Towards Community Standards for Ethical Behavior in Computer Security Research

David Dittrich
University of Washington
dittrich@u.washington.edu

Michael Bailey
University of Michigan
mibailey@eecs.umich.edu

Sven Dietrich
Stevens Institute of Technology
spock@cs.stevens.edu

Stevens CS Technical Report 2009-1, 20 April 2009

ABSTRACT

Since the first distributed attack networks were seen in 1999, computer misuse enabled by *botnets*, *worms*, and other vectors has steadily grown. This rapid growth has given rise to a variety of ethical challenges for researchers seeking to combat these threats. For example, if someone has the ability to take control of a botnet, can they just clean up all the infected hosts? Can we deceive users, if our goal is to better understand how they are deceived by attackers? Can we demonstrate the need for better methods, by breaking something that people rely on today? When one considers the implications of something like botnet cleanup – the blind modification and possible rebooting of thousands of computers without their owners' knowledge or consent – this complexity becomes all the more obvious. To be effective, we must find ways to balance societal needs and the ethical issues surrounding our efforts, lest we drift to the extremes—becoming the very thing we deplore, or ceding the Internet to the miscreants because we fear to act. In this paper, we endeavor to create a dialogue on the ethical issues in computer security and the ethical standards that we intend to enforce as a community.

1. INTRODUCTION

Modern threats such as Denial of Service Attacks, Worms, Viruses, Phishing, and Botnets underscore the need for security research in an increasingly networked and computationally reliant society. Unfortunately, as our understanding of these phenomenon have grown, so has the uncertainty in the computer science security research community on the appropriate ways in which to observe and address these problems.

For example, consider the area of botnet research, which centers around the detection and mitigation of large numbers of infected hosts, or *bots*, networked into a single distributed system, or *botnet* [51]. We have recently seen a steady increase in the amount of criminal activity using botnets, and along with this has come an increase in the number of academic research and federal funding to counter the botnet threat. This criminal activity is compounded by the emergence of politically motivated attacks, such as those against elements of the cyber-infrastructure of Estonia. Responses to these threats are varied, from passive measurement and observation, to calls for the legal right to defend computer systems from attack using aggressive countermeasures [47, 59].

Unfortunately, the structured public discussion of an ethical framework to guide decision making about actions taken while researching and countering botnet attacks, and indeed in a broader set of computer science security research, has not kept pace. Existing structures for determining the ethical behavior (e.g., Institutional Review Boards (IRB), Professional Codes of Conduct) fail to provide detailed actionable guidance due to the absence of technical expertise in this specific domain and a lack community shared values [4]. There is growing frustration expressed by researchers, program committees, and professional organizations about the limits of ethical research and who has responsibility to enforce them [4, 25].

Our primary goal in this work is *to encourage a continuing dialogue on the ethics of computer security research*. Through this dialogue, we hope to build both an expertise that can be used in various policy enforcement bodies (e.g., program committees, IRBs) and will help us arrive at a form of community consensus. To help achieve this goals, this paper provides:

- **An Exploration of Existing Ethical Arguments**. We are certainly not the first authors to grapple with the notion of ethics in general nor ethics in an computer society. Existing work in this field can help us narrow the scope of our efforts and provide guidance on building *consistent* and *coherent* arguments for ethical principles.
- **An Example Framework for Security Research**. We create an amalgam of existing approaches to human subjects research, professional principles, and active response justification in order to create a quantitative framework for judging risk and benefits in computer science security research.
- **Exploration of Ethics through Case Studies**. While our framework explicitly does not draw conclusions about when a piece of research is ethical or unethical, it highlight the relevant ethical issues the research raises. We review 25 recent case studies and apply the framework to a significant fraction of these studies.

This paper is organized as follows. In Section 2 we discuss the general subject of ethics, its relationship to law, and to the communities who are implicated in this discussion. Section 3 describes a potential framework for making ethical decisions and designing computer security related research protocols. In Section 4, we look at a number of case studies inside and outside of academic environments and show how the proposed framework can be used to illuminate the ethical issues implicated in these cases. Finally, we conclude with Section 5.

2. ETHICS, LAW, COMMUNITY, AND STANDARDS

The study of ethics has a long history. While computing and the Internet provide recent twists to long debated ethical issues, the study of even these new applications is a field unto itself. In this section, we provide a context for the remainder of our work by examining the fields of ethics, law, various definitions of community, and existing standards of behavior.

2.1 What is ethics?

Ethics is often defined as a set of morals or guiding principles intended to govern the conduct of individuals and groups within a population (i.e., a profession, a religion, or society at large.) The definition of *computer ethics* has various interruptions in line with this broader definition, and several are explored in Bynum and Rogerson [12]. One of the most often cited of these is from Moor [46]:

> A typical problem in computer ethics arises because there is a policy vacuum about how computer technology should be used. Computers provide us with new capabilities and these in turn give us new choices for action. Often, either no policies for conduct in these situations exist or existing policies seem inadequate. A central task of computer ethics is to determine what we should do in such cases, i.e., to formulate policies to guide our actions. Of course, some ethical situations confront us as individuals and some as a society.

Note that we agree with others in the field who argue that these polices, once developed, are neither absolute laws, nor complete frameworks, nor checklists to be followed blindly [12]. They are never likely to be complete nor the polices mutually exclusive. As such our approach here is close to that of Johnson and Miller [39] in that we are concerned with building expertise in practical decision making. Theoretic ethics and ethical systems are useful in these processes, but not ends in themselves.

2.2 Law versus ethics

The law is in some ways a set of norms that are written to guide behavior within a society. These legal norms can codify another set of moral and ethical norms that are generally agreed upon by that society. These sets of norms are not, however, the same. For example, we may agree that lying to a friend is unethical, but lying to a friend is not always illegal. Lying under oath, on the other hand, is always illegal. In relation to security research, in particular botnet research, there may be many laws in many countries that are implicated by a given action taken by a researcher. But what does this have to do with ethics? Ethical considerations matter to security research in several ways.

- Adherence to ethical principles may be required to meet legal requirements surrounding academic research.
- They may inform academics, security professionals and amateur security researchers as to how to decide on actions to take in response to a criminal botnet.
- They can illuminate the line between beneficial acts and harmful ones.
- They can describe all parties involved, their rights and responsibilities, and how to resolve conflicts between competing interests.

What may be most important in terms of reputation is being able to clearly justifying one's actions should those actions come into conflict with the law, or generate public controversy.

Developing an workable ethical framework is only a first step, however. Having guidelines that embody a set of norms accepted within the security field improves the decision making process. It gives the public a sense of security in knowing that individuals are acting in the best interest of society. Once these norms are accepted, they can then be considered and adopted within the legislative process to advance the common law.

