Cyberwarfare

Benjamin C. Jantzen

INTRODUCTION

Imagine a team of commandos preparing to strike at an unsuspecting nation state. They carry no obvious weapons. They are not dressed in fatigues. In fact, they wear civilian clothing and occupy a nondescript office building on a suburban street. In the operation about to take place, they will not breach any national borders, at least not in person. They will not infiltrate enemy installations or embassies. Instead they attack by sitting at their desks with a radio transmitter connected to a laptop. The transmitter talks the language of the 'smart meter' that the electric company has installed outside their building. The simple computer called a 'microcontroller' in each meter allows it to communicate through a hierarchy of progressively more complex computers to a central control facility of the company that owns the equipment. When the commandos' laptop engages the meter, it exploits a vulnerability in its modest operating system to implant a block of malicious code. This code subtly alters the behavior of the meter, causing it to broadcast to other meters within range and infect them with the same block of code. The malicious code rapidly spreads from meter to meter across a large geographic region. At the same time, each infected meter passes tainted data to the control node above it, until finally the upward cascade ends with the implantation of a payload on the relatively powerful computers of whatever control center that particular meter reports to. At first nothing happens; the weapon has yet to find its target. But eventually, the implanted code leaps the right number of national borders as the number of infected meters grows exponentially. At some point thereafter, the infection spreads upward to infect the control computers of the targeted company. There, the foreign code surreptitiously opens a communication channel to the commandos, and signals them that it's open for business. A much larger piece of code is then uploaded through the hijacked computers of some

innocent bystanders. This much larger digital weapon sets to work coordinating the final stage of the attack. The corrupted control software incites the targeted company's control system to direct a surge of power to a carefully selected set of substations, all at precisely the same moment. At each substation, the surge induces the large transformers at the heart of the station to explode in spectacular fashion, leaving the industrial operation serviced by that particular substation without power from the grid. At each point of attack, diesel generators then roar to life so that critical (and dangerous) manufacturing processes can be safely halted until main power is restored. But the control systems of these generators – also equipped with microcontrollers running tiny programs – have already been corrupted by the commandos. Rather than maintaining a steady flow of electricity, each diesel generator wildly oscillates the throttle, and is quickly reduced to a smoking hulk as the combustion engines tear themselves apart. There is no way to control the manufacturing process that is now running amok. At that point, three munitions plants spread across a country that does not believe itself to be at war simultaneously and – to anyone but the digital commandos – mysteriously explode, resulting in significant loss of life, matériel, and war-fighting capability.

This sort of scenario – or worse – is what those most exercised by the prospect of “cyberwarfare” fear. While nothing quite like this has ever taken place, it is within the realm of practical possibility as I write. The possibility of infecting wireless electrical meters with self-propagating, malicious code has already been demonstrated (Giannetsos et al., 2010; Naone, 2009; Zetter, 2015a, pp. 155–157). The infamous “Stuxnet” – a digital weapon designed and deployed in a collaborative effort by the United States and Israel to slow the advance of Iran's nuclear program – demonstrated the very real plausibility of infecting multiple layers (albeit from the top down) of a remote hierarchical control system like that found in industrial applications throughout the developed world, including systems that run American and European power grids (Zetter, 2015a). Malicious code has already opened a backdoor into regional

electrical supplier control systems that was recently used by hackers to cause blackouts in Ukraine (Goodin, 2016). The destruction of diesel backup generators by remotely injected computer code was demonstrated for reporters as part of Project Aurora (Meserve, 2007a, 2007b). And the fact that many chemical manufacturing processes cannot be halted abruptly without dire consequences is hardly a secret (United States Environmental Protection Agency, 2001).

Such a malicious manipulation of computing technology is possible. But would it be an act of war? If so, is it more of the same as far as innovation in warfighting goes, or do such attacks reflect a use of force that is of a kind genuinely new under the sun? How much risk does cyberwarfare really pose? Does it raise new ethical questions? Despite the hype surrounding the term “cyberwarfare” there exists no clear explication of the concept, and consequently, no consensus on these questions. In fact, there isn't even a consensus on how to spell the term – “cyber war”, “cyber-war”, and “cyberwar” are all contenders.

It is the aim of this article to explicate the phenomenon of cyberwarfare itself. In other words, it's my goal to produce a clear definition of cyberwarfare that captures all or most of the cases on which there is agreement while simultaneously exposing the characteristic features shared by all these cases. Such an explication makes it plain what, if anything, sets cyberwarfare apart from other modes of conflict, and lays the groundwork essential for tackling the harder questions of risk and morality. In order to accomplish this aim, I'm going to follow an indirect strategy. Though an appropriate definition of cyberwar is controversial, many activities are more or less uncontroversially acknowledged to be “cyberattacks”. So I will proceed by explicating the notion of a cyberattack. Then, assuming that warfare essentially involves attacks of one sort or another by one nation state upon another, we can understand *cyber*warfare as warfare involving cyberattacks.

