2011

Challenges in Monitoring Cyberarms Compliance

Rowe, Neil C.; Garfinkel, Simson L.; Beverly, Robert; Yannakogeorgos, Panayotis

Monterey, California. Naval Postgraduate School

International Journal of Cyber Warfare & Terrorism, Vol. 1, No. 1, pp. 1-14
https://hdl.handle.net/10945/36013

Challenges in Monitoring Cyberarms Compliance

Neil C. Rowe[1], Simson L. Garfinkel[1], Robert Beverly[1], and Panayotis Yannakogeorgos[2]
[1]U.S. Naval Postgraduate School, Monterey, California, USA
[2]Air Force Research Institute, Maxwell AFB, Alabama, USA
ncrowe at nps dot edu
slgarfin at nps dot edu
rbeverly at nps dot edu
panayotis.yannakogeorgos at maxwell.af.mil

Abstract

A cyberweapon can be as dangerous as any weapon. Fortunately, recent technology now provides some tools for cyberweapons control. Digital forensics can be done on computers seized during or after hostilities. Cyberweapons differ significantly from other software, especially during their development, and recent advances in summarizing the contents of storage media can locate possible cyberweapons quickly. In addition, use of cyberweapons can be distinguished from the usual malicious Internet traffic by being aimed at targets associated with political, social, and cultural issues that are often known well in advance, and we can monitor those targets. Cyberweapons are relatively unreliable compared to other kinds of weapons because they depend on flaws in software, and flaws can get fixed; cyberweapons therefore require considerable testing, preferably against live targets, and this testing may be observable. So international "cyberarms agreements" could provide for forensics on cyberweapons and usage monitoring. Agreements can also encourage cyberweapons use to be more responsible by stipulating attribution and reversibility. We conclude with a discussion of the kinds of international agreements that are desirable, and examine the recent increasing interest of the United States government in such agreements.

Keywords: cyberweapons, cyberattacks, agreements, monitoring, forensics, reversibility

This paper appeared in the *International Journal of Cyber Warfare & Terrorism*, Vol. 1, No. 1, pp. 1-14.

1. Introduction

Cyberweapons are digital objects that can be used to achieve military objectives by disabling key functions of computer systems and networks. They can be malicious software installed secretly through concealed downloads or deliberate plants by human agents, or they can be malicious data or maliciously delivered data as in denial-of-service attacks. Cyberweapons are a growing component in military arsenals (Libicki, 2007). Increasingly countries are instituting "cyberattack corps" with capabilities to launch attacks in cyberspace on other countries as an instrument of war, either alone or combined with attacks by conventional military forces (Clarke and Knake, 2010). Cyberattacks appeal to many military commanders. They seem to require fewer resources to mount since their delivery can be accomplished in small payloads such as malicious devices or packets that can be primarily delivered through existing infrastructure such as the Internet. They also seem "cleaner" than conventional weapons in that their damage is primarily to data and data can be repaired, although they are difficult to control and usually entail actions close to perfidy, something outlawed by the laws of war (Rowe, 2010). Cyberweapons can be developed with modest technological infrastructure, even

by underdeveloped countries (Gady, 2010) by taking advantages of international resources. So there is a threat of cyberattacks from "rogue states" such as North Korea and terrorist groups that hold extreme points of view, as well as from countries with well-developed cyberweapons capabilities such as China.

Many information-security tools we use today to control threats and vulnerabilities with criminal cyberattacks (Brenner, 2010) help against the cyberweapon threat. Good software engineering practices in design and construction of software, access controls on systems and data, and system and network monitoring for suspicious activity all help. But they are insufficient to stop cyberattacks today because there are ways, albeit challenging, to subvert each of them, and the increasing complexity of cybersystems provides increasing opportunities for finding flaws in software. State-sponsored cyberattacks should be especially hard to prevent because states can exploit significant resources and can use them to develop highly sophisticated attacks. States will likely employ a variety of methods simultaneously to achieve a high probability of success, and will test them considerably more carefully than the hit-or-miss approach of most criminal attacks today. Such challenging state-sponsored cyberattacks will be difficult or impossible to defend against with current information-security defensive techniques.

2. Approach

What can be done against such threats then? We believe that countries must negotiate international agreements similar to those for nuclear, chemical, and biological weapons. Such agreements (treaties, conventions, protocols, and memoranda of understanding) (Croft, 1996) can stipulate the ways in which cyberweapons can be used, as for instance stipulating that countries use cyberweapons only in a counterattack to a cyberattack. Agreements can also stipulate policing of citizens such as "hacker" groups within a country, so that a nation cannot shift blame for cyberattacks and cyberweapons onto them. A few such agreements are in place today for cybercrime, but the growing threat suggests that it is time to plan out what such agreements will entail and how they should be enforced. As an example, the EastWest Institute in the U.S. recently proposed a cyberwar "Geneva Convention" (Rooney, 2011). Deterrence, a key aspect of nuclear weapons control, is not possible with cyberweapons because revealing capabilities significantly impedes their effectiveness.

(Johnson, 2002) was skeptical in 2002 of the ability to implement cyberarms control, citing the difficulty of monitoring compliance. But his arguments are less valid today. Cyberweapons are no longer a "cottage industry" but require significant infrastructure for finding exploits, finding targets, gaining access, managing the attacks, and concealing the attacks. This necessary infrastructure leaves traces even when concealed. The cyberweapon infrastructure needs to be increasingly complex because target software, systems, and networks are increasingly hardened and complex, and because vulnerabilities are being found and fixed faster than ever. Advances in network monitoring make it possible to detect coordinated attacks and remote control of one machine by another as in botnets, since botnets need aggregate effects to be useful to attackers, and aggregate effects can be detected with statistics. Digital forensics has advanced significantly since 2002, making it possible to find many useful things about digital artifacts. Anonymity and encryption techniques that attackers depend upon are easy to see and are good clues to something suspicious. Some techniques central for criminal cyberattacks today such as code obfuscation have little legitimate use and are good indicators of cyberattack development and hence, in the right context, cyberweapons.