This is similar to the field of *computer forensics*, where the issue of the admissibility of scientific evidence in trial is concerned. Based on standards established in a Supreme Court case in 1993, known as *Daubert*, [50] courts will accept testimony involving computer forensic evidence if it meets criteria of (a) relevancy, and (b) reliability. It is the second criteria that matters for this discussion. The court suggested that judges evaluate testimony for *scientific validity* and ensure its proper application to the facts of the case, saying:

> Many considerations will bear on the inquiry, including whether the theory or technique in question can be (and has been) tested, whether it has been subjected to peer review and publication, its known or potential error rate and the existence and maintenance of standards controlling its operation, and whether it has attracted widespread acceptance within a relevant scientific community. The inquiry is a flexible one, and its focus must be solely on principles and methodology, not on the conclusions that they generate.

If the computer security research community develops ethical principals and standards that are acceptable to the profession and integrates those as standard practice, it makes it easier for legislatures and courts to effectively perform their functions. If the broader society also accepts those principals and standards, an even greater benefit results in terms of societal trust in computer security research.

While both influenced by ethical frameworks, and serving as a guide for classes of ethical behavior, a full discussion of the legal issues surrounding computer science security research is beyond the scope of this paper. Interested readers are encouraged to examine an increasing body of work in this field [11, 52, 43].

2.3 What do we mean by community?

In this paper, we use terms like *researcher*, *organization*, *community*, and *society*. These terms apply in two primary contexts: the population taking action, which will use the ethical framework we suggest to guide decision making, and the population implicated in those actions and/or the intended beneficiaries of those actions (e.g., owners of bot infected computers.) The actor populations are, of course, also, directly or indirectly, members of the protected and beneficiary populations. We find three such populations of interest:

- **Computer Users** These are members of the general public who are independent security researchers or computer hobbyists who are interested in computer security. This population has the least regulation or control of their actions especially when harmful acts that are not covered by criminal or civil laws.
- **Professional Community** These are professionals who have roles that involve them in computer security research or incident response as part of their normal job duties. For example, network operators, security operators, forensic analysts, reverse engineers, computer security incident response staff, etc. Control over this population is principally governed by their employers' administrative policies, agreements signed with employers or clients (e.g., non-disclosure agreements) and contract terms. Harmful acts are punished by dismissal, disbarment, or legal actions.
- **Academic Community** These are people who have academic roles in educational institutions, primarily research staff, research faculty, and students (both undergraduate and graduate level) who are studying information security related topics. Control over this population can include both legal restrictions and institutional policies. Harmful acts would be punished by academic sanctions, dismissal, and/or legal actions.

Note that our notion of actors and beneficiaries here are meant to focus this work beyond many of the more general discussions that dominate existing computer ethics work. For example, many texts emphasize the role of general users of technology play in changing how we think about fundamental issues in society such as privacy and intellectual property rights [6]. Some work does focus on the discussion to the role of professionals involved in the application of this technology [12], but often the notion of professional is limited to the roles of software engineers and engineering mangers rather than security researchers. It is our intention in this paper to focus on the most specific of these actor populations: the academic security community.

2.4 What standards exist for guiding ethical behavior in our community?

The security community as noted in previously, already has some standards and legal obligations for adhering to certain requirements related to research protocols. In this next section, we examine some of these standards.

2.4.1 Rules of Engagement

Great attention has been paid to the ethics of *responding* to computer attack in terms of *use of force* alternatives under theories of the *Law of War*, or *Law of Armed Conflict* (LOAC). For the purposes of this discussion, there are parallels to concepts embodied in the LOAC. For example, the LOAC requires *military necessity* as a pre-requisite for the use of force. It requires *distinction*, that is, actions must be directed against lawful combatants and military targets, not against civilians and civilian infrastructure. Lastly, the LOAC requires *proportionality*, that is a use of force less than or equal to the original harm or violation. As a result of international agreements and protocols, such as those defined in the Geneva Conventions, [33] militaries around the world operate under strict *Rules of Engagement* (ROE). These ROE guide decision making on the field of battle to ensure the actions of military personnel do not result in potential war crimes charges.

Yurcik [60, 61] discusses ethics in relation to attack and retaliation using *information warfare* (IW) tactics, and considers whether the lethality of IW operations affects the ethics of employing such operations in defense. This applies to military responses to attacks at the nation-state level, but sets the stage for the equivalent considerations of responses in non-military settings. Yurcik next considers *hack-back*, or aggressive responses to computer attack by attacking back. [37] More complete analyses of the application of international law and the law of war to state-directed IW – also known as *cyberwarfare* – operations were done by Sharp [54] and Wingfield. [58] The actors here are primarily nation-states, not individuals.

Dittrich and Himma [22] discuss the legal and ethical frameworks for responding to computer intrusions. Their research identifies three ethical principles as being central to consideration of aggressive counter-measures: the *Defense Principle*, the *Necessity Principle*, and the *Evidentiary Principle*. Dittrich and Himma build on previous work by Yurcik, specifically focusing on the non-military considerations for response, as well as considering transition of response from civilian to military realms. Himma later expands [30] on previous work with Dittrich to include the *Punishment Principle* and the *Retaliation Principle*.

2.4.2 IEEE/ACM Standards

The Association of Computing Machinery's *Code of Ethics and Professional Conduct* [3] consists of three distinct parts which highlight fundamental ethical considerations, specific professional responsibilities, and leadership imperatives. Section 1 entreats members to: "contribute to society and human well-being" (Section 1.1) and to "avoid harm to others" (Section 1.2), along with six other principles (e.g., don't discriminate, be honest, respect privacy). Professional responsibilities include calls that "ACM members must obey existing local, state, province, national, and international laws unless there is a compelling ethical basis not to do so," (Section 2.3) and to "access computing and communication resources only when authorized to do so." (Section 2.8), along with maintaining competence, accepting review, etc.

IEEE also maintains the "IEEE Code of Ethics" [34], which, although more abbreviated than the ACM version, contains many of the same imperatives. Specifically, the code commits members "to the highest ethical and professional conduct". Members agree to avoid conflicts of interest, be honest, engage in responsible decision making, accept criticism of work, etc. Of particular interest are the mandates to "to improve the understanding of technology, its appropriate application, and potential consequences" and "to avoid injuring others, their property, reputation, or employment by false or malicious action"

These are certainly not the only such codes of conduct for computer professionals. For example, IEEE and ACM have a approved a joint Software Engineering Code of Ethics [27] and there are numerous professional organizations with codes whose headquarters are outside the United States (e.g., Australian Computer Society, Canadian Information Processing Society (CIPS)). In addition some individual companies and academic institutions have their own ethical codes (e.g., Gateway, Texas Instruments, University of Virginia, Howard University), but these are by no means universal.

