2013

Cyber Perfidy / Chapter 29, Routledge Handbook of War and Ethics

Rowe, Neil C.
Monterey, California. Naval Postgraduate School

Routledge Handbook of War and Ethics as chapter 29, ed. N. Evans, 2013
http://hdl.handle.net/10945/36421
Abstract—Perfidy is the impersonation of civilians during armed conflict. It is generally outlawed by the laws of war such as the Geneva Conventions as its practice makes wars more dangerous for civilians. Cyber perfidy can be defined as malicious software or hardware masquerading as ordinary civilian software or hardware. We argue that it is also banned by the laws of war in cases where such cyber infrastructure is essential to normal civilian activity. This includes tampering with critical parts of operating systems and security software. We discuss possible targets of cyber perfidy, possible objections to the notion, and possible steps towards international agreements about it.

This paper appeared in the Routledge Handbook of War and Ethics as chapter 29, ed. N. Evans, 2013.

Index terms—laws of war, cyberspace, perfidy, cyberattacks, cyberweapons, impersonation, product tampering, operating systems, software

INTRODUCTION

Perfidy is an important concept in the laws of warfare. The term covers a category of ruses in which a military force impersonates neutral parties such as the Red Cross to obtain a tactical or strategic advantage [8]. Ruses and impersonation are not generally outlawed by the laws of war; impersonation of an adversary can be justifiable in some situations. But impersonation of neutral parties is generally prohibited because it can hurt both sides of a conflict, causing warfare to deteriorate to chaos. This is because conflicting parties depend on neutral parties to provide food, clothing, shelter, and information. Members of civilian organizations can be especially important in providing humanitarian assistance, evacuating civilians from the area of conflict, providing communications to nonmilitary organizations, and providing connections to what exists as government. These things are not the job of military organizations and it is unreasonable to expect military organizations to provide them. So impersonation of civilians lowers the trust in people appearing to be civilians, making it more difficult for civilians to act freely, and increasing the chances of them being harmed [15].

Article 37 of the 1977 Protocol I Additional to the Geneva Conventions provides one definition of perfidy as acting to "kill, injure, or capture an adversary" by “feigning of civilian, noncombatant status”. This identifies the feigning itself as the crime. This reflects a key component of just-war theory, the discrimination of civilians from combatants. Cyber perfidy can be defined as the feigning of civilian computer software, hardware, or data as a step towards doing harm to an adversary.

A good example of cyber perfidy is the Stuxnet worm and its associated malware [5]. Its ultimate target was command-and-control of particular kinds of centrifuge control equipment from Siemens Corporation, but it used worm methods to spread primarily across many kinds of civilian computers and networks. It thus infected a wide range of objects in cyberspace, masquerading as normal versions of those objects. This infection was quite illegal in most countries because it violated end-user license agreements for Siemens and primarily for the Windows operating
system, even though it did no dramatic direct harm to all but a small fraction of targeted machines. However, the ideas behind Stuxnet are now being used by the criminal cyberattacker community at a net negative cost to world society [10]. Addressing the threats posed by Stuxnet has required work by security-software companies to detect its signatures and implement checking for them. Stuxnet only delayed its intended targets in the Iranian nuclear program a few months, so it appears to have been ineffective against its primary target while having harmful side effects.

Unfortunately, cyber perfidy is more central to cyberwarfare than traditional perfidy is to conventional warfare. That is because computer systems are designed to be highly reliable and resistant to attempts to subvert them for malicious purposes. In fact, computers and digital devices are far more reliable than most non-computer technology per action taken because the digital world is generally far more controllable than the physical world, though it does not always seem apparent because of the tremendous number of actions taken in the digital world. You also need authorization to do something with a computer or digital device. So direct attacks on a computer by sending it unauthorized malicious instructions will just be ignored. Similarly, direct attacks by sending it massive amounts of data will rarely work because most computers have data-rate limiters that control the amount of data they see. Public Web sites accept higher rates but most now have rate limiters too, unlike the country of Georgia for the cyberattacks on its government Web sites in 2008 [21].