It's important to note at the outset, that there there is a tendency to use terms like “attack” and “war” rather promiscuously to label an enormous and heterogeneous collection of events. But, as James Lewis points out, '..it's unhelpful and incorrect to call every bad thing that happens on the Internet a “war” or “attack”'(2011, p. 23). So I'm going to ignore activities that clearly fall under the rubric of espionage, crime, or activism. Specifically, I'll avoid discussion of such issues as cyber-espionage, “hacktivism,” and cyber-terrorism.

WHAT IS A CYBERATTACK?

So what then is a cyberattack? In this section, I present some illustrative examples described along three dimensions: by their mode of influence (i.e., how they affect their computational targets), by the means of influence (i.e., how the attack reaches it's target), and by the aim of influence (i.e., what the attack targets or is intended to accomplish). This scheme allows for easy generalization beyond the contingent collection of modern technologies and methods.

Mode of influence

Resource attacks

In a resource attack, the attacker aims to adversely influence a target system by depleting it's available communication resources. An old (in Internet years) and common kind of resource attack is the Denial of Service (DoS). All DoS attacks work by flooding the communication ports of a target machine with packets of data. There are many variations on this theme, but one prominent method is the Distributed DoS (DDoS) attack. In a DDoS attack, many computers from different locations send connection requests in rapid succession and swamp a web server so that no traffic can get through, rather like the

Three Stooges trying to pile through a door together. It is a way of "denying service" to the normal users of the website or service hosted by the server. Because the computers engaged in a DDoS are often infected by malicious software that unites them, without the knowledge of their owners, into a "botnet" under the control of a distant "bot-herder", the number of attacking machines can be truly staggering. The largest known botnet, Conficker, encompassed some nine *million* compromised machines worldwide ("Clock ticking on worm attack code," 2009).

Configuration exploits

Other attacks exploit properly functioning but ineptly configured systems. For example, many consumer products ship with default passwords that are easily discovered in public forums. If a consumer fails to change the login credentials upon installing the device, it is then trivial to seize control of that device for malicious purposes. The relatively unskilled miscreants who knocked the Sony and Microsoft gaming platforms offline over the Christmas holiday in 2014 were using a botnet built from home routers seized in this way (Krebs, 2015). Similarly, it is easy for the non-initiate to misconfigure tools, such as peer-to-peer sharing software in such a way as to inadvertently share far more than was intended (Singer and Friedman, 2014, pp. 41–42). Even those who should know better can fail to take basic security precautions, such as neglecting to shield a network with a firewall, failing to set passwords for administrative accounts, or, as was the case with the United States Office of Personnel Management, not encrypting sensitive data (Rein, 2015).

Application exploits

More sophisticated attacks exploit weaknesses in software applications to gain control. For example, many websites check user input – such as login credentials – against a stored database. The language generally used to query the database with the user input is called Structured Query Language, or SQL

(pronounced “sequel”). If care is not taken when writing the code for the website, a user can input a carefully crafted string of characters and have this string interpreted as SQL commands instead of input for a query. Such commands might allow an attacker to see the contents of the database, manipulate or destroy data, or even seize control of the machine hosting the website. This sort of application exploit is called “SQL injection.” Another common mode of attack is the buffer overrun. Here, the attacker provides input that is larger than expected and manages to overwrite space in memory that was reserved for the executing program. This allows the running process to be diverted and for arbitrary code to be executed.

Executive system exploits

The most complex class of attacks comprises what one might call *executive system exploits*. I intend the term 'executive system' to encompass a wide range of physically inhomogeneous but logically related adaptive control systems. On one end of a spectrum, there is the complex operating system (OS), such as Microsoft Windows, Apple OS X, Unix, Linux, FreeBSD, Android, and iOS that is executed by powerful microprocessors with access to large dynamic memories (like in a laptop computer). Given their extraordinary complexity, any OS is bound to contain weak spots either by design or accident. Of those I just listed, Windows is infamously vulnerable. To give just one example, the Stuxnet worm used no fewer than four “zero-day” exploits against Windows (a zero-day exploit is one which is unknown to the victim and the broader public when it is deployed – they thus have zero days of warning). One of these involved a specially crafted .LNK file. A .LNK file is a 'shortcut' in Windows that points to another file or directory somewhere on the computer. When the user inserts a USB drive into a computer running Windows, the operating system automatically opens a window to browse the USB contents. When the corrupt .LNK file is read by the file browser, a flaw in the OS allows malicious code inserted in the file to be executed (“LNK Exploits,” n.d.; Zetter, 2015a).

On the other end of the spectrum is so-called 'firmware'. Firmware is a set of programs or control modules that is somewhere between software, which is easily mutable and intended to be updated during the life of the machine that runs it, and hardware, which is fixed and immutable over the life of the machine. Generally, firmware can be updated, but seldom is. An enormous range of devices you probably don't think of as computers do in fact compute and do so under the control of firmware. Examples include TV remotes, elevators, even toasters. Despite being vastly simpler, the fact that it can be changed or that it can process changing data from one or more sensors means that firmware can be attacked. To give just one example, the firmware that allows a USB flash drive or any other USB device to connect to your laptop currently poses a special risk. In 2014, security researchers announced that they had engineered a way to infect the firmware of the USB controller with self-propagating malicious code (Greenberg, 2014a, 2014b; Karsten Nohl and Jakob Lell, 2014; Security Research Labs, n.d.).