2.4.3 The Belmont Report and IRBs

Medical or behavioral researchers in the academic community in the United States have a legal requirement to examine certain ethical considerations. Beyond those requirements, there are other ethical considerations that individuals involved in botnet research in specific should consider. We will look at each of these sets in turn.

In response to a number of incidents of medical research being performed on individuals without their knowledge or consent, the National Research Act was passed in 1974. These incidents included syphilis studies involving low-income African-American males in Tuskegee, Alabama in the 1930s, and medical experiments performed on prisoners of war in World War II (protection of whom was mandated in the Nuremberg Code following Nazi war crimes trials.) This act established the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. In 1979, the National Commission prepared a document known as "the Belmont Report." In 1981, the Department of Health and Human Services (DHHS) released a regulation (45 CFR Part 46, Subpart A) based on the Belmont Report, known simply as the "Common Rule."It defines requirements for research involving human subjects that apply to individual researchers, their institutions, and their related Institutional Review Boards (IRBs).

The three basic ethical principles their application described in the Belmont Report are:

- **Respect for Persons** Individuals should be treated as autonomous agents, whose right to decide about their own best interests is to be respected; Individuals with diminished autonomy, who are incapable of deciding for themselves, are entitled to protection.
- **Beneficence** Do not harm; Maximize possible benefits and minimize possible harms; Systematically assess both risk and benefit.
- **Justice** To each person an equal share in treatments and benefit of research according to individual need, effort, societal contribution, and merit; There should be fairness of procedures and outcomes in selection of subjects.

2.5 Limits on existing standards?

Allman [4] discusses the potential role of conference Program Committees (PCs) in guiding researchers in terms of the ethical foundations for their research methodologies. The ACM code is cited as one guide that PCs may apply in judging academic papers submitted to them for review, however Allman mentions that interpretations can be varied and application of the code to specific actions difficult. One could read the ACM code, Section 2.3, and apply the ethical principles cited in this paper and conclude some research is acceptable, while another could read its Sections 1.1 and 2.8 and conclude the same research is unacceptable. Allman also mentions IRBs as potential arbiters of the ethics of botnet research, but points out that IRBs historically deal with human subjects involved in research, historically from the fields of biology, medicine, psychology, etc. He questions whether IRBs have similar value systems, or domain expertise, to judge network security research and this is consistent with some of the concerns cited by Garfinkel [25].

For example, consider botnet research. In the medical research context, the research subjects themselves are the eventual beneficiaries of the research outcomes (and to a larger extent the rest of society in general.) In the security research context, the research subjects are often criminals and their tools, which happen to involve (most often unknown to their owners) the computers of innocent third parties. This means there are two potential sets of beneficiaries who potentially have an inverse benefit/harm relationship to one another. That is to say, publication of some research results may have a small benefit to society in general, while the criminals whose tools are the subject of research may have a much larger benefit. The criminals may learn how to improve their attacks, or make them harder to detect and mitigate. This is especially true of publication of theoretical research that postulates new and more potent types of malicious software, which could serve as a blueprint for criminals. This is a very complex calculus that sometimes involves initial non-public disclosure of research results, and very carefully timed public disclosure, in order to assist law enforcement or provide lead time for security operations elements to act. Non-public disclosure is diametrically in opposition to typical academic research, where "publish or perish" is often the mantra.

As another example, consider the works of Denning and Spafford who discuss ethics in the context of those engaged in *computer intrusions.* Denning [18] describes the opinions of hackers who were interviewed about computer intrusions as to whether those acts were ethical or not. Some hackers believed that certain malicious actions where wrong and unethical (e.g., "breaking into hospital systems," "reading confidential information about individuals," "stealing classified information," "committing fraud for personal profit.") Some hackers believed that exploring computer systems was ethical, provided that "the objective is to learn and avoid causing damage." Spafford [53] looks at similar acts in terms of *right and wrong*, and whether a *greater good to society* is achievable by computer intrusions. In Spafford's analysis, computer intrusions may only be ethically justifiable in the most extreme cases, such as to save a human life in an emergency. In discussing publication of worm or virus code (which may be capable of resulting in harm to innocent third-parties) he states that, "publication should serve a useful purpose; endangering the security of other people's machines or attempting to force them into making changes they are unable to make or afford is not ethical."

Unfortunately, there is no commonly accepted framework within which decisions can (relatively) clearly and consistently be made. Nor is there currently an accepted venue in which to consider them. Should PCs be the arbiters, as Allman suggests, [4] or is that venue too closed? Both Allman and Garfinkel [25, 4] suggest Institutional Review Boards may have a role, but are they currently capable of making judgements about the issues raised in this paper? If the role of IRBs is to ensure compliance with the National Research Act, "to put a stop to researchers saying 'Trust me'," [25] does this support Burstein's suggestion that researchers, "participate in [legislative] reform efforts... to make known how the lack of a research exception affects them"? [11]. Sicker, et al, [52] offer reasons such legislative reform is neither a timely nor especially effective solution and suggests that prosecutorial discretion may preclude the need for legislative reform (although this has its own risks.). Should computer security researchers be required to receive training similar to the *Education on the Protection of Human*

Subjects mandated by NIH, and if so, what should be covered? How are these issues dealt with internationally?

While the limitations on scope, expertise, and lack of consensus are broader than a single discussion or single work, it is clear that the answers to these and other questions will require community dialogue and effort. For such a dialogue to be successful, we will need to draw from a rich set of experiences and build consistent and coherent arguments for the ethical or unethical behaviors contained therein. In the next section, we describe our efforts in building such a dialogue.

3. SYSTEMATIC EVALUATIONS OF CASE STUDIES

Bynum and Rogerson [12] suggest a multi-staged approach to case study analysis in order to build ethical judgement capabilities. These stages include: identifying key ethical principles, detailing the case study, identifying specific ethical issues raised by the case, calling on your own experience and skills for evaluation, then the abilities of others, and finally, applying a systematic analysis technique.

We detail the case studies and provide an analysis of the ethical issues raised in Section 4. However, in order to make best use of those studies, we first identify the key ethical issues for security researchers and extend existing frameworks so they can be used as a systematic analysis technique.

3.1 General Ethical Issues for Security Researchers

When considering actions related to research or mitigation of malicious or illegal activity, there are many issues that must be considered. These involve issues of (a) proportionality, (b) targeting, (c) necessity, (d) desired outcome, (e) potential consequences, and (f) the greater moral good to society that is expected to result (and whether it outweighs any potential harm to innocent third parties.)