That means that an attacker must almost always impersonate a legitimate user of a computer or digital device to get access to it for a cyberattack during cyberwarfare. If they impersonate a civilian, that is cyber perfidy. That means a military attack on a civilian computer or digital device must almost always be cyber perfidious. But even if attackers identify themselves as military during an attack, cyber perfidy is likely necessary for the attack to work. That is because the key civilian mechanisms that run computers and digital devices, their "operating systems", have many protections to prevent people from doing bad things with them. They have protections against changing their instructions and data in other than approved ways, such as control of access and modification rights, encryption of key data, limitations on what programs can do and where, virus and worm checking, monitoring of behavior of ongoing processes, and monitoring of network connections. So nearly all cyberattacks must involve some form of deception with the goal of modifying a victim operating system [3]. Since operating systems are civilian artifacts, serving predominantly as agents of civilian needs, changing them to accomplish an attack is cyber perfidious.

DEFINING CYBER PERFIDY

Cyber perfidy can be identified whenever malicious software or hardware pretends to be ordinary software or hardware, where its goal is to harm software or hardware as part of a military operation. Impersonation of legitimate software by malicious software occurs frequently with criminal cyberattacks today, and is essential to the most serious compromises such as rootkits (taking over the entire operating system of a computer) [12] and botnets (enslaving a set of computers) [1]. Most current cyberattacks involve subversion of the Windows operating system. Such impersonation enables cybercriminals to make money by sending spam email through compromised computers, hosting phishing scams, attacking sites they do not like, launching "denial-of-service" blockades against sites, and blackmailing sites by threatening attacks against them. Nearly all proposed offensive cyberwarfare techniques use impersonation too. But that could be cyber-perfidious.

Impersonation of legitimate software and hardware is illegal in nearly every country in the world, as it is a form of tampering with a product. Nearly all end-user license agreements (“EULA”s) for software and hardware exclude warranty after modification. Modification means the producer cannot be responsible for what it does, since even a small change in a program can create a big difference in behavior. This has not deterred criminals. But it is a different matter for a legal government to use these tactics. Guns have legitimate uses in police work and hunting, but there are few legitimate uses for cyberweapons in the civilized world, just as there are few legitimate uses for chemical, biological, and nuclear weapons. (A possible exception is "red teaming" to test the cyber-security of
systems, deliberately attacking them using as many tricks as possible, but this is generally a crude and inefficient method of testing systems, and many better tools are available.)

Tampering with consumer products is a serious crime in criminal law [4]. In the United States, the Federal Anti-Tampering Law (USC Title 18 Chapter 65, "Malicious Mischief") defines several categories including "whoever, with intent to cause serious injury to the business of any persons, taints any consumer product or renders materially false or misleading the labeling of, or container for, a consumer product". This has a penalty of up three years in prison; tampering that causes bodily harm receives up to twenty years. These laws clearly apply to software as a consumer product under the classification of "device" where malicious modification renders misleading the labeling of the software.

We are starting to see international cooperation on extending the laws of war to cyberweapons [7, 23]. Identifying and labeling cyber perfidy is in the tradition of “just war” theory where civilians are not legitimate primary targets [22]. Or as [19] puts it, it is wrong to attack those who cannot harm you. General-purpose computer software and hardware is carefully designed not to harm anyone. While military software may be installed on systems that could be legitimate targets, the general-purpose services on systems are like human civilians, and should not be legitimate targets of attacks by analogy to just-war theory. Computer programs have grown so much in sophistication and complexity in recent years that they are no longer always “mechanisms”, but are approaching “artificial intelligences” that are like people in some ways. Most of these programs would qualify strictly as “noncombatants” [11] since they do not contribute to military operations.

**Cyber-resources that are possible targets of perfidy**

Not all cases of software and hardware impersonation should be labeled as perfidious, however. Ruses are accepted as legitimate tactics of warfare when impersonation is done of military personnel and civilian non-neutral parties such as truck drivers bringing materiel. So to qualify as cyber perfidy, the target of the impersonation should be something primarily used by civilians. Then to constitute a war crime, the perfidy must be used to launch an attack. Non-cyber perfidy need not directly damage the object of its impersonation, but some such collateral damage is inevitable with cyber perfidy since the functionality of existing software or data is changed and will not revert to normal functionality without explicit action. That is another argument why cyber perfidy is a serious matter.