In between these two extremes is a spectrum of mutable code that more or less directly influences hardware and serves to varying degrees the role of an operating system. So, for instance, many industrial machines and processes – likely including your washing machine – are under the control of a Programmable Logic Controller (PLC). A PLC is a microprocessor-based controller that stores a single, changeable program typically written in a very simple mid-level programming language. These programs generally consist of a series of operations to be executed sequentially or a set of simple logical rules for responding to input from sensors (e.g., when the water level sensor trips, start the agitator). Perhaps the most dramatic cyberattack to target PLCs was that carried out by the digital weapon, Stuxnet. The supersonic centrifuges used to enrich uranium for making nuclear fuel or weapons are controlled by sophisticated (but relatively uncomplicated) PLCs. Stuxnet altered the

programs in these PLCs to accomplish an act of extraordinarily subtle sabotage. It caused the supersonic centrifuges to speed up and slow down to wear out their bearings, all the while masking the destruction by playing back false, healthy operating data. In total, the Stuxnet attacks destroyed around 1000 centrifuges (Zetter, 2015a).

Means of influence

Each of the sections above described a vulnerability – a feature of a system which makes it possible to subvert the intended function of that system. But I said little about the means by which those malicious influences could be applied. This provides another dimension along which to classify cyberattacks.

Attacks through the communication network

The Internet is exactly what it's name implies: an interconnected network of networks. By the early 1970's there were a handful of computer networks, linked collections of computers capable of communicating with one another. ARPANET, a project of the US Advanced Research Projects Agency (ARPA) was the first, connecting research computers located at various US universities. The Internet was born when a common protocol for network communication was developed that would, despite their internal differences in hardware, allow these disparate networks to communicate with one another (Singer and Friedman, 2014, pp. 16–21). The modern Internet consists of a collection of physical cables and wires, pieces of hardware such as routers and switches, and the computers they connect. They all communicate with one another using a common protocol. It's a combination of Transmission Control Protocol which slices a message into “packets” and handles their reassembly at the intended destination and Internet Protocol, which concerns the addressing system that directs packets of data passing over the physical network. The pair is often just abbreviated TCP/IP. The Internet is often conflated with the World Wide Web, which is a collection of resources (web pages, videos, and other

data) that: (i) reside on machines connected to the Internet, (ii) are accessible through a Uniform Record Locator (URL) system of naming (e.g., www.ratiocination.org), and (iii) which are heavily interlinked logically by “hypertext” links (the links you click on in websites). The Web is often a target of cyberattack, and it comprises a lot of the information exchanged on the Internet. But cyberattacks are carried out via the Internet.

Beyond the Internet, there is the Internet of Things (IoT) (Greengard, 2015). This is the vast and growing array of objects or things which are not general computing devices like PCs or smartphones but nonetheless have an IP address and communicate with the wider Internet. This includes such things as personal fitness wristbands, “smart” bathroom scales, “smart” thermostats, webcams, traffic lights, and just about anything else big enough to fit the requisite communications hardware in. It is the sum total of devices that speak TCP/IP. This broader network provides an expanded means of attack.

By far the largest number of existing cyberattack tools are transmitted over the IoT, and the IoT is usually where discussions of cyberattacks and cyberwar terminate. But this, I suggest, is to confuse a general phenomenon with a particular technology. What matters from the perspective of projecting force or influence remotely via computing devices is causal influence between those devices, not the particular communication protocols that make the connections in a chain. There is an even more vast and shifting network of devices beyond the IoT that are connected to each other and, typically, to devices that are in the IoT by an inhomogeneous collection of protocols. I have a small hobbyist computer called a Raspberry Pi. It talks to a microcontroller using a simple protocol known as Serial Peripheral Interconnect (SPI). This microcontroller in turn talks to a tiny temperature sensor that also speaks SPI (despite being the size of a lentil). Other common protocols include I2C, USB, and RS-232. Many of the connections that use these protocols are ephemeral by design. In particular, systems that

are considered critical, either for reasons of public safety or protecting sensitive information, are often "air-gapped". In other words, they are physically separated (by air) from the network hardware that would allow the systems to be accessed by others via the Internet. But air-gapping is never perfect; there is always a way around the physical separation. This is because, in order for systems of any complexity to be useful, they must be updated occassionally with fresh data or software. That means connecting one or more devices that may themselves have been exposed to the broader network. Even though the computers that controlled the Iranian centrifuges at Natanz were air-gapped, Stuxnet was spread to those systems via USB flash drives that were previously infected by a computer running Windows that was in turn compromised by an attack through the Internet. Nothing is out of reach.

