A Case Study in Ethical Decision Making Regarding Remote Mitigation of Botnets

David Dittrich[1],
Felix Leder[2], and
Tillmann Werner[2] *

[1] University of Washington, Seattle WA 98195, USA
[2] Institute of Computer Science IV, University of Bonn, Germany

Abstract. It is becoming more common for researchers to find themselves in a position of being able to take over control of a malicious botnet. If this happens, should they use this knowledge to clean up all the infected hosts? How would this affect not only the owners and operators of the zombie computers, but also other researchers, law enforcement agents serving justice, or even the criminals themselves? What dire circumstances would change the calculus about what is or is not appropriate action to take? We review two case studies of long-lived malicious botnets that present serious challenges to researchers and responders and use them to illuminate many ethical issues regarding aggressive mitigation. We make no judgments about the questions raised, instead laying out the pros and cons of possible choices and allowing workshop attendees to consider how and where they would draw lines. By this, we hope to expose where there is clear community consensus as well as where controversy or uncertainty exists.

1 Introduction

The first distributed denial of service (DDoS) attacks occurred more than 10 years ago, in the summer of 1999 [7]. These were relatively small attack networks by today's standards, ranging from several hundred to more than two thousand computers. Even at those small sizes, these attack networks were capable of disrupting some of the largest educational and commercial service providers in existence for hours up to days at a time. The motivation for these attacks started out at the level of electronic *drive-by shootings* that were primarily over petty fights on Internet Relay Chat (IRC) channels. That soon shifted to extortion against online gambling sites as early as 2001 [26] and online pornography sites as early as 2003 [35], attacks against commercial competitors as early as 2003 [29], and politically-motivated attacks against national infrastructures in 2007 [2]. Perhaps just as frightening, if less apparent, are highly targeted attacks using small and subtle botnets used for less obvious attacks than brute-force denial of service [9].

Not only are malicious attack networks (or *botnets* as they are commonly known) capable of pure disruption of services, but they also cause harm to both companies and individuals through fraud, identity theft, abuse of computer and network resources and other violations of personal privacy. Some botnets remain under hostile control for many months. This is driving researchers, security product and service vendors, and professionals in the security community to express growing frustration. While we focus in this paper on the former group – *researchers* – the same issues and challenges apply to the latter groups as well. Some in the general public perceieve a lack of visible action by law enforcement agencies or the private sector to stem malicious activity. Their frustration motivates calls for the right to fight back, as if this were an issue of *self-defense* against someone throwing punches.

We acknowledge that research of cybercriminal activity involves ethical choices, legal restrictions, liability concerns, as well as challenging political questions. We also acknowledge that each society and culture has its own norms and laws that must be considered when trying to deal with issues that are global in scope. Our primary goal in this work is to illuminate as many ethical issues as is possible surrounding alternatives for aggressively mitigating today's massive and highly robust distributed attack networks, allowing the reader to draw their own conclusions about what actions are or are not appropriate.

There are many different ethical codes and standards that apply to a greater or lesser degree to professional and academic activities, however that does not mean that any one of these codes or standards are sufficient to guide computer security researchers. [13] For example, Institutional Review Boards (IRBs) in the United States are commonly cited, however IRBs are focused on protection of human subjects of biomedical and behavioral research, only apply to research involving humans, and provide little in the way of guidance for developing new research protocols. Professional standards, industry standards, and the Internet Activities Board's best practices all have limitations. IRBs have a limited form of enforcement (in that they can refuse to approve applications and thus halt research they deem harmful), while the rest leave enforcement to unspecified authorities or membership-specific ethics boards.

Ethicists such as Markham suggest considering *ethic as method* and making conscious decisions about research methodology that reflect one's intentions and their source of, "consciousness, mindfulness, honesty, and sensitivity." [27] In discussing this topic Markham suggests researchers ask themselves self-reflective questions, perhaps along the lines of: "What is the intent in performing this research? Who is the stakeholder being served? How would this stakeholder view my actions and interpret my intent? Would they feel grateful, neutral or resentful?"

Proposing a complete new framework for designing ethical research protocols goes well beyond the scope of a case study. Rather than using a more formal method of analysis [12] and making judgments, we borrow and extend some analytic tools from other domains. By applying them to the specific area of

computer security research involving *criminal botnets* we aim to get to the issues, not the answers.

We next review our cases, delve into the entities and ethical issues involved, then conclude with a call for a thoughtful dialog.

