'W32.Stuxnet Dossier. Version 1.4', Symantec, February 2011.
- 21 See for instance John Markoff, 'A Silent Attack, But Not a Subtle One', *New York Times*, 26 September 2010, p. A6.
- 22 David Sanger, 'America's Deadly Dynamics with Iran', *New York Times*, 6 November 2011, p. SR1.
- 23 It remained unclear how the work affected air-gapped secure systems; see 'Virus Hits Secret Pentagon Network', *BBC News*, 6 May 2000.
- 24 Tom Raum, 'More CIA Employees May Be Disciplined in Computer Case', *Associated Press*, 6 May 2000.
- 25 Jill Slay and Michael Miller, 'Lessons Learned from the Maroochy Water Breach', in E Goetz and S Shenoj (eds), *Critical Infrastructure Protection*, (Vol. 253, 2008), pp. 73–82.
- 26 Tony Smith, 'Hacker Jailed for Revenge Sewage Attacks', *The Register*, 31 October 2001.
- 27 David Fulghum, 'Cyber Attack Turns Physical', *Aerospace Daily and Defense Report*, 27 September 2010, p. 3.
- 28 A CBS News segment of the experiment is at <<http://youtu.be/rTkXgqK1I9A>>.
- 29 Symantec, 'W32.Duqu', October 2011, p. 1.
- 30 *Ibid.*, p. 1.
- 31 Dan Goodin, 'Duqu Targeted Each Victim with Unique Files and Servers', *The Register*, 12 November 2011.
- 32 Aleks Gostev, 'The Duqu Saga Continues', *Securelist*, 11 November 2011.
- 33 Symantec, *op. cit.*, p. 15.
- 34 Sean Sullivan, 'Duqu: Questions and Answers', *F-Secure*, 3 November 2011.
- 35 In English, the malware is called R2D2. See Sean Sullivan, 'More Info on German State Backdoor: Case R2D2', *F-Secure*, 11 October 2011.
- 36 Chaos Computer Club, 'Analyse einer Regierungs-Malware', Berlin, 8 October 2011.
- 37 The assertion was first made in 1996 and has since become part of the debate's standard lore. See John Arquilla and David Ronfeldt, *The Advent of Netwar* (Santa Monica: RAND, 1996), p. 94.
- 38 Department of Defense, 'Cyberspace Policy Report', November 2011, p. 2.
- 39 An ongoing move towards open standards may affect the potential for generic attacks as well as the potential to remotely gather target intelligence, see Vinay M Ijure, Sean A Laughter and Ronald D Williams, 'Security Issues in Scada Networks', *Computers and Security* (No. 25, 2006), pp. 498–506, 500. Also Eric Byres and Justin Lowe, 'The Myths and Facts Behind Cyber Security Risks for Industrial Control Systems', paper presented at Proceedings of the VDE Congress, VDE Association for Electrical Electronic and Information Technologies, October 2004.
- 40 For a comparison between malware and car manufacturing, see Symantec, *op. cit.*, p. 1.
- 41 See Ron Deibert, 'Tracking the Emerging Arms Race in Cyberspace', *Bulletin of the Atomic Scientists* (Vol. 67, No. 1, 2011), pp. 1–8.
- 42 Lynn, *op. cit.*