Attacks through non-communicative interaction chains

Most modern computing devices are "embedded" in larger electro-mechanical systems such as cars, consumer electronics, and military weapon systems. As such, they are intimately bound up with a zoo of sensors and actuators that provide richer ways for machines to project influence on one another, for good or ill, beyond explicit communications protocols. This fact is widely acknowledged with respect to cyber-espionage. The US National Security Agency, for instance, has been aware of (and likely been exploiting) the possibility of surreptitiously reading data from a computer by listening to the electromagnetic radiation generated by the CPU (NSA, 1972; Zetter, 2014) or video display (Kuhn, 2004). Importantly, such unintended channels of influence are also a means of cyberattack. One can, for example, inject malicious signals into a system by exposing its analog sensors to appropriately crafted radio-frequency waveforms (Kune et al., 2013). Alternatively, one can attack by changing the environment around a system's sensors. To give a concrete example, one might exploit a buffer overrun vulnerability in a microcontroller by manipulating the environment in such a way that a sensor connected to the microcontroller feeds a carefully chosen sequence of unexpectedly large values to the

microcontroller. Though this sort of environmental hacking is speculative and, at this point, fairly impractical, this is likely to change the more we saturate the environment with computer-driven actuators and autonomous machines.

Aims of influence

Finally, it is important to consider the aims – the possible targets – of cyberattacks. Obviously, the specific goals can vary even more widely than the tools used to pursue them. Nonetheless, in the context of nation-state conflict we can discern three kinds of objective:

Psychological

Many if not most known cyberattacks aim at psychological disruption. The goals of such an attack include the dissemination of propaganda, intimidation of a particular population, and poisoning the well of information for the public and military planners alike. A notable example took place in Estonia in 2007. That year, this small Baltic nation relocated the Bronze Soldier of Tallinn – a memorial to Russian soldiers who fell in World War II. Russian nationalists took offense and, with the implicit blessing of the Russian government, launched a coordinated DDoS attack against Estonia. Prepackaged tools for conducting a DDoS attack and instructions for using them were disseminated via social media and Russian nationalist websites, and the result was a nation-scale resource attack. It shut down public websites as well as servers running parts of the phone network, the credit-card verification system, and even the nation's largest bank (Clarke and Knake, 2011, pp. 11–16). The effects on computer systems were temporary – the attack lasted three weeks, following which Estonian Internet infrastructure returned to normal. But the psychological effect was clearly more profound (Singer and Friedman, 2014, pp. 110–111). The Estonian foreign minister asserted that, “The attacks are virtual, psychological and real” (Urmas Paet, 2007). Making a case before fellow NATO members that the rules of their

alliance demand a joint response, the Estonian prime minister likened the cyberattack to a naval blockade (Rid, 2013, p. 7).

Tactical

On a quiet night in September of 2007, a squadron of Israeli jets leveled a large building complex located well within Syria's borders. The complex was an illicit nuclear weapons facility designed and built with the assistance of North Korea. Despite sinking billions of dollars into electronic air defense systems, the Syrians never saw the Israeli planes coming or going. Ultimately, it was revealed that Israel had used a cyberattack to blind the Syrian radar. Precisely how they did this remains uncertain – possibilities include hacking the radar directly by feeding it a carefully crafted return signal from a stealth Unmanned Aerial Vehicle (UAV), or the incorporation of a malicious code by an Israeli agent into the software that ran the Russian-built systems (Adee, 2008; Clarke and Knake, 2011, pp. 1–8; Follath and Stark, 2009; Singer and Friedman, 2014, pp. 126–128). Whatever the mechanism used in 'Operation Orchard' as it was called, it is an excellent illustration of the tactical aims to which cyberattacks can be put. Another is the coordinated use of cyberattacks in support of the Russian air, land, and sea assault on Georgia during the 2008 dispute over South Ossetia. A combination of cyberattacks, including DDoS and SQL injection, were closely coordinated with the use of conventional forces to shut down Georgian government websites in order cut off communication with the Georgian people and foreign governments (Carr, 2011, p. 3; Singer and Friedman, 2014, p. 125). More recently, Iran captured a US RQ-170 stealth drone, apparently by spoofing the GPS signal it uses to navigate (Johnson et al., n.d.; Peterson and Faramarzi, 2011; Rawnsley, 2011). Such attacks – by disabling or commandeering enemy equipment, clouding communications on the battlefield, or misdirecting enemy attention – are designed to play key roles in small-scale, short-term military maneuvers.