2 Storm, Conficker, and beyond

2.1 Storm

In April 2007, Holz, et al, at the University of Mannheim [20], performed Storm botnet enumeration experiments in which they infiltrated the Storm botnet and used features of the distributed hash table (DHT) that is used by Storm to enumerate the bots. They were able to observe the effect of other researchers who were simultaneously doing their own enumeration experiments, and specifically noted UCSD and Georgia Tech (among other unnamed sites) as being observable participants in the Storm botnet. They discuss two attacks – eclipsing, or *Sybil attack*, and poisoning – that could be performed to degrade or render inoperable the Storm botnet. Both could be argued to be positive outcomes. While not stated by Holz, these two attacks would also not have negative effects on the owners of compromised computers. While potentially disabling the botnet, at least temporarily, these attacks do nothing to help mitigate the botnet by assisting in cleanup efforts of individually compromised hosts.

On December 29, 2008, researchers from the University of Bonn and the RWTH Aachen University presented a talk at the 25th Chaos Communication Congress (25C3) in Germany on "0wning the Storm botnet." This research was inspired by the Storm enumeration research at the University of Mannheim. The group demonstrated how knowledge gained from reverse engineering the Storm botnet's command and control (C&C) protocol allowed them to take control of Storm nodes. They showed how Storm bots could be commanded to download and replace Storm with *any chosen binary executable*. Such reverse engineering is required for comprehensive understanding of emerging malware threats [14, 22, 20, 5, 4]. Partial source code for their program that implements the counter-attack on the Storm botnet (named *Stormfucker*) was released on the `full-disclosure` mailing list. In their 25C3 presentation, and an interview following the conference [8], they caution that affecting compromised computers is illegal in many countries, but speculate that someone who resides in a country where there are no laws preventing such action might use the knowledge embodied in the released code to dismantle the Storm botnet, or complete their own working code and publish it. They reasoned that publication could have the positive effect of informing the owners of infected computers, who were likely unaware of these infections, could clean up their *zombie* computers. This work was not presented in an academic setting. Had it been, a program committee may have provided anonymous feedback and/or initiated more public discussion of the ethical principles that could justify attempting to clean up thousands of infected computers (e.g., offering guidance such as Denning [10] or Spafford [36]

that could help guide those with access to the source code in deciding how to use it.)

Two of the Bonn researchers presented this research at a conference at the Cooperative Cyber Defence Centre of Excellence in Tallinn, Estonia, in June 2009. The abstract of their talk [25] "asks urgently for political discussions about authorization and legal feasibility" of taking offensive measures to clean computers without their owners' knowledge or consent, and argues that, "pro-actively fighting botnets requires immediate political and international consensus."

2.2 Conficker

Conficker (a.k.a. Downadup) version .A was first reported to have been found in the wild on November 21, 2008 [39]. Conficker.A exploits the Windows RPC vulnerability MS08-067 for propagation and uses a set of 250 randomly generated domain names as C&C rendezvous points. On December 29, 2008 (38 days later), version .B was released which added more propagation methods targeting hosts on the local and remote networks, as well as blocking access to Microsoft's patching servers, AV companies and other mitigation tool web sites. More notable was its switch from use of the SHA-1 hashing algorithm to MD-6, released on October 27, 2008 (64 days prior). A third version, .C[3], was observed on February 20, 2009 which also implemented a limited peer-to-peer (P2P) protocol for command and control and moved to random selection of a daily rotating subset of up to 50,000 domain names. 12 days later an updated .C release occurred that fixed a bug in MD-6 that was only publicly announced 16 days earlier. Another analysis of Conficker [24] was released on March 30, 2009, that described some weaknesses in Conficker that allowed for remote infection scanning. Again, a new release of Conficker.D on April 8, 2009 (8 days after [24]) rendered the first scanning method useless and required major changes to the scanners in order to stay effective. This shows ample evidence that the authors of Conficker are studying publications about Conficker and are capable of quickly responding when they wish. Furthermore, it illustrates the arms race that exists between open publication of defense methodologies and reactive counter-measures by attackers.

Conficker infected nodes have only been observed to attempt its HTTP-based update protocol via its domain generation algorithm (DGA). An active defense mechanism against those update attempts is the sinkholing of domains performed by the Conficker Working Group in cooperation with registrars all over the world. The power of this approach prompted Conficker's authors to add P2P functionality in order to be able to perform updates by another means. Eventually, the .D update was pushed using the new protocol showing researchers and defenders that sinkholing, while necessary, was not sufficient to completely stop updates. Again, this illustrates how information made public can degrade defensive mechanisms and hinder the ability of defenders to monitor malicious activity.

[3] Following the naming scheme in [24].

There have been no major releases of Conficker since the .D release, however millions of infected hosts remain active on the internet and DNS-based mitigation methods continue to be pursued and research into Conficker continues. A detailed analysis of Conficker.C's P2P algorithm was done by SRI and released on September 21, 2009 [33]. This analysis discusses several technical aspects of the design that have not been publicly discussed to date.