An analogy in conventional warfare would be a well. Generally speaking, poisoning a well is not acceptable by the laws of warfare, although it could provide the important tactical advantage of forcing a civilian population to move on [22]. In a village where a communal well is the only source of water, poisoning it would be attacking a resource too central to the civilian community to satisfy the criterion of discriminability of civilian targets from military ones. It is even more a war crime if the poisoning is not announced and people start dying without knowing the cause [6]. Cyber perfidy is similar to the poisoning without announcement, since the effectiveness of cyberattacks generally depends on keeping them secret as long as possible.

Consider some cyber-services whose modification would fulfill this definition of cyber perfidy and which would be especially useful to a cyberattacker:

- The security kernel of an operating system. This controls what rights programs have to do things, in particular to change items in memory or secondary storage. It maintains the functionality of the operating system (its "integrity"), and without it many kinds of harm are possible. Most criminal cyberattacks want to compromise it and exploit it.
- The file management features of an operating system. Software is useless unless it can manage files. Attacks do not need to affect all files, just certain critical ones.
- The update manager. Since software is usually created without guarantees, updates are essential to quickly block new kinds of attacks. Attacks on the manager can prevent timely defenses or could
install malicious ones.

- The networking software. Some important applications critically depend on access to a network. This is particularly true of small wireless devices which need networking to complete basic tasks for which they lack resources.
- Electronic mail services. People need to communicate.
- The hardware (integrated circuits) that implements the central-processing unit (CPU) of the computer. If these can be changed, then the computer can be run any way an attacker wants. However, this requires physical access to the machine and cannot be done over the Internet.

Attacks on the above cyber-services should be considered cyber perfidious if they happen on civilian computers or networks. For cyber-services on a military computer or network, it is more arguable whether the term applies. At issue is whether an attack could spread to civilians; for instance, if a virus attack on a military computer could easily spread to civilian computers, then it could be cyber perfidious. (U.S. military networks are connected to civilian networks in ways not always realized.) But tampering with a product or poisoning a well used by military personnel could be considered as violating the laws of war anyway even if it does spread to other instances.

By contrast, consider some cyber-services that could be impersonation targets [2] but for which this could not be cyber perfidy:

- Weapons systems. These are legitimate targets of cyberweapons.
- Military planning and coordination systems (“command and control”). These are legitimate targets of cyberweapons.
- Management of specifically military organizations and their contractors.
- Specific manufacturing software for weapons. That was the intended target of Stuxnet, but does not excuse its collateral damage.
- Web pages on military or government Web servers. These are an important target of military propaganda or "influence" operations.

Methods of tampering

The tampering with software or hardware in cyber perfidy can have several effects:

- The service can be modified to support an attack, such as browsers infected by malware that direct Web users to certain attacker-designed pages that help launch attacks.
- The service can be modified to malfunction so that users cannot do what they normally can, such as electronic mail services modified to malfunction during an attack and thus fail to provide warnings to other sites.
- The service can be modified to directly create harmful new effects, such as launching cyberattacks on other sites.
- The service can be modified to actually harm people, such as launching attacks against water treatment plants or hospital computer systems.

A "strict constructionist" approach might consider only the last to be cyber perfidy. But a "loose constructionist" argument can be made for all of these. There is rarely any military necessity for cyberattacks to do these things, since there are many better weapons for achieving military goals. Cheapness and simplicity are not a justification for using a weapon, as there are many horrible cheap and simple weapons like mines.

[14] points out that cyberwarfare, like most innovations in warfare, was initially hoped to be more precise in effects than conventional warfare. But this has not been true because of the dependence of most cyberweapons on flaws in software – and flaws can get fixed unexpectedly – and the easy ability to use deception in cyberspace by potential
targets. Thus the tampering with software and data necessary to accomplish a cyberattack must often be substantial to ensure a militarily significant effect. This makes it easier for the tampering to spill into civilian software and data.

**Perfidy involving human cyber-combatants**

Another aspect of the definition of perfidy in the laws of war is identification of combatants. Cyberwarfare risks violating this because the combatants tend to be unseen; they may be programmers or software users in a military organization who do not think of themselves as soldiers. However, if they help launch attacks, they are legitimate targets of retaliation in the form of bombs and targeted assassinations. That means that cybercombatants should announce the geographical locations of their cyberwarfare groups in advance of hostilities to avoid perfidy.