Strategic

Finally, as suggested by the scenario I sketched at the outset of this essay, cyberattacks can be used strategically to degrade the infrastructure, economy, and war-fighting capability of an adversary. Stuxnet was aimed at slowing the Iranian nuclear program. Though it's actual impact is debatable, it's certain that the worm could have been far more destructive if it hadn't been designed for stealth. It had the capability to destroy the centrifuges at Natanz all at once. Similar strategic attacks against a nation's warfighting capabilities are both possible and doubtless under active development. We are also likely to see cyberattacks aimed at general civilian infrastructure with military implications, such as the power grid, manufacturing facilities, and communications. For example, in January of 2015, hackers seriously damaged a German steel mill by preventing a controlled shutdown of a blast furnace (Zetter, 2015b). In 2008, the United States's Federal Bureau of Investigation reported that counterfeit routers, switches, and interface cards – key pieces of equipment for directing traffic on computer networks – imported from China posed a risk to American networks (Krigsman, 2008). Specifically, a cyberattack could be launched using a “backdoor” in the code controlling the rogue equipment that would allow an attacker (presumably China) to shutdown large swathes of the US communications network. While there is some controversy over the possibility of achieving lasting strategic effects with a cyberweapon (see, e.g., Lewis, 2011), it is clear that much of the critical war-fighting infrastructure of a wired country like the United States is at least temporarily vulnerable.

WHAT MAKES AN ACTION A CYBERATTACK?

Having briefly surveyed some examples of cyberattacks, we are now in a position to look for uniting features. What makes an activity a cyberattack? What, specifically, what makes it “cyber”? What, if anything, separates such an attack from any other act of violence against property or persons?

In light of the above examples, we can reject two of the most common proposals. To begin with, some (e.g., (Shakarian et al., 2013, p. 2)) speak of cyberwarfare as warfare that takes place in a special domain of combat called “cyberspace,” just as the ocean is the special domain of naval warfare. But where exactly is “cyberspace”? It's variably identified with “the information environment” (“DOD Dictionary of Military and Associated Terms,” n.d.), “inter-connected networks of information technology infrastructures” (Tabansky, 2011, p. 78), or merely a “notional environment” (“cyberspace, n.,” 2015). It isn't a literal spatial volume in which you can walk around. But the centrifuges Stuxnet destroyed were most definitely to be found in real, physical space, a space not consisting of information technology infrastructures.

Others have characterized cyberattacks in terms of their intended target. The National Research Council, for example, defines cyberattacks as “deliberate actions to alter, disrupt, deceive, degrade, or destroy computer systems or networks or the information and/or programs resident in or transiting these systems or networks” (Owens et al., 2009, p. 1). This, too is somewhat unsatisfying. While is true that the aim of a DDoS attack is to cripple a piece of “information technology infrastructure” (e.g., someone's webpage, or, more seriously, the network used by the military to coordinate remote actions) that is not obviously the case in general. For instance, when the Iranians downed the RQ-170 drone, they were aiming to influence a piece of physical hardware, not disrupt information. In my introductory scenario, the targets were weapons manufacturing plants, not information infrastructure. This approach thus seems to confuse the means with the aims of cyberattack.

A final and more promising approach views cyberattacks as a special mode of attack. For example, Singer and Friedman (2014, pp. 68–69) say that cyberattacks “use digital means, a computer action of

some sort... Instead of causing direct physical damage, a cyberattack always first targets another computer and the information within it". Clarke and Knake define "cyber war" as "...actions by a nation-state to penetrate another nation's computers or networks for the purposes of causing damage or disruption" (2011, p. 6). And the Tallinn Manual, an internationally crafted, NATO sponsored document often cited as an authoritative source on cyber conflict defines a cyberattack as a "cyber-operation … reasonably expected to cause injury or death to persons or destruction to objects" (Schmitt, 2013). These explications are close to capturing all of the actual known examples of cyberattacks, and at the same time clearly delineating what sets them apart from conventional kinetic attacks. Like a kinetic attack, the targets and venues of cyberattacks are wide and varied. But *unlike* a kinetic attack, cyberattacks are essentially indirect – they are mediated by one or more computers, things that are not inherently weapons. Taking care not to obscure an underlying feature by imposing too much detail that is only an accidental feature of existing technology, we can refine the proposal this way:

cyberattack **-** an intentional effort to apply force against an adversary where the intended effect is the terminus of a causal chain in which at least one link crucially involves a computation that: (1) is specified or determined in advance by the attacker; and (2) induces behavior contrary to that intended by the designers of the device carrying out the computation.

The term "force" in the above definition requires some clarification. As Stone (Stone, 2013) points out, it is important to distinguish three concepts from one another: force, violence, and lethality. All of these have something to do with the notion of warfare. Force, as the great theoretician of war, Carl von Clausewitz (Clausewitz, 1993) understood it is literally that which effects or compels physical change. This is more general than either violence, which involves damage or destruction of persons or things, or

lethality which involves the killing of people. Force does not entail violence, nor does violence entail lethality. Though war overall generally involves lethality, many commonly recognized conventional attacks or acts of war involve only violence or force. For instance, bombing an empty bridge is violent but not lethal. Jamming a radar station with a radio beam is neither lethal nor destructive. Yet both are generally considered attacks. While relatively few are potentially lethal or even violent, all of the cyberattacks described above involve force.

CYBERWARFARE

Recall that my strategy for clarifying the notion of cyberwarfare was to explicate the more straightforward notion of cyberattack and then to understand cyberwarfare as warfare – whatever exactly that may be – involving one or more cyberattacks:

cyberwarfare - warfare involving one or more attacks in which there is an intentional effort to apply force against an adversary, and where the intended effect is the terminus of a causal chain in which at least one link crucially involves a computation that is specified or determined in advance by the attacker and that runs contrary to the intended operation of the computing system.