2.3 Alternative Countermeasures

In the examples just presented, there are several alternative means for trying to counter or mitigate these advanced threats. The encryption mechanisms in advanced bots like Nugache [14] and Conficker are sufficiently robust to prevent taking over the C&C channel and directly controlling the bots: the bots will ignore commands without proper signatures. Storm, on the other hand, was weak enough that someone could control the bots. It is likely that all three had programming vulnerabilities that could be exploited to attack via buffer overflow errors, etc., allowing the running bot to be hijacked.

Infected bots can be identified in one of several ways. One can passively monitor botnet activity to learn which peers are active; one might be able to write a crawler that can walk the botnet and enumerate all active bots; or one can scan for active bots that are listening for such connections. Of course NAT and firewalls can limit the ability to reach a subset of bots, which may limit the ability to communicate with bots to only using the in-band C&C channel (which may be hardened to the point that it is not usable). The fact that the entire infected population cannot be reached at any given moment means that there is no absolute *strike once* potential for completely taking the botnet out of the hands of the criminal.

There are two primary ways to attempt to remotely mitigate them (i.e., *clean up infected hosts or disable the malware* by exploiting weaknesses: some form of *targeted* attack that uses *hit lists*, or some form of *autonomous, self-propagating* mechanism like those used for other worms. The former method can be controlled very precisely, limiting its scope, rate, and timing. The latter method is typically less predictable, more prone to secondary side-effects on network infrastructure, and very indiscriminate. Staniford, et al [37] describe various methods to speed up worm propagation that could also be used to more precisely target and control worms (e.g., localized scanning, hit-list scanning, and topological scanning.) In terms of ethical principals, *proportionality* requires that actions be properly targeted (not indiscriminate) and the *Defense Principal* requires the actions be necessary for repel or prevent harm directed at the entity taking action. Once an anti-worm is spreading autonomously *in-the-wild*, the propagation effects may be impossible to predict or control.

An informal poll of university system administrators in 2003 [11] found 80% of the 76 respondents believed an autonomous *white worm* was unethical. When asked whether a more targeted (non-worm) method of worm mitigation used by Laurent Oudot to clean up Blaster infections within networks over which he had responsibility [31] was described, the response flipped and only 20%

believed Oudot's method was unethical. This remaining hesitancy was partly due to the final command shown in the article (`shutdown -r -f -t 0 exit`) which immediately reboots the computer without the owner/operator's knowledge or consent.

At the less aggressive end of the spectrum are actions that researchers can take to try to identify those controlling malicious botnets. Crawling malicious P2P networks without leaving noticeable traces or at least attempting to conceal this activity, may allow identification of computers used to control the botnet without impacting law enforcement investigations or affecting enumeration activity of other researchers.

3 Ethical questions raised

As we have seen, today's sophisticated botnets pose a serious threat that is difficult to mitigate through end-user action alone, especially when those end-users do not possess the knowledge, skills, or tools to allow them to easily and effectively counter advanced malware. We have chosen to focus this case study on resilient botnets for several reasons.

First is the inability of the average computer user to either protect themselves against malware infection through social engineering attacks, or effectively respond when attacked. Even if it were possible to inform users that their computer was infected with resilient malware, it is extremely difficult for them to effectively cleanup the infection without resorting to wiping and re-installing the entire operating system, enlisting the help of costly expert assistance, and taking many hours or days of down-time to complete the task. This places the emphasis on finding ways of helping users who cannot help themselves.

The concept of the *Active Response Continuum* was developed to describe the problems resulting from differences in *capacity to respond* and in *aggressiveness of actions taken* to counter wide-spread malicious attack. [15, 11] Both of these concepts are useful for this discussion and are adapted for use here in Tables 1 and 2. In terms of the ARC, most computer users operate at *Level 0* and a lesser number only operate at *Level 1* or higher. In addition, it must be mentioned that the use of protective software alone does not save users from getting infected because none of the known products has a detection rate of 100% [3]. They typically miss several hundreds of thousands of malware specimens.

The second interesting issue is that service providers and enterprises who manage computer systems for thousands or millions of users are capable of operating at higher ARC *Level 3*, but are often prevented (for various reasons, mostly non-technical) from being able to individually assist all infected users and/or cleanup the computers by hand. As malware gets more sophisticated and resilient to detection and mitigation, the problem grows.

Security researchers are also continuing to improve their skills, to the point where it is now common to obtain enough information to control a botnet and its infected hosts [28, 18, 38, 20, 22, 32, 23, 8, 9].