**Repairing the damage of cyberattacks**

Some of the damage of cyber perfidy is short-term: time delays from added malicious activity, space wasted, reduced resources to deal with concurrent criminal cyberattacks, and personnel time for investigation since most critical-infrastructure systems have humans monitoring for anomalous behavior and investigating it. But cyber perfidy also results in long-term damage and entails costs to clean up afterwards. Because careful deception is generally necessary for the perfidy, and the attack must be overwhelming to have a significant tactical effect, many things may need to be repaired. It is very hard to recognize unauthorized changes to software and data. While software may be restorable from backup, that may not be possible with data that needs to be timely. And the backup may itself be tainted, so restoration from backup may be impossible. Thus most cyberattacks are “dirty” weapons that leave a mess. Nonetheless, it is possible to design “cleaner” cyberweapons by making their effects reversible and that can be encouraged.

Another serious problem is that cyberattacks of all kinds are quickly analyzed and dissected by the information-security community in an effort to prepare for similar future attacks. After some dispute in the 1990s, the consensus today is that it is better for the public interest to report new attacks and get public fixes than to conceal attacks and the vulnerabilities that led to them (TechRepublic, 2005). This analysis is done publicly on sites such as www.kb.cert.org (the US-CERT Vulnerability Notes Database), eve.mitre.org (the Common Vulnerabilities and Exposures site), web.nvd.nist.gov (the National Vulnerability Database), and www.securityfocus.com (the Bugtraq wikis for reporting vulnerabilities). The results of such analysis and reverse engineering are quickly available to criminal cyberattackers. Stuxnet is an example because many of its new propagation methods targeted the widespread Microsoft Windows operating system and had nothing to do with centrifuges. Analysis or reverse engineering of a conventional weapon is usually not helpful for accomplishing new attacks because most conventional weapons either use well-understood principles (like mines), are expended upon attack (like missiles), or do not provide key information contributing to their effectiveness (such as bullets, which do not reveal the military intelligence and planning used to target them). On the other hand, a good part of the damage of a cyberweapon will be in its subsequent enabling of criminal cyberattacks against civilians far from the scene of the original attack.

**OBJECTIONS TO THE CONCEPT OF CYBER PERFIDY**

Various objections have been raised to this notion of cyber perfidy. For one thing, cyber perfidy does not seem to invoke strong emotions. Traditional perfidy can, since attacking neutral parties elicits strong moral condemnation in most legal and ethical systems. Warfare that attacks human health like poison gas often invokes a particularly visceral fear [16]. Such visceral responses generally occur in regard to threats that have been with the human race for millions of years. Unfortunately, computers are very new in the evolution of the human race and our emotional systems are not tied to them (except maybe in tampering with computer-based children’s toys). But maybe attacks on them should create equal fear and outrage too considering how important they are in our lives. So we need to
evaluate objectively whether the damage due to cyber perfidy can be similar to the damage due to traditional perfidy, not from a gut feeling. We argue here that it is.

Some have argued that the notion of perfidy is based on the threat to human lives [19]. Thus shooting Red Cross personnel would be murder but disabling a computer would just be vandalism. Certainly cyber perfidy could threaten human lives if it results in violence, as with Stuxnet causing uncontrollable acceleration in centrifuges. But the Geneva Convention section on perfidy includes injuring and capturing adversaries as a result of perfidy as well as killing. We argue that is because the key casualty of perfidy is trust, not lives. If we cannot trust neutral parties, we will be unwilling to use them for their intended purposes [22]. Even when Red Cross personnel are not in mortal danger, the value of the Red Cross is greatly reduced when they cannot be trusted. Damage to trust, when spread over many instances, can be equal or worse than a harm done to any one individual. Consider financial services such as banking, on which everyone depends. A cyberattack that targets banking and reduces its trust could be highly significant. This could be done by targeting online banking with malicious software that changed monetary amounts randomly so no one could tell how much money they had in their accounts. If the damage to a banking system affects many customers, the harm done to each of them could add up to the harm of several murders. Militaries often risk soldiers in achieving tactical and strategic goals that are considered more valuable than the individual lives of soldiers.