With this more or less precise definition in mind, we can identify a number of special features of cyberwarfare. Perhaps the most widely discussed of these is the *difficulty of attribution*. Cyberattacks of all sorts are notoriously difficult to attribute to any particular actor. Simple attacks like DDoS can be made through botnets whose constituent machines are scattered around the world in innocent countries, making it difficult if not impossible to determine the source. For more sophisticated attacks, it's possible for the attackers to leave clues behind, such as names or other information buried in compiled code, but there is no way to distinguish a genuine clue inadvertently left behind from a malicious plant

that incriminates the wrong nation state. Practically speaking, this uncertainty can be exploited to produce mayhem of all sorts. In 2007, Israel conducted a military simulation that ended with the US on the brink of conflict with Israel and the Israeli's preparing to invade Syria and Lebanon, largely on the basis of misattributed cyberattacks that were actually (within the simulation) launched by Iran (Zetter, 2015a, pp. 379–380).

Another characteristic feature of cyberwarfare vis-a-vis traditional kinetic warfare is *asymmetry*. Actually, there are two sorts of asymmetry to worry about. The first, and most widely cited, concerns the low cost of producing and deploying a cyberweapon relative to a conventional weapon. That cost is so low that anyone in the world with access to a computer and an Internet connection can launch a cyberattack. This has led to some hyperbolic hand-wringing over a lone teenage hacker bringing a nation state to its knees. But this concern is based on an oversimplified view of the nature of cyberattacks. As we saw above, these range widely in sophistication, destructive power, and strategic value. But the greater the power and precision of a cyberattack, the greater the resources required to pull it off. It's true that a lone hacker can launch a very successful DDoS attack, but shutting down webpages is unlikely to cripple the warfighting capability of a modern nation. On the other hand, attacks like Stuxnet do have substantial strategic value. But they require enormous infrastructure and talent to produce. Stuxnet, for instance, was built by multiple teams of highly trained and talented programmers and required detailed knowledge (and almost certainly, physical copies) of the hardware being used in the nuclear facility at Natanz. Only a nation state is likely to be able to assemble such resources.

There is, however, a more worrisome asymmetry in vulnerability. It is the case that even isolated, otherwise technologically stunted nations like North Korea have the resources to build extensive

cyberwarfare capabilities (Carr, 2011, pp. 246–247). In fact, it's in principle possible for these capabilities to rival those of a large, wealthy, technologically advanced nation like the United States. But a nation like North Korea – with virtually no infrastructure to speak of – is not itself vulnerable to cyberattack. While an effective cyberattack on the US electrical grid could result in significant loss of money, life, and warfighting capacity, destroying the electrical infrastructure of North Korea means relatively little. This means that cyberwarfare offers a new and dangerous kind of warfighting assymetry between nation states, one for which there can be no analog of the “mutually assured destruction” that kept nuclear arms from being deployed.

Finally, cyberwarfare is, at least in principle, subject to greater automation than conventional warfare. As many are fond of pointing out, a cyberattack travels at the speed of electrical impulses. Targets distributed throughout the world can be hit nearly instantaneously. With little time to react, some – such as the National Security Agency (NSA) of the United States – have thought it prudent to build systems to retaliate or to attack automatically when trigger conditions are met (Gillum, 2014). There are at least two sorts of automation to consider. First, systems can be deployed to retaliate or attack automatically when certain pre-determined conditions are met. If more than one nation state employs such a system, there is the unprecedented risk of a chain reaction of increasingly dire automatic retaliation from both sides in a dispute (see (Danks and Danks, 2013) for an overview of this possibility and its moral implications). Second, one can also imagine more adaptive cyberweapons that are endowed with the capacity to adjust plans and learn from experience in order to reach their targets. This sort of weaponized machine learning algorithm raises a slew of worries about doomsday scenarios (e.g., Barrat, 2013).

AN INEVITABLE TREND

As I indicated at the outset, the aim of this article is to get a handle on just what cyberwarfare is so that we can begin to answer important questions about its nature and morality. The upshot of this effort is a view of cyberwarfare as warfare incorporating attacks that are constituted by causal chains involving computation in at least one critical step. I've tried to frame this conception independent of the details of current technology, details that change fast and seem superfluous to the phenomenon in question. While my proposal leaves some key ideas vague, it highlights a very robust trend in the recent history of technology. It is widely known that computing power has grown ever cheaper. What is not often discussed is the extent to which causal chains involving computation continue to lengthen and ramify. Devices with a spectrum of computing power permeate the environment and connect ever more physical processes by computational causal chains. In the late 1960's and early1970's we began to computerize our industrial processes. With the birth of the Internet in the 90's, distant control systems were linked together – however indirectly, a PLC in an automobile factory could influence the temperature of a blast furnace at a smelting plant. With the “sensor revolution” of the early twenty-first century, computational measurement and control spilled out into the broader world of consumer products and private dwellings. Now the refrigerator in my neighbor's house is connected (albeit by a very lengthy chain) to that same blast furnace. All signs suggest that this trend will continue – more sensors and cheaper computing power mean that great swathes of the built environment and also, perhaps, the natural environment will fall under the sway of long causal chains that depend on computation. By extension, the turf of cyberwarfare, the ways in which cyberattacks can impact the infrastructure, environment, and daily life of an adversary will only grow. The future will offer modes of coercion utterly alien to conventional armed conflict.