Thus many international agreements on cyberweapons could be feasibly monitored despite the challenges. The situation is similar to that with chemical weapons for which there are, for example, many methods for making mustard gas that can use easily available chemicals with legitimate uses. Although proving that a

facility is used for chemical or biological weapons production is difficult, the type of equipment at a facility can provide a good probability that it has been used to manufacture such weapons, as U.N. inspectors realized in Iraq in the 1990s when they discovered evidence of airlocks in alleged food-production facilities. International conventions banning chemical and biological weapons have been effective despite the difficulties of verifying production and stockpiling of such weapons (Price, 1997). We think that similar examinations, and therefore conventions, should be possible in the cyberdomain. For instance, even if developers of cyberweapons delete or hide evidence on their disks, there are often ways to reconstruct it such as finding data deleted but not yet overwritten, data assembly from fragments (Garfinkel, 2006), and examination of magnetic residues.

Cyberinspection technology can have other uses too. It helps law enforcement, military organizations, and intelligence communities within a country in examining captured computer systems belonging to suspected criminals or terrorists for cyberweapons.

We realize that policy is too often driven by crises, so it may take a serious cyberattack to interest a country in negotiating cyberarms limitations. Such a cyberattack is technically feasible (Clarke and Knake, 2010) and could happen at any time. Model agreements can be developed in advance of a crisis. In the meantime, progress can be made by international organizations in negotiating broad cyberarms agreements as well as more specific agreements that can be used against rogue states and organizations.

3. Models for cyberweapons use

Two recent cases provide possible models for future cyberwarfare. One is the cyberattacks on Georgia in August 2008 discussed in (Rowe, 2011). Attacks were launched to coincide with a military invasion of Georgia by Russia (the "South Ossetia War"), and appeared to be well planned and timed. These were primarily denial-of-service attacks against predominantly Georgian government Web sites, including some Web-site defacement (USCCU, 2009). Some of attacking machines were known malware hosts, some were new sites created specifically for the attack, some were botnets of otherwise innocent computers, and some were machines of people recruited to attack from social-networking sites. None of these were government or military sites.

The targets of the attack were government and business organizations in Georgia that were viewed as key in withstanding the conventional military attack by Russia that followed shortly thereafter. They included government agencies associated with communications as well as news-media organizations, apparently with the goal of making it difficult for Georgians to determine what was happening. Later attacks broadened the scope to financial and educational institutions, as well as businesses associated with particular kinds of infrastructure. These cyberattacks were clearly targeted at civilians, and were targeted precisely. Reconstruction of the attacks was possible from a variety of international resources since Internet traffic is routed through many countries. While the attribution of the perpetrators of the attacks does not meet standards of international law, the circumstantial evidence is strong for the involvement of sources in Russia. This conclusion was reached by the U.S. Cyber Consequences Unit by piecing together Internet traffic records (USCCU, 2009).

The other important recent case is the so-called "Stuxnet" worm and corresponding exploits targeting industrial-control systems (Markoff, 2010). These used traditional malware methods for modifying programs. Since Stuxnet targeted systems with no financial incentive, it was most likely developed by an information-warfare group of a nation-state. After the attack, forensic investigators discovered many

distinctive properties of the attack software. For example, it used previously unknown attacks and a variety of concealment methods, and it appears to have targeted a specific industrial control system associated with uranium enrichment. These features were unusual in cyberattacks. Stuxnet was discovered because it spread far beyond its intended target although its damage was highly targeted. This dissemination was necessary to propagate it to its targets, and was a clue to the international community that something was happening. So even though Stuxnet was a highly sophisticated cyberattack, it was recognized quickly by the international community.

4. Technical obstacles

We discuss three key technical challenges to achieving international cyberarms control: (1) locating cyberarms on computers; (2) noticing cyberarms use; and (3) developing more responsible kinds of cyberweapons. Cyberarms also raise important challenges to the laws of warfare that we do not have space to discuss here, including distinguishing a cyberweapon from other malicious software, assigning legal responsibility for a cyberattack, and setting norms for proportional and discriminatory counterattacks (Wingfield, 2009).

4.1 Analysis of drives to find cyberweapons

The U.S. analyzed a number of captured computers and devices in its recent military operations in Iraq and Afghanistan. This was useful in identifying insurgent networks and their interconnections. Similarly, a good deal can be learned about a country's or terrorist group's cyberweapons from the computers used to develop or deploy them. Alternatively, a country may agree to forego cyberweapons as part of a negotiated settlement of a conflict, and may agree to submit to periodic inspections to confirm this (United Nations, 1991).

Detection of cyberweapons might seem difficult. But there are precedents in the detection of nuclear, chemical, and biological weapons (O'Neill, 2010). Cyberweapons development generally requires unusual computer usage in secret facilities since most cyberweapons require secrecy to be effective, which rules out most software development facilities. Clues to cyberweapons can also be found inside computers. Certain types of software technology such as code obfuscation and spamming tools are good clues to malicious intent. Code for known attacks (for providing reuse opportunities) and stolen proprietary code such as Windows source code (for testing attacks) are other good clues. Technologies such as systematic code testers, "fuzzing" utilities, and code for remote control of other computers provide supporting evidence of cyberweapons development though they have some legitimate uses. Data alone can be a clue, such as detailed reconnaissance information on adversary computer networks. Diversity of software techniques is a clue to cyberweapons development because the unreliability of cyberweapons encourages the use of multiple methods. Once suspected cyberweapons are found, they can be studied systematically to confirm their nature using malware analysis (Malin, Casey, and Aquilina, 2008).

A cyberweapons inspection would have to be performed on-site and with automated tools, as a party to a cybermonitoring regime would not allow a potential adversary to remove materials from a secret facility. Cyberweapon monitors would likely be required to use bootable read-only storage that would contain programs to analyze the contents of a computer system and look for evidence of cyberweapon development. Inspection would require a scheme for obtaining temporary use of the necessary passwords and keys for the systems inspected, which could be aided by "key escrow" methods. Inspection regimes should also require "write-blockers" to assure that the monitors did not themselves plant cyberweapons on the systems being

monitored. Other useful ideas from monitoring of nuclear capabilities (O'Neill, 2010) include agreed inspector entry into the inspected country within a time limit, allowed banning of certain inspectors, designation of off-limits areas, and limits on what kind of evidence can be collected.