For example, there are situations in which great tension exists between releasing information to claim first discovery, or holding it back to prevent harm. This is uncommon to most academic research, where discoveries are primarily applied immediately to benefit society. We must remember that much computer security research is focused on *criminal activity* that is actively causing harm to the public, and the potential for harm from unethical actions could extend to the entire internet population. Take public/private partnerships aimed at responding to cyber threats, which involve government, commercial entities, academic researchers, and select members of the public who specialize in computer crime activity. Here the tension to go public with new knowledge is more intense. Vendors of anti-malware products may wish to be first to disclose to increase their customer base and sell more products. Academics may wish to be first to disclose to enhance their academic positions and increase chances of future funding. Private citizens may wish to improve their chances of getting a new high-paying job. In research into curative treatments in healthcare, premature disclosure of certain information will not be used by viruses, micro-organisms, or cancer, to improve pathological efficiency. In security research, such premature disclosure *can and does* result in improvements of malicious software and tactics that make the task of responders much harder, and the potential harm to the public greater.

The kinds of questions that researchers must ask themselves include (but are not limited to) the following:

- Are the research results intended to protect a specific population, and if so, which population? (E.g., the owners of infected hosts, the victims of secondary attacks using a botnet, the researchers' own institution, or the general internet user?)
- Is there a way to achieve multiple benefits to society simultaneously when studying criminal botnet behavior? (E.g., developing new defenses, while aiding investigation of criminal acts and assisting victimized network sites?)
- Who will benefit more from publication of research findings, and in which order: Victims of criminal acts; authorities responsible for protecting their citizens; the researchers themselves; or the criminals who are perpetrating computer crimes?
- Is there any other way to accomplish the desired research result(s)?
- What is the safest way to disseminate research results without risk of improper use by individuals who may not share the researchers' ethical standards?
- If all security research is halted because 100% safety cannot be guaranteed, is the result a greater *harm* to society because no new defenses are developed, or is taking the risk of some small number of potential infections worth the thousands or millions of hosts protected by resulting new defenses?

While these general questions get to some of the issues, they are not sufficient to give fine-grained guidance in a form that could be evaluated. We are encouraging researchers to include in publications an indication that they have made the effort to evaluate their work against the ethical questions raised in this paper in a way that is uniform across all research situations and topics. Further, using a simple and uniform methodology supports consistent evaluation by outside parties in a manner that improves trust in computer security research protocols.

3.2 Towards a Systematic Approach

Table 1 shows a potential ethics scoring guide that includes the salient ethical principals and their sub-components as identified in the previous Sections. The table is split into two sets. The issues on top of the table come from the Active Response Continuum, and are primarily aimed at situations involving direct interaction with hosts outside one's own administrative control. The issues on the bottom come from the Common Rule, and are primarily aimed at protection of research subjects (and for the purposes of this paper, other indirectly involved third-parties.)

A set of 16 representative case studies where chosen from Section 4 and evaluated as to the ethical issues raised. Filled circles indicate the issue is *central* to the case study, and empty circles indicate the issue is *tangential*, or of lesser importance. Those cases involving active engagement with third party systems in some way (e.g., internal botnet enumeration, disinfecting, monitoring, taking control of botnet command and control, copying files, etc.) obviously tend to involve the most issues, while those that are more narrowly focused on vulnerability disclosure involve fewer.

Following Bynum and Rogerson, we have identified how these issues pertain to the selected cases. We do not take the

Table 1: Potential Ethics Issues. (● = Central Ethical Issue, ○ = Tangential Ethical Issue)

Principle	Question	Case Number																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Defense	Population being protected is identified?	●	●	●	●	●	●	●	●	●	○	○	○	●	○	○	○	○
Defense	Looks like use of *force*?	●	●	○	○	●	●	○	●	○			○	●			○	
Defense	Actions are proportional?	●	●	○	○	●	●	○	●	○			○	●				
Defense	Necessary to repel or prevent harm?	●	●	●	●	●	●	●	●	●	○	●	●	●	○	○	○	○
Defense	Benefits of disclosure favor victims over attackers?	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Defense	Actions are appropriately directed?	●	●	●	●	●	●	●	●	●		○	○	○			○	
Necessity	Greater moral good defined?	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Necessity	No other reasonable options available?	●	●	●	●	●	●	●	●	●	○	○	○	●	○	○	●	○
Necessity	Otherwise respectful of rights?	●	●	●	●	●	●	●	●	●		○	○	●			●	○
Punishment	Avoids punitive motives?	○	○				●		●					○				
Retribution	Avoids retributive motives?	○	○				●		●					○				
Evidenciary	Adequate reason to think preconditions of applying other principles are met?	●	●	●	●	●	●	●	●	●	○	●	●	●	●	●	●	●
Respect for Persons	Individuals treated as autonomous agents?	●	●		●	●	●		●	●		●	●	●			○	
Respect for Persons	Individuals (or their providers) informed and allowed to consent?	●	●	○	●	●	●	○	●	●		●	●	●			●	●
Respect for Persons	Individuals with diminished autonomy are protected?	●	●	○	●	●	●	○	●	●		○	●	●			●	●
Respect for Persons	Identities of innocents are protected?	●		●	●	●	●	●	●	●		●	●	●			●	●
Beneficence	Low potential to inflict harm?	●	●	○	●	●	●	○	●	●	○	●	●	●	○	○	●	●
Beneficence	Maximize possible benefits and minimize possible harms?	●	●	○	●	●	●	○	●	●	○	●	●	●	○	○	●	●
Beneficence	Risks and benefits systematically assessed?	●	●	○	●	●	●	○	●	●	○	●	●	●	○	○	●	●
Justice	Who benefits?	●	●	○	●	●	●	○	●	●	○	●	●	●	○	○	●	○
Justice	Fairness (neutrality) of procedures & outcomes in selection of subjects?	●	●	○	○	●	●	○	●	●		●	●	●			●	○

next step and perform an evaluation, however this could follow in future work. The work we have done to date provides a straw-man proposal for a foundation on which a scoring methodology could be built, and a simple and clear set of issue laid out to guide researchers in developing their research protocols in a unified manner. The result can establish a basis of trust with the general public, who are implicitly involved as central stake holders.

Our aim intent is to find a way for researchers to take risks that are acceptable to the general public and address the advancing threat landscape. Being overly conservative may lose ground to these threats. Being overly aggressive or risky, especially if some harm results, may generate a backlash that likewise loses ground. A reasoned and measured approach, based on accepted ethical standards, can result in a decreased threat landscape. It can also result in something akin to the legal concept of a *reasonable person* standard (i.e., "Would a reasonable person, in the same circustances, chose to act in the same way?")