Level	Victim Posture	Characteristic Actions
0	Unaware	No activity: passively rely on system to stay functional
1	Involved	Use and maintain protective software and hardware
2	Interactive	Modifies software and hardware in response to attacks
3	Cooperative	Implements joint traceback and investigation with other victims
4	Non-cooperative (Active Response)	Invasive tracebacks, controlling malware infected hosts, cease-and-desist measures, retaliatory counter-strike

Table 1. Levels of *Capacity*. (Original source: [15])

Level	Impacts	Characteristic Actions
Benign	Limited to victim's own systems	Sniffing, scanning, re-addressing hosts, honeypots
Intermediate	Impacts on remote systems, but not calculated to produce damage	Invasive tracebacks, remote evidence collection, interaction with (controlling) malware
Aggressive	Impacts calculated to alter function of remote system or affect integrity	Remote exploitation, corruption of data, patching, re-installation of software/ malware, denial of service

Table 2. Levels of *Aggressiveness*. (Original source: [15])

3.1 Who?

When engaged in what Markham calls "world-fixing," one needs to "[derive their methods] through constant, critical reflection on the goals of research and the research questions," understanding not only the problems to be solved, but the potential effects on all parties involved. [27] Before diving into our questions, let us first answer this question of, "who is involved with criminal botnets?"

The *Owners/Operators* of infected computers have responsibility for protecting the information and information systems that they own. When their systems are attacked, they have responsibility for taking actions to regain control of their assets. They are the least capable of detecting and responding to attacks. They have rights of privacy and autonomy of operation within their domain.

When botnets are used to perform secondary attacks, such as defrauding customers of specific banks through "phishing," distributed denial of service, spamming, etc., there are two kinds of *Victims*. Some may be entirely unrelated to the *Owner/Operators*, such as the banks and service providers suffering DDoS mentioned above. Alternatively, there are other *Victims* who are related only in terms of sharing the infected computing resources (e.g., friends, family, customers.) Actions taken to remove bots from the control of attackers benefit all *Victims*, but if the action harms the infected computers, the people sharing those computers are also harmed. *Victims* have little or no responsibility for the systems they use, or for the systems used to cause harm to them. They do not have authority to change those systems or to monitor network activity. They are similar to typical *Owners/Operators* in terms of capability to protect themselves, and have similar rights (e.g., to privacy.)

Service Providers provide network connectivity to the *Owners/Operators* of infected computers or *Victims*. In the case of enterprises these may also be the *Owners/Operators*, while in the case of home users these are Network Service Providers (NSPs), such as broadband, wireless, and DSL companies. *Service Providers* are similar to *Owners/Operators* in terms of responsibility to protect their computer assets. In some cases, they are granted *provider exemptions* from various computer crime or privacy laws that allow certain activities, such as monitoring real-time communications, that would otherwise be an illegal wiretap. They may also have contractual terms extending their authority to the information systems of their customers (e.g., limited control of anti-malware software on customer computers.)

Researchers are the ones capable of reverse engineering today's advanced malware and developing methods to detect, cleanup, and possibly counter-attack the botnet via exploitation of design weaknesses. Their role is to help identify and analyze malicious software, deriving generalizable knowledge that can then be disseminated to corporations for improvement of their products and services, to service providers to improve the efficiency of their response, helping law enforcement understand computer crime tools and techniques, and helping the general public with awareness and training. They have an obligation to act responsibly. They are not exempt from computer crime statutes, and in academic settings may have legal obligations to submit their research protocols to institutional review boards for human subjects protection evaluation. They themselves have no authority to make changes to computer systems owned by others without the knowledge and consent of those owners, but they can work in consultative or advisory roles to those who do.

Since we are talking about criminal activity, two other classes of people involved are *Law Enforcement* and the *Criminals* themselves. *Law Enforcement*, as agents of sovereign governments, are the only other parties who have legitimate responsibilities to bring criminals to justice. They are bound to protect the legal rights of all others, while performing their duties with a minimum of negative impact on innocent or victimized parties. While they do some research, their role is not to provide generalizable knowledge and disseminate their research results to the public as is the case with *Researchers*, on whom they rely heavily for advanced applied and theoretical research. They also rely on *Owners/Operators*, *Service Providers*, and *Victims* to report crimes and provide evidence to further investigations. On the other side are the *Criminals*, who drive much of computer security research today. They act without regard for harm to anyone, may negatively impact the lives of millions, yet (arguably, to some) still have rights.

3.2 What, where, why, when, and how?

We now examine some of the most common and problematic ethical questions surrounding criminal botnet research.