A good prototype of what can be done in analysis of drives is our work on the Real Data Corpus, our collection of drive images (computer disks, mobile device storage, and storage devices) from around the world. Currently this collection includes 1467 images. Recent work has characterized disks and drives as a whole, including understanding what is distinctive about the files in each of several dimensions such as file size, number of image files, number of deleted files, and number of files frequently edited (Rowe and Garfinkel, 2011). Clusters of files that have no counterpart for others in a corpus are particularly interesting, and can be the focus of more detailed forensic analysis. For quicker analysis, random pieces of files can be selected, and this can be surprisingly good at identifying many types of data (Garfinkel et al., 2010). Deception markers in particular can be sought since illegal cyberweapons development would need to be concealed. Deception could be in the form of oddly named, renamed, or encrypted files, and could be enhanced by other techniques such as changing the system clock or manipulating a log file.

Figure 1 shows an example. This is a histogram of the mean suspiciousness metric per drive of 325 Windows disk drives in our corpus (Rowe and Garfinkel, 2011). The suspiciousness metric included the use of double file extensions, long file extensions, rarity of the file extension, and presence of large numbers of nonalphabetic characters in the file path. These are clues to concealment because double file extensions and nonalphabetic characters suggest concealment of purpose, and long and rare file extensions suggest anomalous usage of a computer. Figure 1 identifies several drives as unusually suspicious, beyond what would be expected in the Poisson distribution of most values, and most of these were in fact suspicious. This analysis only took a few seconds per drive and could save inspectors time in hunting for cyberweapons.

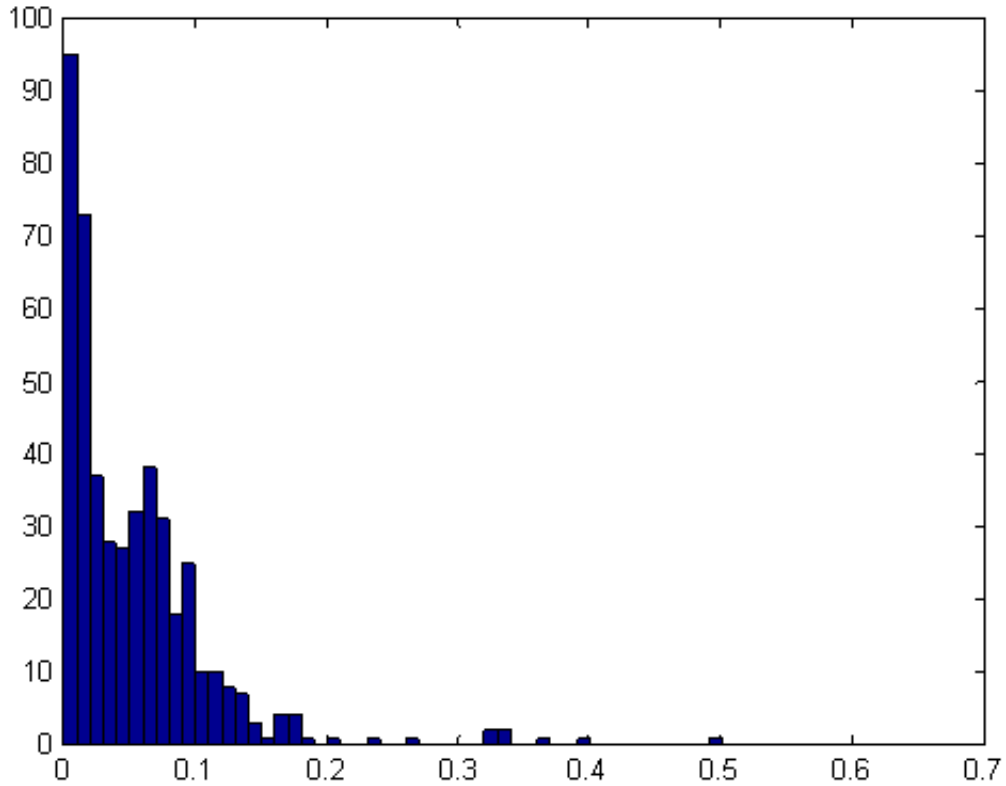

Figure 1: Histogram of counts (vertical axis) of mean suspiciousness (horizontal axis) of files on a set of drives.

Another way to simplify inspections is to cluster the files of a drive into meaningful groups such as images, spreadsheets, and programs, and calculate statistics on groups. That way an inspector could first focus on group differences and not be overwhelmed by large numbers of files. Figure 2 shows an example cross-drive clustering ("superclustering") of individual-disk clusters on the corpus as Figure 1. This plots the two principal components of 32 properties of the clusters where size of the circle indicates the number of clusters in the supercluster. The small superclusters represent file types with few counterparts on other drives, such as unusual software, and thus should be higher priority for investigation. For instance, an investigator should examine files in small superclusters of executable files to have a better chance of finding malicious software.