4. CASE STUDIES AND FRAMEWORK APPLICATION

We will now look at several case studies both inside and outside of the academic research setting, in terms of the ethical principles mentioned earlier. Not all of these are research specific, but all serve to illustrate the ethical questions involved.

4.1 Participating, Observing, or Breaking Something to Understand How Something Works

"Shining Light in Dark Places: Understanding the Tor Network. [45] McCoy et al. participated in the Tor network to analyze the types of traffic, countries using Tor, and possible abuses of the network. By running a modified Tor server, they were able to observe all traffic either being relayed (they were a relay for two weeks) or exiting the network (they were an exit node for another two weeks). Fully aware that the payload collection would be a problem, they tried to limit the amount of payload data being collected in the experiment. The main purpose of the work was one of discovery and measurement, and how to possibly limit the exposure of sensitive data, as they devised a method to detect logging by malicious routers. However, suggestions for improving and fixing Tor also emerged from this paper. (Case 11)

Why and How to Perform Fraud Experiments. [36]. In this work the authors discuss their experiences with conducting fraud experiments (i.e., phishing). In particular, they focus on two studies: one in which they explore the impact on phishing source (i.e., someone trusted versus someone random) [35] and one in which they explore the impact of cousin domains (i.e., those which sound similar to the real domain) [24]. The purpose in this article was not to explore these studies in depth, but rather to highlight three important ethical issues associated with conducting these experiments. The first of these issues, that of informed consent, centers around whether it is ethical to not allow the participants to choose whether to participate in the study. Here the ethical considerations of the value from the study must be weighed against the fact that the study results change if the users know it is happening. A similar set of arguments are used in discussing the next issue, that of explicit fraud. As mentioned in the article, lying to users must be done with the utmost care, be overseen by full IRB boards, and generally be avoided by most researchers. Finally, the authors explore the notion of debriefing, that is, informing users after the study that they participated without their knowledge. This debriefing is generally a requirement of the waived informed consent, especially in the cases where one COULD have received informed consent, but the argument was made that it would impact behavior. In the case of their experiments, the authors successful argued for IRB approval to waive informed consent, conduct fraud experiments, and not to debrief the subjects afterwards based on a series of risk/benefit arguments that were not articulated in the summary article. (Case 12)

BBC TV: Experiments with commercial botnets. [44] In March 2009, the British Broadcasting Company (BBC) *Click* technology program chose to perform an experiment. Unlike the situation in Kraken, direct control of the botnet was exercised. The BBC staff purchased the use of a malicious botnet identified after visiting internet chat rooms. They used that botnet for several purposes: (1) They sent thousands of spam messages to two free email accounts they set up on Gmail and Hotmail; (2) They obtained permission to perform a distributed denial of service attack against a site willing to accept the flood; (3) They left messages on the infected computers that made up the botnet; and finally (4) issued unspecified commands that disabled the bots on those computers, killing the botnet. There was immediate reaction to the news of this experiment by a law firm in the United Kingdom, citing probably violation of the British Computer Misuse Act by the unauthorized access and use of computer resources, and unauthorized modification of the configuration of the involved computers. The BBC's response to the criticism was to state they had no intention of violating laws, and believed their actions were justified by citing, in their words, "*a powerful public interest in demonstrating the ease with which such malware can be obtained and used; how it can be deployed on thousands of infected PCs without the owners even knowing it is there; and its power to send spam e-mail or attack other Web sites undetected.*" [49] (Case 1)

Measurements and mitigation of peer-to-peer-based botnets: a case study on storm worm. [32] In April 2007, Holz, et al, 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 affects 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. (Case 3)

Spamalytics: an empirical analysis of spam marketing conversion. [41] Kanich, et al, (2008) performed a study of the conversion rate of spam campaigns. Their analysis was achieved by infiltrating the Storm botnet and manipulating spam being relayed through systems they controlled by altering command and control (C&C) traffic, and using a fake web site that looked like web sites advertised by those responsible for setting up the illicit Storm botnet. The ethical considerations used to justify their experiments follow the principle of the use of *neutral actions* that *strictly reduce harm*. This was the first time research was performed to learn the conversion rate of spam campaigns. Alternative actions that *could also result from manipulation of C&C traffic*, which may result in an equal or greater moral good to society, are not discussed. (Case 4)

Studying Spamming Botnets Using Botlab. [38] John, et al, (2008) researched spam-generating botnets through analysis of email messages identified by email filters at the University of Washington (UW). Through the use of a botnet monitoring architecture incorporating malware analysis and network behavioral analysis, they were able to develop several functional defenses. e explicit about the risks that result from doing behavioral analysis of malicious botnets, and conclude that, "a motivated adversary can make it impossible to conduct effective botnet research in a safe manner." Observing that an attacker could design even benign looking C&C traffic that could result in the researchers' bots causing harm to third-party systems, they chose to be conservative and halted all network crawling and fingerprinting activity that would identify new malware binaries. They also stopped allowing any outbound connections to hosts other than a small set of known central C&C servers, which meant they halted all analysis of Storm (which uses variable ports for its obfuscated C&C servers.) By taking a very conservative stance, they are minimizing potential harm yet simultaneously limiting their future ability to do beneficial research. (Case 5)

P2P as botnet command and control: a deeper insight. [21] In 2006, Dittrich and Dietrich, began analyzing the Nugache botnet. Nugache, the first botnet to successfully use a heavily encrypted pure-P2P protocol for all command and control, was nearly impossible to observe through passive monitoring of traffic flows from the point-of-view of local networks. After fully reverse engineering the Nugache P2P protocol, a crawler was written that took advantage of weaknesses in the P2P algorithm. Several enumeration experiments were performed with the crawler, carefully crafted to ensure minimal impact on the botnet. This crawler, and the enumeration experiments performed with it, are similar to later efforts to enumerate the Storm botnet. [42, 32]. The authors cite two key issues with botnet enumeration experiments: *accuracy* in counting, and *stealthiness*. They note the potential for researchers doing aggressive enumeration experiments to inflate counts obtained by other researchers, to hinder mitigation efforts, or to impede law enforcement investigations. (Case 7)