[*Question* 1] **Is it ethical to perform research that alters an active crime scene without coordinating with law enforcement?**

[**YES**] Researchers in academia often value independence of thought, speech, and from involvement with investigative activities of the government. In some ways, there is a societal obligation for academics to be independent and to avoid the appearance of acting as unrestricted agents of law enforcement. It is not their role to collect and deliver evidence to the state, but to study the world and derive generalizable knowledge from their studies to enlighten the public. In certain situations where researchers are studying criminal behavior there are *certificates of confidentiality* limiting compelled disclosure of research data to the state, even under subpoena. Under the U.S. regulation governing the protection of human subjects in research (45 CFR 46, also know as "the Common Rule" [1]) there exists an exemption (§101(b)(2)) for, "Research involving [...] the observation of public behavior: unless: (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation." Certainly identifying a criminal suspect to law enforcement could result in damage to their "financial standing, employability, or reputation," but is that the intent of this federal regulation?

In the case of the Storm worm, it was only possible to trace the origin of the network because researchers reverse engineered Storm's communication protocols and were actively collecting information about active nodes. Those behind Storm have not been identified, but commands originating from networks believed to be associated with criminal activity were clearly observed. It would be unlikely for law enforcement, unaided, to do the research necessary to learn these facts.

[**NO**] Let us assume that the researcher's goal is to maximize benefit to the public by learning about how criminal tools work and developing new detection mechanisms, better investigative capabilities, or new methods of protecting systems. How does protecting the privacy of criminals, or withholding research results for several months to fit conference publication cycles, impact law enforcement? Certificates of confidentiality were designed to (a) protect criminals who are *consenting research subjects*, and (b) are involved in biomedical research under the authority of the Department of Health and Human Services [30]. Even if such a certificate could be obtained, it may be difficult to argue that protection of the privacy rights of criminals results in a greater moral good than providing information to law enforcement officers protecting the public.

Beyond potentially identifying criminals to law enforcement officers, there are other potential impacts of certain research activities on active criminal investigations and thus a need for *deconfliction*. How do actions that fall into the *Aggressive* level in Table 2, which may introduce false evidence into an active crime scene as might result from a *Sybil* attack [16] on a P2P network, impact law enforcement? What may seem a joke – that the Storm botnet would "shrink to a handful of real bots [while] an army of rabid researchers [fight] with each other to measure whatever was left [17]" – has serious implications. What if

researcher's actions divert law enforcement, causing them to issue one or more subpoenas before eventually learning they had "caught" a white hat instead of a black hat? Could this in some way even assist criminals? Should researchers even be allowed to perform such experiments without coordination, or some prior arrangement to work with security operators (who have legal exemptions and responsibility to protect information and information systems)? Is there a need for regulation limiting research to only non-criminal activities, or researching criminal botnet activity only under tightly controlled conditions similar to research into biological agents and toxins like anthrax, ricin, and smallpox (e.g., Public Law 107-188 in the United States)? Or is a government-mandated ethical review model like the Embryonic Stem Cell Research Oversight (ESCRO) [6] committees in the United States, which are separate from IRBs, necessary?

Cybercriminals today possess advanced technical expertise that demands constant study in order to keep up. This is hard enough for researchers, but it is impractical to expect law enforcement to be experts at *both* researching malware and performing complex investigations at the same time. For this reason, law enforcement relies heavily on private sector research. Cooperative efforts between law enforcement and the private sector are vastly improving the situation, but when does this close relationship risk the independence of researchers?

Finally, there arise questions related to *responsible conduct of research*. Should researchers be initiating experiments that alter cyber crime scenes without at least knowing how and when to contact law enforcement, reporting this activity before (or as soon as possible after) performing the experiment? What if the experiment uncovers evidence of very serious financial crime, industrial espionage, or possible national security espionage (e.g., as in the Ghostnet [9] investigation)? Shouldn't actions with potential risks to the researchers (or their institutions) require considering these issues in advance to minimize potential of loss of control of the experiment, or possibly being reported as suspects in criminal activity themselves? Again, the Storm worm is an example of of how actions by researchers alter an active crime scene. In 2008, researchers from different institutions all over the world were actively participating in the P2P network [17]. Research activity made up a large amount of the network traffic and complicated making a distinction between research machines, infected computers, and possible sources of actual malicious C&C traffic.

[*Question* 2] **Is it ethical to restrict researchers to only performing actions that are guaranteed to be risk-free, or avoid any potential ambiguity in laws?**

[**YES**] Since researching botnets necessitates interaction with the bots, which may alter data inside the botnet, there is always the chance of unpredictable side-effects. This is especially true because it is not always possible to know how communications with the botnet will influence the bots when analysis starts. Certain actions might affect or even break the systems of innocents. Simply introducing another zombie into the malicious botnet may result in that host becoming part of a DDoS attack, sending spam, or allowing a criminal to hide behind a proxied connection.