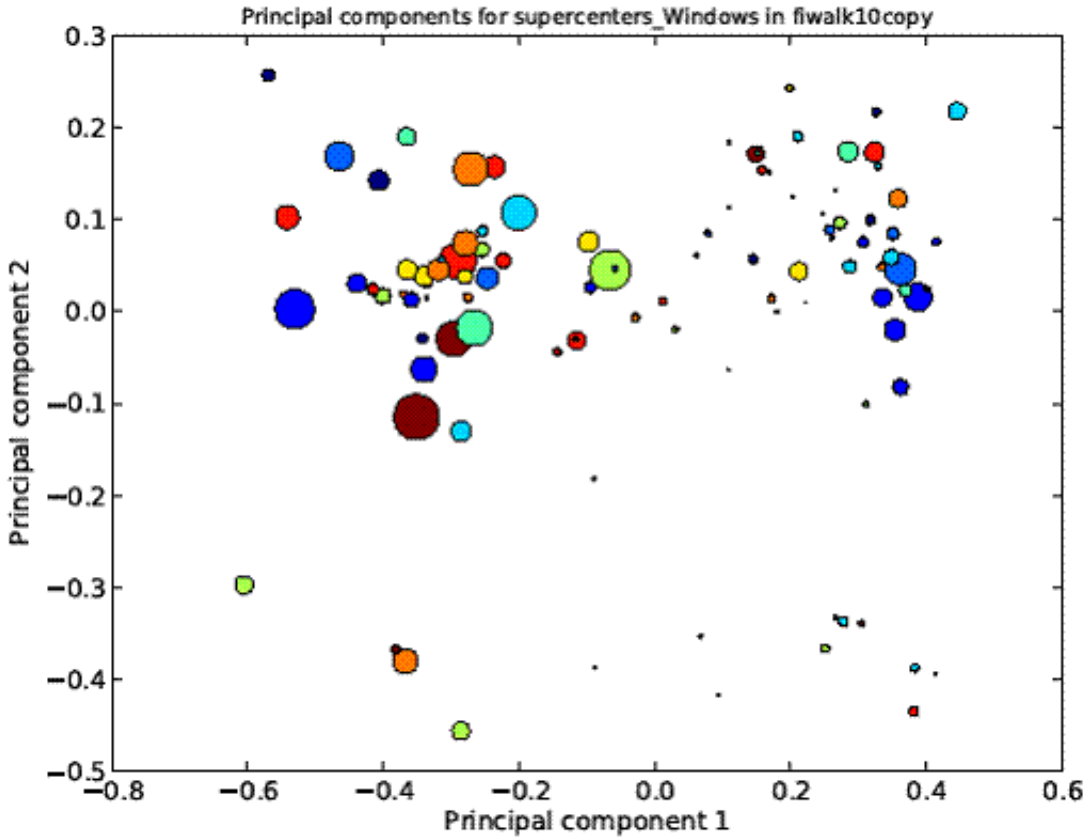

Figure 2: Cross-drive clusters of drive-file clusters on the Figure 1 corpus.

4.2 Network monitoring for cyberweapons

Many tools can discriminate legitimate from abusive network traffic. Such inferential intrusion detection has limitations due to the difficulty of defining malicious traffic in a sufficiently general way without incurring a large number of false positives (Trost, 2010; Sommer and Paxson, 2010). But the attack landscape is different for politically and economically motivated state-sponsored cyberattacks:

1. ***Targets***: State-sponsored attacks will be targeted to particular regions and political agendas, in

contrast to criminal attacks which usually target victims indiscriminately.

2. ***Sophistication***: Cyberarms will be the product of well-funded nations with significant resources. Thus they will use new and sophisticated techniques rather than those of the common attacks we see on the Internet. That means that we can ignore most malicious traffic we see when searching for cyberweapons usage. While some initial stages of cyberweapons activity will be hard to detect--this is why we need international agreements about them--to be useful weapons, cyberweapons must eventually produce a significant effect, and that effect should be easy to see. The Georgia attacks, for instance, were obvious.

3. ***Attribution***: As with conventional warfare, the warring parties will likely follow specified (nondigital) protocols. Protocols will likely dictate that combatants reveal who they are at least in general terms so that the attacks will achieve the desired political effect.

These features provide three kinds of clues to cyberweapons use which we can detect by network monitoring. This does not mean detecting the setup of an attack, merely the active or "attack" phase because these attacks will be sophisticated and stealthy in their setups. Detection of active attacks does require a sufficiently broad deployment of network-traffic vantage points, secured both physically and virtually from tampering, run by an international organization such as the International Telecommunications Union (ITU). One approach to deploying them is to have the vantage points be entirely passive and communicate over separate infrastructure via encrypted and authenticated channels. Centralized collection of data would be efficient for an international organization. Ideally, a vantage point should exist at the ingress to each important network of a country, capable of full-rate traffic processing. If this is difficult, random sampling of traffic can be done. The monitoring infrastructure could be realized via government mandate as it is in many countries today including the United States.

Cyberweapon usage is likely to be quite focused. A cyberweapon might attack a particular country, a type of service (e.g. electrical grid or water systems), or systems used by a certain political, ethnic or religious persuasion. Both the Georgia and Stuxnet attacks employed focused targeting (insufficiently focused according to critics). However, we should also be able to see cyberweapons testing in Internet traffic. That is because potential vulnerabilities and attack vectors will not correlate well with desirable targets, and there must be significant testing, something generally unnecessary for criminal cyberattacks. Also, cyberweapons by their nature are complex pieces of software that include components for penetrating remote systems, controlling the remote systems, and propagating to other systems. Understanding the behavior of a cyberweapon in isolation, or in simulated environments is difficult – the more secret the testing, the less like the real world it will be, and the less accurate it will be at predicting real-world performance. We can see this demonstrated in the poor initial performance of complex new conventional weapons systems such as aircraft. We expect that countries wishing to employ cyberweapons will first unobtrusively try them against real targets to understand their real-world efficacy. An example is the attacks on Estonia in 2007 prior to the attacks on Georgia in 2008. This initial testing provides a clue to forthcoming cyberweapons use.

Thus, detecting pre-hostility events at the network level is possible. It can be aided by metrics for detecting national, political, social, or cultural bias in the targets of malicious network traffic. Standard statistical techniques can suggest that the victims represent a particular political perspective or country's interest more than a random sample would (Rowe and Goh, 2007). For instance, a significance test on a linear metric encoding political or social agendas can provide a first approximation, while the Kullback-Leibler divergence can characterize the extent of difference between expected and observed traffic distributions. How do we

identify the political or social agenda to search for? This requires help from experts on international relations. Nations have longstanding grievances with other nations, and particular issues are more sensitive in some nations than others. We can enumerate many of them and identify associated Internet sites.