4.2 Hack back and Aggressive Response

Tracing Anonymous Packets to Their Approximate Source. [10] Burch and Cheswick show a method that uses controlled flooding of a link using the UDP chargen service to achieve a form of IP traceback to the attacker's source, or close enough to it. At a time when DDoS was on the rise, many methods were being explored to tackle the problem. The researchers even dedicate a small section at the end to the ethics of their approach: they admit that their method could be questionable, perhaps even just as bad as the attack they were trying to trace. However, they argue that their intent was the benefit of the Internet community, whereas the intent of the attacker was to harm the community. (Case 13)

Learning More About the Underground Economy : A Case-Study of Keyloggers and Dropzones. [31] - In order to study impersonation attacks, typically achieved using keyloggers, Holz et al. used a malware analyzer to locate so-called dropzones within malware samples. These dropzones are the places where keylog information gathered from the users is sent by the malware, to be later retrieved by the malware operators. At these dropzones, the researchers discovered 33GB of data from 173,000 compromised computers, containing 10,000 bank account and 149,000 e-mail account passwords. This study was conducted over a period of seven months in 2008 and aimed to study the underground economy and to automate the analysis process. The collected data was eventually handed to AusCERT, which acted as a notification broker for the victims. (Case 9)

Symbiot: Active Defense. In March 2004, the Austin, Texas based company *Symbiot, Inc.* announced a product named the *Intelligent Security Infrastructure Management Systems* (iSIMS) platform posessing counter-strike capabilities. [26] Their product was positioned as a means for victims to not only block detected attacks, but to automatically identify "attackers" and direct retaliatory strikes, or even launch preemptive Denial of Service (DoS) attacks to stop attackers. Critics said the system encouraged vigilantism, and noted that true attribution of attackers was not actually being done, only *last-hop* identification, thus targeting of innocents for the counter-strikes was highly likely. The system was also promoted in terms of allowing retributive and punitive actions. (Case 8)

Lycos Europe: "Make Love not Spam" Campaign. [20] In 2004, Lycos Europe – a service company with roughly 40 million e-mail accounts in eight European countries – decided it was time to do something to counter unsolicited commercial email (also known as *spamming*). Lycos created a screen saver designed to impact sites associated with spam emails by consuming the majority of bandwidth available to those sites. The system, and campaign associated with it, was named *Make Love not Spam* (MLNS). The MLNS campaign began operating in late October 2004, and was ended the first week of December 2004 after the screen saver was installed by over 100,000 users. Their two principle stated goals were punitive and retributive: (1) to annoy spammers and to thereby convince them to stop spamming by (2) increasing their costs and thus decreasing their profits. Lycos did not show they had no other options, such as law suits, by which to achieve the same goals. Lycos could not guarantee specific targeting of only culpable parties, nor did they correlate *illegal* spamming with targeting. Some targets could have been innocent of any criminal acts. The final analysis, based on the principles expressed by Himma, showed Lycos had failed to meet the preconditions of the Defense Principle, the Necessity Principle, or the Evidentiary Principle. (Case 6)

Tipping Point: Kraken botnet takeover. [48] In May 2008, researchers at TippingPoint Technologies' Digital Vaccine Laboratories reverse engineered the encryption used by the Kraken bot, and were able to infiltrate and take control of the 400,000 host botnet. This is the same activity performed by some academic research groups, and results in the same situation: the potential to fully control a malicious botnet. One of the researchers interviewed, Cody Pierce, suggests they were, "one click away from [shutting] down the communication between the people sending commands to these [infected] computers." While they may have had no intention of taking action, the discussion surrounding the situation is applicable here. A statement by Endler (tipping point) is interesting to consider: *If you see someone breaking a window to go into someone's house, that really doesn't give you the right to break another window and go in after them.* [48] Implicitly, Endler is talking about violating a third-party's property rights by breaking in to take action (either punitive or retributive) against a criminal. This would not be justifiable, according to Himma, under any of the ethical principles he cites. There is at least one state court decision, however, that aligns with the *Necessity Principle* [1] in suggesting that an emergency private search may be allowable. The reasoning involves allowing a private citizen to break and enter into another's property to *retrieve and protect the stolen goods* of a victim of theft if they are easily destructible or concealable.

University of Bonn: Stormfucker. On December 29, 2008, a research group from the University of Bonn presented a talk at the 25th Chaos Communication Conference (25C3) in Germany on "Owning the Storm botnet." This research was inspired by the Storm enumeration research at the University of Mannheim. [32] 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. [21, 41, 32, 14, 7] 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, [16] 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. This work was not presented in an academic setting. Had it been, a discussion of the ethical principles that could justify attempting to clean up thousands of infected computers, such as with Denning [19] or Spafford [53], would help guide those with access to the source code in deciding how to use it. (Case 2)

Information Warfare Monitor: GhostNet. [17] Between June 2008 and March 2009, researchers in Canada conducted a multi-phase investigation of a malicious botnet. The victims included the foreign embassies of dozens of countries, the Tibetan government-in-exile, development banks, media organizations, student organizations, and multi-national consulting firms. Initial research involving passive monitoring of suspected victim networks confirmed the intrusions and identified the malware, which was then reverse engineered. *Honeypots* were then infected and used to collect intelligence on the botnet's operation and control servers. The researchers "scouted these servers, revealing a wide-ranging network of compromised computers." Gaining access to the attackers' command and control front end, they were able to, "derive an extensive list of infected systems, and to also monitor the systems operator(s) as the operator(s) specifically instructed target computers." [17] This activity falls within the lower- to mid-level of aggressiveness in the Active Response Continuum, [22] and most certainly involves unauthorized access to systems outside of the authority of the researchers. While there is not mention of ethical considerations, the researchers' actions appear to conform with the ethical issues of proportionality, defense, necessity, and are narrowly targeted at attacker-controlled systems. It is assumed from the structure of the report that it was delivered to law enforcement agencies directly or indirectly through the victims being assisted.

4.3 Vulnerability and Disclosure

Exploiting open functionality in SMS-capable cellular networks. [23] - Enck et al. suggest a bandwidth-exhausting attack on cellular networks by sending enough text messages (SMSs) to prevent establishment of voice channels for legitimate callers. Since text messages and voice-setup messages use the same medium, this attack is possible, which is what the authors clearly demonstrate in their paper. According to the authors, a sufficiently dedicated attacker can disrupt voice traffic for large cities such as New York, and a truly dedicated attacker can target a large continent with the help of a DDoS network. They provide the required message rate for a successful attack on cities like New York or Washington, DC. They offer some thoughts on how to solve or mitigate this problem, but tthe solution does not appear to be complete without a complete rearchitecture of the cellular network. They suggest that this problem should be investigated further to protect this critical infrastructure.