I've suggested that the phenomenon of cyberattack and, by extension, cyberwarfare is genuinely novel. In a sense, a cyberattack is a meta-tool: a technology for manipulating technologies that in turn

manipulate the physical world. This crude understanding of the nascent phenomenon of cyberwarfare raises a host of questions. Some concern the technology itself. What does this capacity for indirect influence and coercion mean for the development of international conflict? Given the increasing complexity of cyberattacks and the systems they exploit, there is pressure to develop autonomous responses. But is it even possible to devise effective automated systems for cyber defense or attack that learn fast enough and respond appropriately? There are hints already of an extraordinary arms race underway, one that requires ever more sophisticated methods for generating automated military software, perhaps to the point where those methods of generation themselves will have to be automated. Will such weapons constitute human artifacts or something else? Other questions concern the nature of warfare: will cyberwarfare increase or decrease violence in international conflict? Will it make war more or less likely? There is a case to made on either side (Clarke, 2009; Rid, 2013). Finally, and perhaps most pressingly, cyberwarfare raises many new ethical issues. Because of the difficulty of attribution, one must worry whether is it ever possible to satisfy the epistemic demands of the just use of force – if we can never be reasonably sure who attacked us, how can any retaliation be just (Eberle, 2013)? Assuming retaliation can be justified, what about retaliation carried out by machines without human intervention (Danks and Danks, 2013)? Is the dominant moral framework of 'Just War Theory' even applicable to these novel modes of violent and non-violent coercion (Bringsjord and Licato, 2015)? These questions only scratch the surface of what we and the next generation will want to know about cyberwarfare and the curious confluence of technologies that make it possible.

Adee, S., 2008. The Hunt For The Kill Switch. IEEE Spectr. 45, 34–39. doi:10.1109/MSPEC.2008.4505310
Barrat, J., 2013. Our Final Invention: Artificial Intelligence and the End of the Human Era, 9.1.2013 edition. ed. Thomas Dunne Books, New York.
Bringsjord, S., Licato, J., 2015. By Disanalogy, Cyberwarfare Is Utterly New. Philos. Technol. 28, 339–358.
Carr, J., 2011. Inside Cyber Warfare, 2nd Edition, 2nd ed. O'Reilly Media, Inc.
Clarke, R., 2009. War From Cyberspace. Natl. Interest 31–36.
Clarke, R.A., Knake, R., 2011. Cyber War: The Next Threat to National Security and What to Do About It, Reprint edition. ed. Ecco, New York.
Clausewitz, C. von, 1993. On war, Everyman's library. Knopf, New York.
Clock ticking on worm attack code, 2009. . BBC.
cyberspace, n., 2015. . OED Online.
Danks, D., Danks, J.H., 2013. The Moral Permissibility of Automated Responses During Cyberwarfare. J. Mil. Ethics 12, 18–33. doi:10.1080/15027570.2013.782637
DOD Dictionary of Military and Associated Terms [WWW Document], n.d. URL http://www.dtic.mil/doctrine/dod_dictionary/data/c/10160.html (accessed 10.26.15).
Eberle, C.J., 2013. Just War and Cyberwar. J. Mil. Ethics 12, 54–67.
Follath, E., Stark, H., 2009. The Story of "Operation Orchard": How Israel Destroyed Syria's Al Kibar Nuclear Reactor. Spieg. Online.
Giannetsos, T., Dimitriou, T., Krontiris, I., Prasad, N.R., 2010. Arbitrary Code Injection through Self-propagating Worms in Von Neumann Architecture Devices. Comput. J. 53, 1576–1593.
Gillum, J., 2014. NSA planned automated cyberwarfare program. Christ. Sci. Monit.
Goodin, D., 2016. First known hacker-caused power outage signals troubling escalation [WWW Document]. Ars Tech. URL http://arstechnica.com/security/2016/01/first-known-hacker-caused-power-outage-signals-troubling-escalation/ (accessed 1.7.16).
Greenberg, A., 2014a. The Unpatchable Malware That Infects USBs Is Now on the Loose [WWW Document]. WIRED. URL http://www.wired.com/2014/10/code-published-for-unfixable-usb-attack/ (accessed 10.3.15).
Greenberg, A., 2014b. Why the Security of USB Is Fundamentally Broken [WWW Document]. WIRED. URL http://www.wired.com/2014/07/usb-security/ (accessed 10.4.15).
Greengard, S., 2015. Internet of Things. MIT Press, S.l.
Johnson, R., Dec. 5, 2011, 28, 849, 25, n.d. Meet The Russian Avtobaza — Iran's Possible Drone Killer [WWW Document]. Bus. Insid. URL http://www.businessinsider.com/meet-the-russian-avtobaza-irans-possible-drone-killer-2011-12 (accessed 9.20.15).
Karsten Nohl, Jakob Lell, 2014. BadUSB - On Accessories that Turn Evil. Black Hat 2014.
Krebs, B., 2015. Lizard Stresser Runs on Hacked Home Routers — Krebs on Security. Krebs Secur.
Krigsman, M., 2008. FBI: Counterfeit Cisco routers risk "IT subversion" [WWW Document]. ZDNet. URL http://www.zdnet.com/article/fbi-counterfeit-cisco-routers-risk-it-subversion/ (accessed 10.9.15).
Kuhn, M.G., 2004. Electromagnetic Eavesdropping Risks of Flat-Panel Displays, in: Martin, D., Serjantov, A. (Eds.), Privacy Enhancing Technologies, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 88–107.
Kune, D.F., Backes, J., Clark, S.S., Kramer, D., Reynolds, M., Fu, K., Kim, Y., Xu, W., 2013. Ghost Talk: Mitigating EMI Signal Injection Attacks against Analog Sensors, in: 2013 IEEE Symposium on Security and Privacy (SP). Presented at the 2013 IEEE Symposium on Security and Privacy (SP), pp. 145–159. doi:10.1109/SP.2013.20