This comparison monitoring needs to recognize that cyberattacks are bursty, however, and rates should only be compared during bursts (and there may be no comparable bursts at some targets due to the randomness of targeting). As an example, Table 1 shows the burstiness of attacks on one Internet site we ran. The quantity measured is the number of alerts counted by the Snort intrusion-detection system for consecutive days (the columns) broken down by the category of alert (the rows). The ICMP bursts in the third through sixth days, the INFO bursts on the fourth and fifth days, and the SHELLCODE bursts on the third, seventh, and eighth days are likely the only data worthy of comparison between sites. Both the number of total attacks in each burst and the number of distinct bursts are useful metrics.

Table 1: Example daily counts of alerts of malicious traffic on a honeypot.

Snort Alert Daily Totals, Jan29 – Feb4	312	329	11538	18602	17028	46437	339	312	226	226	435
BAD-TRAFFIC	0	0	0	0	0	0	0	0	0	0	0
ICMP	279	289	11466	15618	16727	46397	282	253	198	206	395
INFO	0	0	1	2952	245	0	1	0	0	0	0
MS-SQL	28	31	40	26	36	19	20	24	22	16	25
NETBIOS	4	7	2	1	7	5	10	8	3	2	6
SCAN	0	0	0	0	0	0	0	0	0	0	0
SHELLCODE	1	2	29	1	13	16	26	23	3	2	9
WEB-IIS	0	0	0	0	0	0	0	1	0	0	0
WEB-PHP	0	0	0	0	0	0	0	0	0	0	0

Other broad properties of the observable network traffic can provide precursors to attack such as the number of packets, the number of bytes transferred, the size of an average "flow" (set of related packets), the frequency of flows, and so on (Munz and Carle, 2007). Measurement of relatively crude properties works well in tracking and analyzing attacks supported by amateurs such as the Chinese hacker groups that are harnessed to attack Western organizations at times of political or social grievances against them (Hvinstendahl, 2010). Feature selection methods in finding discriminating network traffic features (Beverly and Sollins, 2008) can provide a more rigorous basis for choosing more sophisticated properties. We also can look for particular sequences of events indicative of a systematic attack, say a broadcast of many footprinting packets followed by more specific footprinting, something not seen much in criminal cyberattacks.

An additional tool useful in detecting cyberweapons development is a decoy, a site deliberately designed to encourage attacks. A decoy can be designed to be more useful than a normal site by narrowing its content to just that necessary to invoke a response. For instance, for the Georgia attacks it would have been useful to monitor decoys giving government announcements. Decoys need to be situated in plausible Internet sites,

however, so that a government decoy is on a government computer system. We need to then design "differential honeypots" that compare attacks on a decoy with those on a similar non-decoy system. A decoy can also be equipped with more detailed monitoring of its usage that would not be possible for most sites, and should use honeypot technology to implement attack resilience and intelligence-gathering capabilities that are not easily disabled. Decoys do not generally raise ethical concerns because they are passive, but guidelines should be followed in their use (Rowe, 2010) since decoys are also used by phishers.

Data fusion on World Wide Web usage can complement our network monitoring. If a country's government shows a sudden increase in visits to hacker Web sites, it may also suggest cyberweapons development since such activity is knowledge-intensive.

Finally, the aforementioned forensics techniques can enhance network monitoring. For instance (Beverly, Garfinkel, and Cardwell, 2011) showed the presence of residual network packets on nonvolatile storage may be correlated with observed traffic and attacks.

4.3 Encouraging more-responsible cyberweapons

International agreements can also stipulate acceptable types of cyberweapons. Two important aspects of this are attributability and reversibility of attacks. For attribution, a responsible country will find it in their interests to make their attacks clear in origin to better enable desired political and social effects of an attack, which are often more important than the actual military value. The ability of the USCCU to trace the Georgia cyberattacks back to people in Russia says that Russia was sending a political message to Georgia. Contrarily, it could be useful to a country to be able to prove it was *not* the source of a cyberattack for which it is being blamed. Attribution can be done by using digital signatures attached to attack code or data, identifying who is responsible for an attack and why. They could be concealed steganographically (Wayner, 2002) to avoid giving advance warning to the victim that they are being attacked, but allowing it to be demonstrated later to the international community. For attacks without code like denial of service, a signature can be encoded in the low-order bits of the times of the attacks.

Nations should also be encouraged to use attack methods that are more easily repairable, following the same logic behind the design of more easily removable landmines. (Rowe 2011) proposed four techniques that can be used to make cyberattacks that are easier to reverse by the attacker than by the victim even when the victim tries to restore from backup (Dorf and Johnson, 2007). As illustrated in Figure 3, four methods are: (1) "locking up" the operating system of the victim's computers by encryption of key software by the attacker, where the victim does not have the key to decrypt it; (2) obfuscation of a victim's system by the attacker by data manipulations that are hard to decipher yet algorithmic and reversible (such as turning document "document" into "stnemucod" by reversing its bits); (3) intercepting and withholding by the attacker of key information that is important to the victim, while saving it in backup; and (4) deception by the attacker of the victim to make them think their systems are not operational when they actually are. In the first two cases, reversal can be achieved by software operations by the attacker; in the third case, the attacker can restore missing data; and in the fourth case, the attacker can reveal the deception. Note that reversal can be done at a distance so the attacker does not necessarily require visiting the victim's territory.