Pacemakers and Implantable Cardiac Defibrillators: Software Radio Attacks and Zero-Power Defenses. [28] Implantable cardiovascular defibrillators (ICD) are implanted medical devices used to sense a rapid heartbeat and administer a shock to restore a normal heart rhythm. These devices are configurable through a device programmer which connect to the ICD wirelessly. This paper demonstrates several attacks on the privacy and and integrity of one such medical device using a software programmable radio. The proof of concept attacks described in the paper determined if the patient had such a device, its type, personal information about the patient and reception of real time telemetry data. More importantly, the attacks showed the ability to change or disable therapies (what the device does in certain conditions) and the ability to deliver commands to shock the individuals heart. The potential risks of such a disclosure have immediate and life threatening impacts. As such, they are fairly anomalous when compared with the risks associated with most security research. The authors go to great lengths to avoid discussing of attacks from distances ($\gtrsim$ 1 CM), attack or protocol specifics, or descriptions of how their attacks could impact the health of an individual. The authors intentionally explore the rational for their disclosure, in spite of the risk, describing the benefits in terms of increased privacy and integrity for future such devices. (Case 16)

Black Ops 2008 – Its The End Of The Cache As We Know It. [40]. In the summer of 2008, Dan Kaminsky (IOActive, Inc.) found a practical attack on an old bug involving a weak random number generation algorithm used for creating transaction IDs. These transaction IDs were meant to ensure clients were talking to the real DNS server. The bug existed in dozens of popular DNS implementations serving between 40% to 70% of internet users.

Attackers exploiting this bug could poison DNS cache entries and control where victims' computers connected. As DNS is critical to operation of all services on the internet, and plays a key role in a wide variety of trust chains, significant damage could result from widespread exploitation of this bug. Balancing the huge risk, the author intentionally set about the process of notification and correction *before* publication/presentation at Blackhat, including the controversial step of requesting that other researchers *not speculate* on the bug or develop attacks of their own. As a result of patient and coordinated disclosure and mitigation efforts, hundreds of millions of users were protected prior to the vulnerability being announced.

RFID Hacking. [9] Bono et al. revealed a hardware mechanism, built from publicly accessible resources, for breaking RFID devices used in the SpeedPass, a payment token for purchasing gasoline and other items at a US gas station, and also in RFID-enabled car ignition keys. Their approach included reverse engineering the device, showing that it was possible to crack the 40-bit key in roughly an hour, and creating a cloned device with which they purchased gasoline, and also starting a car with a cloned device.

Heydt-Benjamin et al. [29] built a device to capture and clone first-generation RFID-enabled credit cards. This earned them a related episode in the then popular US television show '24'. As they show, the credit card owner's data can be captured at a distance, e.g. by pointing a reader at the person or their purse to access the RFID chip. To demonstrate their work, they successfully completed a purchase with their cloned device using a commercial credit card reader.

How to Own the Internet in Your Spare Time. [55] Staniford et al. start by analyzing Code Red, comparing it to Nimda, and speculate about future worms by exploring various propagation vectors. They create conceptual worms, such as an improved Code Red (aptly named Code Red II), flash worms, hit-list scanning worms, the Warhol worm, and the topological worm, and muse about their propagation speeds and control vectors. They also explore the concept of a stealthy contagion of users via file-sharing networks. In summary, they provide several recipes for creating massive disruptions within a short period of time. (Case 10)

Botnet design. In "Army of botnets" [56] and "An advanced hybrid peer-to-peer botnet" [57], the authors devise botnets based on smaller disjoint botnets that collude to form a much larger botnet, or advanced command and control mechanisms for P2P botnets. In either case, the level of description for the mechanisms is very high, from pseudo-code to the key exchanges necessary to create and maintain such advanced botnets. (Case 15)

WORM vs. WORM: preliminary study of an active counter-attack mechanism. [13] Castaneda et al. propose the concept of anti-worms, an automated process that generates a variant of the worm in question. They created a Windows-based prototype and tested it in a smaller run, and simulated its effects at a larger scale. Some of the proposed mechanisms include a patching worm, one that would either remove an existing worm infection or prevent it altogether The authors do realize that there are some legal issues (accessing a remote computer without the consent of the user) and network implications (disruptions by spreading just as fast as the original worm) for their approaches and present a short discussion to that effect. When this paper was published, concepts like Code Green and CRClean, anti-worms for Code Red, had already been publicly discussed. (Case 14)

A pact with the devil. [8] Bond and Danezis create the Satan Virus, aka The Devil Worm, a hypothetical ultimate worm that plays the participants against each other. The propagation of the malware is drawn by temptation of access to another user's machine, mails, and documents in general. It further tempts the infected user to recruit more targets for it, since it watches the infected user for remote access to the machines it originally gave the infected user access to. By threatening to disclose this unauthorized access, the malware then blackmails the user to continue gathering new users for its network, and then eventually double-crosses the user and "sells" his or her information as well. While the malware is hypothetical, the authors do describe implementation issues and sample temptations and threats that the malware can use.

4.4 Publication of Results and Data

Protected Repository for the Defense of Infrastructure Against Cyber Threats (PREDICT). [2]. The availability of realistic network data plays a significant role in fostering collaboration and ensuring U.S. technical leadership in network security research. Unfortunately, a host of technical, legal, policy, and privacy issues limit the ability of operators to produce datasets for information security testing. The Virtual Center for Network and Security Data is a unique effort to organize, structure, and combine the efforts of the network security research community with the efforts of the Internet data measurement and collection community. Under the umbrella of the Protected Repository for the Defense of Infrastructure against Cyber Threats (PREDICT) initiative of the Department of Homeland Security Science and Technology directorate, the Virtual Center will provide a common framework for managing datasets from various Internet data providers. It also will formalize a process for qualified researchers to gain access to these datasets, in order to prototype, test, and improve their Internet threat mitigation techniques, while ensuring that the privacy and confidentiality of Internet users are not compromised. Does publication of network data effect the privacy of Individuals? Can the government sponsor this research? Are current privacy protection methods (anonymization) sufficient? Is it legal for providers to collect this data? (Case 17)

Playing Devil's Advocate: Inferring Sensitive Information from Anonymized Network Traces. , "Issues and etiquette concerning use of shared measurement data" [5] - unanonymized a shared network trace and the dataset publisher response Coull et al. [15] divulged deanonymization techniques for recovering both topology and heavy-hitter (e.g. major web servers) information from anonymized datasets. While such datasets are necessary for scientific validation of research results, researchers rely on strong anonymization techniques to protect sensitive and proprietary information about their internal networks. In this case, the authors applied their technique to three datasets, two from their own respective institutions, as well as one well-known and publicly accessible dataset prominently used in the security community. To prove the correctness of their result, they published key information about the public dataset in their paper, thus revealing internals about that researcher's network.