A further objection to the concept of cyber perfidy is that it does not involve explosives, a concern of the laws of war. However, cyber perfidy can have an analogous effect to an explosion in cyberspace (hence the term "logic bomb") in that it can reduce a sophisticated software device to a set of unconnected fragments. Cyber perfidy can also be like a booby trap since it must be activated primarily by normal activities of its victim to do its damage. Booby traps are outlawed by the United Nations Convention on Certain Conventional Weapons Protocol II (1996 as amended) [20]. Many clever forms of booby traps were used with improvised explosive devices (IEDs) by Iraqi insurgents 2003-2011 in violation of the Convention and were universally condemned. For instance, explosive devices were hidden in trash or other everyday objects, violating the prohibition against using “apparently harmless portable objects which are specifically designed and constructed to contain explosive material”. Cyber perfidy tries to hide its malicious intent inside standard or “everyday” code in an analogous way.

A different concern is that cyberweapons may just seem awful because they are new, and cyberweapons may well become an accepted part of future warfare. Objections were raised to torpedoes early in the twentieth century for this reason, and some argued they violated the “civilized” nature of naval warfare [9]. Today torpedoes are well accepted in military arsenals. However, torpedoes are visible objects that use conventional munitions. They can be seen leaving the attacker who must be nearby, they have a direction, they cause clear damage all at once at a target, and the damage is generally localized with a clear cause. None of these things applies to cyberweapons, which makes them much more difficult to monitor. Thus objections to cyberweapons should not be dismissed merely because they are new.

Some may consider that cyberweapons are the natural result of the evolution of warfighting technology towards greater control of technology [13], and we should not object to “progress”. One response mentioned earlier is that countermeasures to cyberweapons like better information-security practices by defenders and deliberate deception are making it uncertain that these weapons provide commanders with greater control of conflict. Another is that “progress” has brought us chemical, biological, and nuclear weapons in the last hundred years, yet international laws and other incentives prevent countries from using them.

Following just-war theory, some might hold that cyber perfidy exploits a “dual-use” technology of information systems and such double effects are accepted in warfare provided the primary effect is military [22]. However, the most effective targets of cyberweapons as described earlier are things like operating systems that are inherently civilian in nature. Furthermore, the fraction of military traffic carried on the Internet is a small fraction of the total traffic, so attacking the Internet has little dual-use effect on military systems. However, attacking a specifically
military network that is substantially unconnected to the Internet would be an exception, and could be appropriately justified by the dual-use principle.

A final objection is that cyber perfidy is generally used as a stepping stone to enable well-focused attacks on a few key targets, so it is desirable for attackers to design its perfidy so that it is mostly harmless to the stepping stones. However, an important principle in the laws of war is that neutrality of non-participants must be respected [22]. Operating systems have no stake in either side of a conflict. To co-opt them for launching an attack on another country, as the Russians did with computers in a range of countries including the United States during the Georgia attacks of 2008, is like violating the neutrality of a country as Germany did to Belgium in World War I and was roundly criticized for. Attacks that use more substantial resources of a country such as setting up botnets within them could be considered similar to forced conscription or at least forced labor in munitions plants as the Germans did in France in World War II, again tactics outlawed in the laws of war.

INTERNATIONAL AGREEMENTS ON CYBERPERFIDY

Since cyber perfidy can lead to uncontrollable consequences in warfare, it is important to seek international agreements to control it. Three useful ideas that should be included in such agreements are international cooperation in detection, policy on attribution of attacks by attackers, and policy mandating selection of nonperfidious methods for attacks.

Cyber perfidy can be detected by comparing the bit patterns of critical software before and after an attack on it. This can be done efficiently by comparing hash codes. When software is installed on a computer, hash codes using one of the standard algorithms like SHA-1 can be computed on their files, and can be recomputed periodically. If they change that means that the software has been modified. Hash codes on a broad range of software and its updates can also be obtained for free from the National Software Reference Library (NSRL) maintained by the U.S. Government organization NIST, and commercial companies like Bit9 provide supplementary hash codes. If comparison of hash codes is so simple, why isn’t it done routinely to protect computer systems? It would slow systems considerably since there are many files to check. Also, a rootkit that commandeers an operating system can disable checking software or make it give false results, so the hash codes are unreliable once a serious attack has succeeded against a computer system. And certainly there are legitimate cases where software must be changed to update it in response to discovered bugs.