Running malware in sandboxes to observe behavior is now the standard investigation method for new malware. The risk of further spreading while doing so is high. Many researchers ran Conficker samples in order to investigate the exploitation mechanism. Limiting researchers to performing less dangerous analysis steps first (e.g., black-box analysis in closed lab environments, or using only static analysis techniques) reduces risk while often gathering the same information as *live infection.*

[**NO**] Limiting research to only that which is completely risk-free would result in "no research" at all. By not allowing any research the bot developers and criminals using the botnets would have a massive advantage. When no new analysis and mitigation tactics can be investigated, the botnet problem would grow. But where should we draw the line? Conficker, for example, does not appear to include any logic that does direct harm to infected computers (e.g., data destruction), nor has any such harm been observed. Without allowing interaction with the C&C server, it is impossible to know the malware's behavior *in the wild* and the threat it poses.

Attackers have learned to avoid simple means of detection, e.g., automated use of *sandbox* analysis of malware. Researchers must sometimes run malicious code for a long time in order to become a "trusted" node in the infrastructure and to see the heart of botnets. These nodes have to act "undercover" and must behave like regular infected machines. This includes sending spam and participating in DDoS attacks, which inflicts some amount of harm on third parties [21]

[*Question* 3] **Is it ethical to clean up infected computers owned by others without their knowledge and consent?**

[**YES**] Worm infected hosts can crash. They can disrupt networks and harm other hosts. Leaving them infected prolongs this harm. Worm infected hosts have been seen to disable medical facilities, prompting them to seek emergency active countermeasures against Conficker to restore network stability immediately. But how can one calculate the risk vs. benefit for *uncoordinated* cleanup?

[**NO**] It is hardly feasible to clean up only specific computers from remote locations because it is often not known whether commands are proxied to another machine or consumed right by the communication peer. Thus, a remote cleanup must be regarded similar to the actions of self-spreading worms that do not know the next victim machine in advance. There are no examples of *white worms* that were 100% effective and harmless at automatically cleaning up malicious worm infected hosts, but there are many examples of ones that caused more harm than good. The very first attempt at a helpful worm in 1978 left the entire Xerox Palo Alto Research Center (PARC) network useless for a couple of days while each computer had to be manually cleansed of a rampant worm. The *Code Green* worm and Linux *Cheese* worm in 2001, and the *Welchia (a.k.a., Nachi)* worm in 2003, all had problems that caused some systems to crash. There far were fewer systems connected to the internet in 2001-2003 than there are today, and vastly fewer systems involved in critical processes like patient care, emergency call routing, process control, etc.

Conficker arguably has infected several million computers (and is still spreading, one year after its first appearance.) At least two hospitals and one municipal government have reported hundreds of Conficker infected hosts involved in patient care and law enforcement activities. If someone were to release a *white worm* to clean up all Conficker infected hosts, without anyone knowing this was going to occur, there is no guarantee that patient care would not be disrupted or that a serious criminal might be let go on a minor traffic offense because a background check was not possible.

[*Question* 4] **Is it ethical to provide information or tools to third parties for them to use as they see fit, even if they chose to use this information in unethical ways?**

[**YES**] Let us assume that researchers are operating under clear ethical guidelines for responsible research. They expend a significant amount of time and energy doing complex reverse engineering, analysis, and programming, to develop a countermeasure against a criminal botnet. They do not use the tools they develop in ways that could impact third-party systems, by altering processes of file systems contents, but significant details and tools they have developed are released to the general public. Publication of research is necessary to disseminate knowledge and improve technology. The assumption is that those who consume this published knowledge will use it only to produce benefits for society. Anyone with sufficient skills could do the same analysis, in which case use of this knowledge for unethical purposes is solely their responsibility.

[**NO**] It is significantly less effort for someone to take the researchers' output. Does this mean that the researchers have some responsibility if someone who otherwise would not be able to do harm uses the researchers' output, acts without the same degree of ethical consideration, and causes harm? Perhaps, but perhaps not. The issue here is not full disclosure vs. no disclosure. It is one of *responsible disclosure* by evaluating the comparative benefits and harms resulting from disclosure. Disclosing a buffer overflow vulnerability in a commercial product allows consumers of that product to protect themselves, but disclosure of a vulnerability in *malicious software* that they are not able to detect, let alone patch, cannot increase benefit to owners and operators only capable of acting at ARC Level 0. It does provide other criminals with knowledge that could allow them to do further harm, which decreases the benefit to society. How should a middle course of partial, responsible disclosure, be used to minimize benefit to criminals and maximize benefit to other parties involved (owners/operators, service providers, related victims, law enforcement, etc.)