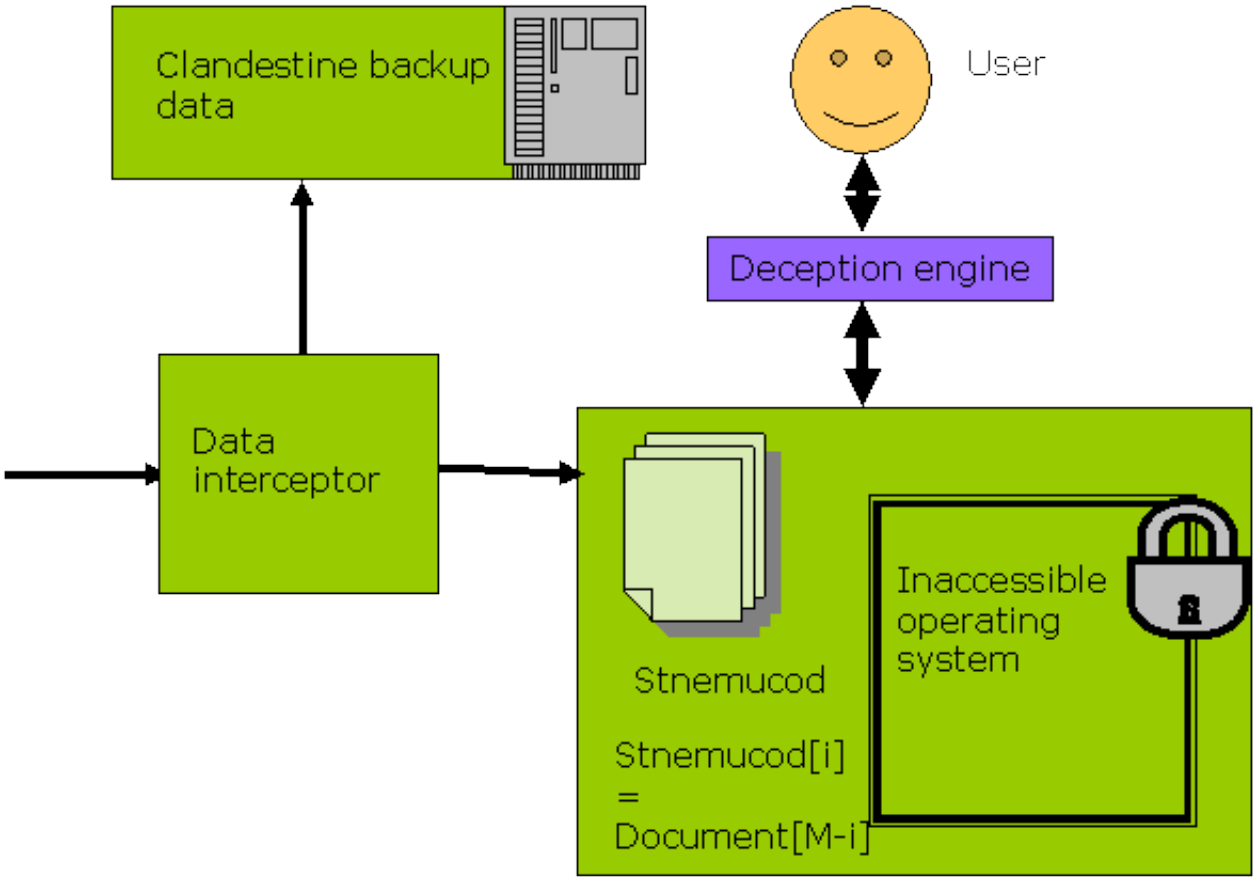

Figure 1: Four ways to mount reversible cyberattacks.

How do we encourage attackers to use reversible attacks? There are several possible incentives. One would be if the attacker will eventually need to pay reparations, as the United Nations could stipulate as part of a negotiated settlement of a conflict (Torpey, 2006). Even in an invasion or regime change, it is likely that the impacts of cyberweapons will need to be mitigated—indeed, the perceived possibility of mitigation will likely drive the adoption of cyberweapons. Another incentive comes from international outcry at using unethical methods and the resulting ostracism of the offending state, as with the use of biological weapons. Another incentive is if a victim can respond in like kind, wherein use of a reversible attack could encourage an adversary to do the same to avoid appearing to escalate the conflict (Gardam, 2004). Also, nonreversible attacks may in the future be interpreted as violating the laws of warfare in regard to unjustified force when

reversible methods are easily available. Responses of the international community to analogous such violations with traditional arms include sanctions, boycotts, fines, and legal proceedings (Berman, 2002).

4. Support for international cyberarms cooperation in the United States

Many of the ideas mentioned here benefit from international cooperation, but obtaining such cooperation has been difficult. We focus here on the role of the United States. Until recently, the United States would not discuss international cooperation in matters of military cybersecurity (Yannakogeorgos 2010; Yannakogeorgos 2011). This policy began to shift in early 2010. It is now recognized that although a cyberarms control treaty is most desirable, negotiating it will take decades since it requires agreement on global norms of behavior built from multilateral institutions of diplomacy.

Embryonic global norms found in international agreements dealing with cybercrime can serve as models for cyberarms control. The Council of Europe Convention on Cybercrime, adopted in November 2001, seeks to align domestic substantive and procedural laws for evidence gathering and prosecution, and to increase international collaboration and improve investigative capabilities for coordinating E.U. efforts on cybercrimes. Adopted and ratified by the U.S. in 2007, it is considered a model law for the rest of the world. The UN General Assembly and the World Summit on the Information Society Declaration of Principles endorsed a global culture of cybersecurity that is promoted, developed, and implemented in cooperation with all stakeholders and international expert bodies. The ITU and U.N. General Assembly have also passed several resolutions addressing the criminal misuse of information. The efforts of the ITU have culminated in the International Multilateral Partnership against Cyber Threats (IMPACT) although the United States does not currently support it. IMPACT is a Global Response Centre based in Cyberjaya, Malaysia. It was set up in 2009 to serve as the international community's main cyberthreat resource by proactively tracking and defending against cyberthreats. The center's alert and response capabilities include an Early Warning System that enables IMPACT members to identify and head off potential and imminent attacks before they can inflict damage on national networks.

A major obstacle to any international agreement is the concept that states need to acknowledge responsibility for malicious cyber actions within their borders. Several recent studies of cyberespionage, and some corporate investigations, have traced recent attacks on the U.S.'s commercial infrastructures to China (Areddy, 2010). Denying its official involvement, the government of China bemoaned its fate as the greatest victim of cybercrime. (The Chinese definition of cybercrime includes content, and thus using social networks to mount revolutions would be considered a crime in China, whereas the U.S. considers such actions as part of democracy.) The individuals responsible were not caught, and China received only a slap on the wrist via a State Department "note verbal" and the launching of the Department of State's Internet Freedom Agenda in 2010. Recent policing of internal hacker groups in China has not improved. Increasing the consequence of a state for cyber attacks originating within their territory is an appropriate course of action for the United States.