International cooperation will aid in detecting cyberattacks and their precursors because cooperation greatly helps in controlling criminal cyberattacks today as discussed above. Early reports of possible attack patterns suggesting nation-sponsored cyberwarfare can be identified in the far larger volume of normal attacks, and summarized and reported to international agencies.

One the most disturbing features of cyber perfidy is the difficulty of attributing it. Attribution is possible in some cases: A country may announce responsibility for an attack, or we may be able to trace its origin across the Internet using a number of techniques, or we may be able to find source code for a distinctive cyberweapon during a police raid. But attribution is more difficult in cyberconflict than in conventional conflict. When a missile is launched at a country, there are clues to who launched it in its path and velocity vector. When software or hardware is maliciously changed, there are frequently no clues; Stuxnet has still not been attributed two years afterwards. Detecting the source of a cyberattack is difficult because cyberattacks frequently hide or falsify their origins. Detecting the true origin of an external cyberattack requires broad monitoring of the Internet, and to do it truly effectively, agencies must access information not normally available. But also attacks can be planted by insiders and do not need to come in over the Internet at all, or they can originate from tainted storage media. Thus attribution is very difficult with
cyberattacks involving cyber perfidy. A country may just assume it is due to a long-time adversary when it has actually been caused by a trouble-making third party.

Therefore, we have argued elsewhere that responsible countries, or just countries wishing to ensure that their attacks have a precise effect, label their cyberattacks in some way [18], much as the Hague Convention stipulates that belligerents should have a “fixed distinctive emblem recognizable at a distance”. It is hard to conduct a war fairly when it is unclear who is attacking whom. So a label or signature is needed, and must be concealed from casual inspection, so techniques of steganography must be used to hide its existence until the attacker chooses to reveal it. In addition, the signature itself needs to be encrypted with a key known to the attacker or a neutral agent so that by decrypting it the attacker can prove that they are the source of the attack. International agreements can specify standards on the form and methods of signatures.

Another possible subject of international agreements is restrictions on cyberattacks to avoid cyber perfidy. The section "Defining Cyber Perfidy" gave a list of perfidious targets and methods that could be stipulated. While nonperfidious attacks may have harder challenges to achieve desired damage, the damage can still be severe. In fact, it is often easier to attack applications software than the operating system because most protections are designed for operating systems. That is because most criminal attackers want broad effects and will get most leverage from attacking the operating system. However, cyberwarriors do not have these goals and, if desired, can rely on more targeted methods such as modifying a small set of applications software by methods such as viruses. The situation is similar to bombing a specific building in Baghdad without damaging the neighboring civilian buildings, which was done successfully during U.S. military operations in 1991 and 2003 Computers permit more precise control than bombs and ought to offer more precise effects, particularly when their targets are not as powerful as an operating system.

What recourse does the world community have if countries persist in engaging in cyber perfidy? Sanctions that target a country's ability to do Internet commerce are appropriate. With international cooperation, sanctions could be powerful tools. In addition, since cyber perfidy generally violates product-tampering laws, those laws could be enforced by traditional criminal legal proceedings in the countries victimized.

CONCLUSIONS

Does defining something as cyber perfidy really matter? We believe so because there are policy implications: Perfidious actions can be considered off-limits to a civilized and decent society. Less scrupulous countries may still engage in cyber perfidy. But then we will have an additional mark against them that we can use in motivating international coalitions against them.

Most discussion of cyberweapons treats them as similar to other weapons. But cyberspace is a fundamentally different technology than that of explosives because its weapons need not be physically localized, need not be large compared to their effect, and need not have damage that can be easily recognized. For these reasons the laws of war need to address cyberweapons from a fresh perspective. Clearly certain aspects of cyberweapons could be highly dangerous. Cyber perfidy would seem a good thing to prohibit in the laws of war because of its uncontrollability and destabilizing effects. It is, however, just one of the many ethical problems raised by cyberwarfare [17].

ACKNOWLEDGMENT

The views expressed are those of the author and do not represent those of the U.S. Government.

REFERENCES