[*Question* 5] **Is it ethical to violate the ownership (privacy) rights of others in order to obtain information that helps mitigate a criminal botnet?**

[**YES**] Different studies, like [34], have been used to enumerate the groups of users affected by botnets. Holz, et al [19], have taken another step further and even investigated the private, mostly financial, information found in various

drop zones. The data helps to understand the collection process of bots and can be useful to derive new preventive methods.

The Storm worm inflicted harm on users who, by themselves, would not have been able to handle it. The ability to perform remote disinfection would have been helpful for those with infected systems, removing the threat for all internet users.

[**NO**] These researchers looked at and used very private information, like credit-card numbers, banking data, and credentials for all kinds of web-sites, without the owners' knowledge. The point at which the benefit for potential future victims outweighs the violation of the privacy of victims in the present is hard to estimate, but it is not a binary function. When, if at all, is it ethical to violate the privacy rights of others in order to mitigate botnets?

4 Conclusion

We have seen how complicated it can be to develop effective countermeasures to today's advanced botnets. It is hard to calculate losses, estimate risks/benefits and achieve an acceptable balance. When an attack raises to the level of national impact – a reasonably predictable event, given past examples of financially and politically motivated attacks – policy makers will face decisions about taking control of computers owned by private citizens or corporations to limit further harm. We hope this work will help inform their discussion of options, how they weigh the potential benefits or harms and choose a series of actions to build into contingency plans. We also hope our peers will contribute additional questions, suggestions and their own opinions about where they believe the lines to be. Consensus is an important requirement for achieving ethical guidelines that are acceptable to the community and can be enforced as much through peers as through some official body.

The authors wish to thank the anonymous reviewers, and Aaron Burstein, for their valuable comments.

References

1. 45 CFR 46. http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm.
2. Author unknown. Estonia urges firm EU, NATO response to new form of warfare: cyber-attacks. http://www.smh.com.au/news/Technology/Estonia-urges-firm-EU-NATO-response-to-new-form-of-warfarecyberattacks/2007/05/16/1178995207414.html, May 2007.
3. Author unknown. On-demand detection of malicious software. Technical Report No. 23, Anti-Virus Comparative, August 2009.
4. M. Bailey, J. Oberheide, J. Andersen, Z. M. Mao, F. Jahanian, and J. Nazario. Automated classification and analysis of internet malware. In *Proceedings of the 10th International Symposium on Recent Advances in Intrusion Detection (RAID'07)*, September 2007.
5. K. Chiang and L. Lloyd. A case study of the rustock rootkit and spam bot. In *HotBots'07: Proceedings of the First USENIX Workshop on Hot Topics in Understanding Botnets*, 2007.
6. N. R. C. Committee on Guidelines for Human Embryonic Stem Cell Research. *Guidelines for Human Embryonic Stem Cell Research*. The National Academies Press, 2005.
7. P. J. Criscuolo. Distributed denial of service. Technical report, Department of Energy, Computer Incident Advisory Capability (CIAC), February 2000.
8. D. Danchev. Legal concerns stop researchers from disrupting the storm worm botnet, Jan. 2009. http://blogs.zdnet.com/security/?p=2397.
9. R. Deibert, A. Manchanda, R. Rohozinski, N. Villeneuve, and G. Walton. Tracking GhostNet: Investigating a cyber espionage network, March 2009. http://www.scribd.com/doc/13731776/Tracking-GhostNet-Investigating-a-Cyber-Espionage-Network.
10. D. E. Denning. The ethics of cyber conflict, June 2008. Chapter 17 in The Handbook of Information and Computer Ethics.
11. D. Dittrich. Second Agora workshop on Active Defense. http://staff.washington.edu/dittrich/arc/AD-workshop-091203.pdf, September 2003. Sponsored by Cisco Systems, Inc.
12. D. Dittrich, M. Bailey, and S. Dietrich. Have we Crossed the Line? The Growing Ethical Debate in Modern Computer Security Research. *Poster to be presented at the 16th ACM Conference on Computer and Communication Security*, November 2009.
13. D. Dittrich, M. Bailey, and S. Dietrich. Towards Community Standards for Ethical Behavior in Computer Security Research. Technical Report CS 2009-01, Stevens Institute of Technology, April 2009.
14. D. Dittrich and S. Dietrich. P2P as botnet command and control: a deeper insight. In *Proceedings of the 3rd International Conference On Malicious and Unwanted Software (Malware 2008)*, pages 46–63, Oct. 2008.
15. D. Dittrich and K. E. Himma. Active Response to Computer Intrusions. Chapter 182 in Vol. III, Handbook of Information Security, 2005. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=790585.
16. J. R. Douceur. The sybil attack. In *IPTPS '01: Revised Papers from the First International Workshop on Peer-to-Peer Systems*, pages 251–260, London, UK, 2002. Springer-Verlag.
17. B. Enright, G. Voelker, S. Savage, C. Kanich, and K. Levchenko. Storm: When researchers collide. In *USENIX ;login: vol. 33, no. 4*, August 2008.