Catalyzed in part by events involving the Google corporation in China, and discussions with the Russian delegation in January 2010, senior leaders began to talk publicly about global norms of behavior in cyberspace, including military cyberspace (Lynn, 2010). Formal shifts in policy began in 2011, with the *National Military Strategy* identifying the cyber threat as being "expanded and exacerbated by lack of international norms, difficulties of attribution, low barriers to entry, and the relative ease of developing potent capabilities." (USDOD, February 2011). Subsequently the Department of Defense Strategy for Operating in Cyberspace stated that DOD "will work with interagency and international partners to encourage responsible

behavior and oppose those who would seek to disrupt networks and systems, dissuade and deter malicious actors, and reserve the right to defend these vital national assets as necessary and appropriate." (USDOD, July 2011). These are steps in the right direction towards a U.S. international cyber policy that holds states responsible, and represents high-level acknowledgement that in addition to being a law enforcement, diplomacy, and development issue, cyber is also a military issue. However, the U.S. government appears to be playing a game of forum picking with "like-minded states," rather than supporting international initiatives already underway at the ITU. Indeed, there is no mention of elements related to cybersecurity within the World Summit for the Information Society *Tunis Agreement* and *Geneva Action Plans*, work of the High Level Experts Group of the Global Cybersecurity Agenda, and the ITU's IMPACT program. Further refinement of U.S. cyber policy is required to enhance future positions within international cyberarms control discussions.

So the U.S. international cyber strategy currently encourages the development of global norms, while fighting the institutionalization of cyber issues within preexisting frameworks that have been under development for the past decade. This is done to avoid diplomatic hurdles within them on the grounds that there are challenges posed to U.S. norms of openness "by some governments and international institutions intent on imposing pre-Internet-era telecommunications regulatory schemes to provide them control over the flow of information (and money) they enjoyed in the old days of the monopoly phone company" (Kornbluh and Weitzner, 2011). The alternative of setting up parallel dialogues for norms discussion could be unworkably complex. By extension, the U.S. aspiration to lead the world in setting global norms could fail as others fail to follow. The result could be that the U.S. could lose an opportunity to focus its diplomatic resources on the parallel structures that the Europeans, Russians and Chinese have forged over the past decade with the United Nations and its specialized agency, the ITU. Like-minded states are important allies within diplomatic forums. Creating an alliance of cyber security with state and private-sector partners that could push the norms discussion within the ITU would serve U.S. interests well in what will likely become heated diplomatic debates. However, those states that are resisting this talk of state responsibility within cyberspace will still need to be compelled in one way or another to cooperate in investigating cyber attacks. The U.S. should begin documenting and issuing reports on the overall capacity of each nation's efforts to both create and enforce legal mechanisms within their countries to assure their people can be prosecuted, and also to measure to what extent said state is cooperating in investigations. On this basis, international arms control agreements will be guided by norms and customs of behavior that have been qualified by years of documentation.

Criminal prosecution of a nation's hacker groups by its government could be an important detail in the stipulations of agreements. For instance, when Philippine hackers in 2000 launched a virus that attacked computers worldwide and the Philippine government was initially unhelpful, improvements under international pressure were subsequently made by it, both legally and managerially, to enable a better response in the future. Other possible agreements could follow those of traditional arms control, as for instance a commitment to use cyberweapons only in self-defense, or agreed export controls on cyberweapons technology. We do need to make legal distinctions between cybercrime, cyberconflict, cyberespionage and cyberterror as this is necessary when creating a regulatory regime for cyberweapons (Wingfield, 2009). One model that could be studied is the Wassenaar Arrangement for export controls, which could be extended to information technology products.

Other specific technical details can be negotiated as part of cyberarms agreements. An example would be a mandate for countries to use IPv6 instead of IPv4 to enable better attribution of events on the Internet; rogue states could be told that they cannot connect to the Internet unless they use IPv6. Other mandates could

stipulate architectures in which attribution of traffic is easier such as minimum requirements on persistence of cached records. Others could prohibit less-controllable attacks such as worms and mutating viruses, to achieve better discrimination of military from civilian targets in cyberattacks (Shulman, 1999).

5. Conclusion

Cyberarms agreements have been said to be impossible. But technology is changing that. We can seize and analyze drives on which cyberweapons were developed; we can detect attacks and the necessary testing of cyberweapons; we can create incentives for self-attributing and reversible cyberattacks; and we can develop and ratify new kinds of international agreements. While we cannot stop cyberweapons development, we may be able to control its more dangerous aspects much as we control chemical, biological, and nuclear weapons, and limit it to responsible states. It is time to consider seriously the possibility of cyberarms control.

References

Areddy, J. (2010, February 20). *Wall Street Journal*, A1.

Baron, K. (2011, July 14). U.S. cyber defenses 'way too predictable' says Cartwright. *Stars and Stripes Central*. Retrieved July 15, 2011 from www.strips.com/blogs/stripes-central/stripes-central-1.8040.

Berman P. (2002). The globalization of jurisdiction. *University of Pennsylvania Law Review, 151* (2), 311-545.

Beverly, R., & Sollins, K. (2008). An Internet Protocol address clustering algorithm. In *Third Workshop on Tackling Computer Systems Problems with Machine Learning Techniques*. Berkeley, CA: Usenix.

Beverly, R., Garfinkel, S., & Cardwell, G. (2011). Forensic carving of network packets and associated data structures. In *Digital Forensics Workshop*.

Brenner, S. (2010). *Cybercrime: criminal threats from cyberspace.* Santa Barbara, CA: Praeger.

Clarke, R., & Knake, R. (2010). *Cyberwar: the next threat to national security and what to do about it*. New York, NY: HarperCollins.