Lewis, J.A., 2011. Cyberwar Thresholds and Effects. IEEE Secur. Priv. 9, 23–29. doi:10.1109/MSP.2011.25
LNK Exploits [WWW Document], n.d. . Lavasoft. URL http://www.lavasoft.com/mylavasoft/securitycenter/whitepapers/lnk-exploits (accessed 9.29.15).
Meserve, J., 2007a. Sources: Staged cyber attack reveals vulnerability in power grid [WWW Document]. CNN. URL http://www.cnn.com/2007/US/09/26/power.at.risk/index.html (accessed 1.11.16).
Meserve, J., 2007b. Staged cyber attack reveals vulnerability in power grid.
Naone, E., 2009. Meters for the Smart Grid [WWW Document]. MIT Technol. Rev. URL http://www.technologyreview.com/hack/414820/meters-for-the-smart-grid/ (accessed 1.11.16).
NSA, 1972. Tempest: A signal problem.
Owens, W.A., Dam, K.W., Lin, H.S. (Eds.), 2009. Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities. National Academies Press, Washington, D.C.
Peterson, S., Faramarzi, P., 2011. Exclusive: Iran hijacked US drone, says Iranian engineer (Video). Christ. Sci. Monit.
Rawnsley, A., 2011. Iran's Alleged Drone Hack: Tough, but Possible [WWW Document]. WIRED. URL http://www.wired.com/2011/12/iran-drone-hack-gps/ (accessed 9.20.15).
Rein, L., 2015. Top House Republican calls on OPM director to resign over employee data breach. Wash. Post.
Rid, T., 2013. Cyber War Will Not Take Place. Oxford University Press, New York.
Schmitt, M.N. (Ed.), 2013. Tallinn Manual on the International Law Applicable to Cyber Warfare. Cambridge University Press, Cambridge.
Security Research Labs, n.d. Turning USB peripherals into BadUSB.
Shakarian, P., Shakarian, J., Ruef, A., 2013. Introduction to Cyber-Warfare: A Multidisciplinary Approach, 1 edition. ed. Syngress, Amsterdam ; Boston.
Singer, P.W., Friedman, A., 2014. Cybersecurity and Cyberwar: What Everyone Needs to Know, 1 edition. ed. Oxford University Press, Oxford ; New York.
Stone, J., 2013. Cyber War Will Take Place! J. Strateg. Stud. 36, 101–108. doi:10.1080/01402390.2012.730485
Tabansky, L., 2011. Basic Concepts in Cyber Warfare. Mil. Strateg. Aff. 3, 75–92.
United States Environmental Protection Agency, 2001. Chemical Accidents from Electrical Power Outages (Chemical Safety Alert No. EPA 550-F-01-010).
Urmas Paet, 2007. Statement by the Foreign Minister Urmas Paet [WWW Document]. Eesti Päeval. URL http://epl.delfi.ee/news/eesti/statement-by-the-foreign-minister-urmas-paet?id=51085399 (accessed 12.23.15).
Zetter, K., 2015a. Countdown to Zero Day: Stuxnet and the Launch of the World's First Digital Weapon. Broadway Books, New York.
Zetter, K., 2015b. A Cyberattack Has Caused Confirmed Physical Damage for the Second Time Ever [WWW Document]. WIRED. URL http://www.wired.com/2015/01/german-steel-mill-hack-destruction/ (accessed 9.23.15).
Zetter, K., 2014. How Attackers Can Use Radio Signals and Mobile Phones to Steal Protected Data [WWW Document]. WIRED. URL http://www.wired.com/2014/11/airhopper-hack/ (accessed 11.13.15).