18. T. Holz, M. Engelberth, and F. Freiling. Learning more about the underground economy : A case-study of keyloggers and dropzones. Technical Report TR-2008-006, Department for Mathematics and Computer Science, University of Mannheim, December 2008.
19. T. Holz, M. Engelberth, and F. Freiling. Learning more about the underground economy: A case-study of keyloggers and dropzones. In *Reihe Informatik*, 2008.
20. T. Holz, M. Steiner, F. Dahl, E. W. Biersack, and F. Freiling. Measurements and mitigation of peer-to-peer-based botnets: a case study on storm worm. In *LEET'08: First USENIX Workshop on Large-Scale Exploits and Emergent Threats*, Apr. 2008.
21. J. P. John, A. Moshchuk, S. D. Gribble, and A. Krishnamurthy. Studying Spamming Botnets Using Botlab. In *Proceedings of the 6th USENIX Symposium on Networked Systems Design and Implementation (NSDI '09)*, April 2009.
22. C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxson, and S. Savage. Spamalytics: an empirical analysis of spam marketing conversion. In *CCS '08: Proceedings of the 15th ACM conference on Computer and communications security*, pages 3–14, 2008.
23. S. Kelly. BBC team exposes cyber crime risk, Mar. 2009. `http://news.bbc.co.uk/2/hi/programmes/click_online/7932816.stm`.
24. F. Leder and T. Werner. Know Your Enemy: Containing Conficker. `https://www.honeynet.org/papers/conficker/`, April 2009.
25. F. Leder, T. Werner, and P. Martini. Proactive Botnet Countermeasures – An Offensive Approach. In *Cooperative Cyber Defence Centre of Excellence Tallinn, Estonia*, March 2009.
26. J. Leyden. DDoS protection racket targets online bookies. `http://www.theregister.co.uk/2001/11/26/ddos_protection_racket_targets_online/`, November 2001.
27. A. Markham. Method as ethic, ethic as method. *Journal of Information Ethics*, 15(2):37–55, 2006.
28. R. Naraine. Kraken botnet infiltration triggers ethics debate, May 2008. `http://www.eweek.com/c/a/Security/Kraken-Botnet-Infiltration-Triggers-Ethics-Debate/`.
29. D. of Justice. Criminal Complaint: United States of America v. Paul G. Ashley, Jonathan David Hall, Joshua James Schichtel, Richard Roby and Lee Graham Walker, 2004. `http://www.reverse.net/operationcyberslam.pdf`.
30. Office for Human Research Protections (OHRP). Guidance on Certificates of Confidentiality. `http://www.hhs.gov/ohrp/humansubjects/guidance/certconf.htm`, February 2003.
31. L. Oudot. Fighting Internet Worms With Honeypots. `http://www.securityfocus.com/infocus/1740`, October 2003.
32. H. Phong. Korean agency accuses BKIS of violating local and int'l law. `http://english.vietnamnet.vn/reports/2009/07/859068/`, July 2007.
33. P. Porras, H. Saidi, and V. Yegneswaran. Conficker C P2P Protocol and Implementation, September 2009.
34. M. A. Rajab, J. Zarfoss, F. Monrose, and A. Terzis. My Botnet Is Bigger Than Yours (Maybe, Better Than Yours): Why Size Estimates Remain Challenging. April 2007.
35. N. Shachtman. Porn purveyors getting squeezed. `http://www.wired.com/news/print/0,1294,59574,00.html`, July 2003.

36. E. H. Spafford. Are computer hacker break-ins ethical. In *Deborah G Johnson and Helen Nissenbaum, editors, Computers, Ethics & Social Values*, pages 125–135. Oxford University Press, 1992.
37. S. Staniford, V. Paxson, and N. Weaver. How to 0wn the internet in your spare time. In *Proceedings of the 11th USENIX Security Symposium*, pages 149–170, August 2002.
38. B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna. Your Botnet is My Botnet: Analysis of a Botnet Takeover. Technical report, University of California, May 2009.
39. Symantec. The Downadup Codex: A comprehensive guide to the threat's mechanics Edition 2.0. `http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/the_downadup_codex_ed2.pdf`, June 2009.