Croft, S. (1996). *Strategies of arms control: a history and typology.* Manchester, UK: Manchester University Press.

Dorf, J., & Johnson, M. (2007). Restoration component of business continuity planning. In Tipton, H., and Krause, M. (Eds.), *Information security management handbook, Sixth Edition*, (pp. 1645-1654). Boca Raton, FL, US: CRC Press.

Erbschloe, R. (2001). *Information warfare: how to survive cyber attacks.* Berkeley, CA: Osborne/McGraw-Hill.

Gady, F.-S. (2010, March 24). Africa's Cyber WMD. *Foreign Policy*.

Gardam, J. (2004). *Necessity, proportionality, and the use of force by states.* Cambridge, UK: Cambridge

University Press.

Garfinkel, S. (2006, September). Forensic feature extraction and cross-drive analysis. *Digital Investigation, 3*, Supplement 1, 71-81.

Garfinkel, S., Roussev, V., Nelson, A., & White, D. (2010). Using purpose-built functions and block hashes to enable small block and sub-file forensics. Ithaca, NY: Digital Forensics Research Conference.

Ghernouti-Helie, S. (2010). *A national strategy for an effective cybersecurity approach and culture.* New York, NY: IEEE Press.

Johnson, P. (2002). Is it time for a treaty on information warfare? In Schmitt, M., and O'Donnell, B., *Computer Network Attack and International Law (International Law Studies Volume 76)* (pp. 439-455). Newport, RI, US: Naval War College.

Hvistendahl, M. (2010, March 3). China's hacker army. *Foreign Policy*.

Kornbluh, K., & Weitzner, D. (2011, July 14). 21st century statecraft: foreign policy of the Internet. *The Washington Post.*

Libicki, M.(2007). *Conquest in cyberspace: national security and information warfare.* New York, NY: Cambridge University Press.

Lynn, W. (2010, September/October). Defending a new domain: the Pentagon's cyberstrategy. *Foreign Affairs*.

Malin, C., Casey, E., and Aquilina, J. (2008). *Malware forensics: investigating and analyzing malicious code.* Waltham, MA: Syngress.

Markoff, J. (2010, September 26). A silent attack, but not a subtle one. *New York Times*, A6.

Mel, H., & Baker, D. (2000). *Cryptography decrypted, 5th edition.* Boston, MA: Addison-Wesley Professional.

Munz, G., & Carle, G. (2007, May). Real-time analysis of flow data for network attack detection. In *10th IFIP/IEEE Intl. Symposium on Integrated Network Management* (pp. 100-108). New York: IEEE.

O'Neill, P. (2010). *Verification in an age of insecurity: the future of arms control compliance.* New York, NY: Oxford.

Price, R. (1997). *The chemical weapons taboo.* Ithaca, NY: Cornell University Press.

Rooney, B. (2011, February 4). Calls for Geneva Convention in cyberspace. *Wall Street Journal*.

Rowe, N. (2010). The ethics of cyberweapons in warfare. *Journal of Technoethics, (1)*1, 20-31 [JTE].

Rowe, N. (2011). Towards reversible cyberattacks. In J. Ryan (Ed.), *Leading Issues in Information Warfare*

and Security Research, Volume I (pp. 145-158). Near Reading, UK: Academic Publishing International.

Rowe, N., & Garfinkel, S. (2011) Finding anomalous and suspicious files from directory metadata on a large corpus. *3rd International ICST Conference on Digital Forensics and Cyber Crime*, Dublin, Ireland.

Rowe, N., & Goh, H. (2007, June). Thwarting cyber-attack reconnaissance with inconsistency and deception. In *Eighth IEEE Information Assurance Workshop* (pp. 151-158). New York: IEEE.

Shulman, M. (1999). Discrimination in the laws of information warfare. *Columbia Journal of Transnational Law, 37,* 939-968.

Sommer, R., & Paxson, V. (2010). Outside the closed world: on using machine learning for network intrusion detection. In *IEEE Symposium on Security and Privacy.* New York: IEEE.

Torpey J. (2006). *Making whole what has been smashed: on reparations politics.* Cambridge, MA: Harvard University Press.

Trost, R. (2010). *Practical intrusion analysis.* Upper Saddle River, NJ: Addison-Wesley.

United Nations (1991). *Final document: third review conference of the parties to the convention on the prohibition of the development, production, and stockpiling of bacteriological (biological) and toxin weapons and on their destruction.* BWC/DONF.II/23, Geneva, Switzerland.

USCCU (United States Cyber Consequences Unit) (2009, August). US-CCU special report: Overview by the US-CCU of the cyber campaign against Georgia in August of 2008. Retrieved November 3, 2010 from www.usccu.org.

USDOD (United States Department of Defense) (2011, February). National military strategy. Retrieved September 11, 2011 from https://hsdl.org./hslog/?q-node/5994.

USDOD (United States Department of Defense) (2011, July). Strategy for operating in cyberspace. Retrieved September 11, 2011 from www.defense.gov/news/ d20110714cyber.pdf.

USGAO (United States Government Accountability Office) (2010, March 5). Cybersecurity: progress made but challenges remain in defining and coordinating the comprehensive national initiative. Washington, D.C., US: Government Accountability Office.

Wayner, P. (2002). *Disappearing cryptography: information hiding: steganography and watermarking.* San Francisco, CA: Morgan Kaufmann.

Wingfield, T. (2009). International law and information operations. In Kramer, F., Starr, S., and Wentz, L. (Eds.), *Cyberpower and National Security* (pp. 525-542). Washington DC: National Defense University Press.

Yannakogeorgos, P. (2010, October). Cyberspace, the new frontier - and the same old multilateralism. In Reich, S., *Global Norms, American Sponsorship and the Emerging Patterns of World Politics.* Houndsmills, UK: Palgrave.

Yannakogeorgos, P. (2011). Promises and pitfalls of the U.S. national strategy to secure cyberspace. Carlisle, PA, US: Army War College.

The views expressed are those of the authors and do not represent those of any part of the U.S. Government.