5. FUTURE WORK

We have reviewed many security research situations where those involved faced questions about what they should and should not do with knowledge they possess. In some cases, actions taken were questioned by observers. In other cases, actions were not taken and we will never know if a greater good to society would have resulted, or if any damage to property, lives, or reputations would result. We have also seen frustration expressed by those witnessing the growth in computer crime and desiring something be done about it, and growing interest by researchers and defenders to respond to do just that. But there is insufficient guidance today for researchers to follow, or standards by which to judge research activities. There is even a question of whether academic or private researchers should actively be involved in computer crime activities solely for research purposes (as opposed to supporting protective or investigative activities.) [11]

More questions are typically raised about the ethics of computer security research activities than are answers provided, which can illuminate topics for future work. What constitutes risk, and who is placed in harms way? Do some research activities themselves come sufficiently close to a *use of force* that they warrant special consideration? Is federal regulation of research of computer crime activity necessary, similar to research into biological agents or toxins like anthrax, ricin, and smallpox (Public Law 107-188)? Could the Information Assurance concepts of *integrity, availability, and confidentialy* be used to untangle a complex mix of many inter-related actions when making risk/benefit evaluations? How might a scoring system be developed for uniformly evaluating risk, benefit, and appropriate actions?

To help understand these issues better and define a workable ethical framework, we believe that a more structured series of public discussions is urgently needed. We look forward to seeing these discussions accompanying future botnet research.

Acknowledgments

This work is supported in part by the Cisco Systems Critical Infrastructure Assurance Group (CIAG), the U.S. Department of Homeland Security Science & Technology Directorate under Contracts No. NBCHC060090, NBCHC080037, and AFRL cooperative agreement FA8750-08-2-0147 and by NSF under contracts CNS 0627445 and CNS 0831174. We wish to thank Fabian Monrose, Jose Nazario, Christopher Kruegel, John Heideman, Dorothy Denning, Kaiti Carpenter, Michael Collins, Rachna Dhamija, and the anonymous reviewers of an earlier version of this text for their valuable suggestions.

6. REFERENCES

[1] People v. Williams, 53 Misc. 2d 1086, 1090, 281 N.Y.S.2d 251, 256 (Syracuse City Ct. 1967).

[2] Protected repository for the defense of infrastructure against cyber threats (PREDICT). http://www.predict.org.

[3] ACM Council. Code of Ethics and Professional Conduct, Oct. 1992. `http://www.acm.org/about/code-of-ethics`.

[4] M. Allman. What ought a program committee to do? In *WOWCS'08: Proceedings of the USENIX Workshop on Organizing Workshops, Conferences, and Symposia for Computer Systems*, pages 1–5, 2008.

[5] M. Allman and V. Paxson. Issues and etiquette concerning use of shared measurement data. In *IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pages 135–140, 2007.

[6] S. Baase. *A gift of fire: social, legal, and ethical issues in computing*. Pearson Prentice Hall, 3rd edition, 2009.

[7] 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.

[8] M. Bond and G. Danezis. A pact with the devil. In *NSPW '06: Proceedings of the 2006 Workshop on New Security Paradigms*, pages 77–82.

[9] S. C. Bono, M. Green, A. Stubblefield, A. Juels, A. D. Rubin, and M. Szydlo. Security analysis of a cryptographically-enabled RFID device. In *Proceedings of the 14th USENIX Security Symposium*, pages 1–16, 2005.

[10] H. Burch and B. Cheswick. Tracing anonymous packets to their approximate source. In *LISA '00: Proceedings of the 14th USENIX conference on System Administration*, pages 319–328, 2000.

[11] A. J. Burstein. Conducting cybersecurity research legally and ethically. In *LEET'08: Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*, pages 1–8, 2008.

[12] T. W. Bynum and S. Rogerson. *Computer Ethics and Professional Responsibility: Introductory Text and Readings*. Blackwell Publishers, Inc., Cambridge, MA, USA, 2003.

[13] F. Castaneda, E. C. Sezer, and J. Xu. WORM vs. WORM: preliminary study of an active counter-attack mechanism. In *WORM '04: Proceedings of the 2004 ACM Workshop on Rapid Malcode*, pages 83–93, 2004.

[14] 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.

[15] S. E. Coull, C. V. Wright, F. Monrose, M. P. Collins, and M. K. Reiter. Playing devil's advocate: Inferring sensitive information from anonymized network traces. In *Proceedings of the Network and Distributed System Security Symposium*, pages 35–47, 2007.

[16] D. Danchev. Legal concerns stop researchers from disrupting the storm worm botnet, Jan. 2009. `http://blogs.zdnet.com/security/?p=2397`.

[17] 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.

[18] D. E. Denning. Concerning hackers who break into computer systems. In *Proceedings of the 13th National Computer Security Conference*, pages 653–664, 1990.

[19] D. E. Denning. The ethics of cyber conflict, June 2008. Chapter 17 in The Handbook of Information and Computer Ethics.

[20] D. Dittrich. How bad an idea was 'Make Love Not Spam?' Let me count the ways, Mar. 2005. http://staff.washington.edu/dittrich/arc/workshop/lycos-response-v3.txt.

[21] 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.

[22] 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.

[23] W. Enck, P. Traynor, P. McDaniel, and T. La Porta. Exploiting open functionality in sms-capable cellular networks. In *CCS '05: Proceedings of the 12th ACM conference on Computer and communications security*, pages 393–404, 2005.

[24] P. Finn and M. Jakobsson. Designing ethical phishing experiments. *Technology and Society Magazine, IEEE*, 26(1):46–58, Spring 2007.

[25] S. L. Garfinkel. IRBs and security research: Myths, facts and mission creep. In *Proceedings of UPSEC '08 (Usability, Psychology and Security)*, Apr. 2008.

[26] S. Gaudin. Plan to counterattack hackers draws more fire. http://www.internetnews.com/article.php/3335811, April 2004.

[27] D. Gotterbarn, K. Miller, and S. Rogerson. Software engineering code of ethics. *CACM*, 40(11):110–118, Nov. 1997.

[28] D. Halperin, T. Heydt-Benjamin, B. Ransford, S. Clark, B. Defend, W. Morgan, K. Fu, T. Kohno, and W. Maisel. Pacemakers and implantable cardiac defibrillators: Software radio attacks and zero-power defenses. In *IEEE Symposium on Security and Privacy*, pages 129–142, May 2008.

[29] T. S. Heydt-Benjamin, D. V. Bailey, K. Fu, A. Juels, and T. O'Hare. Vulnerabilities in First-Generation RFID-enabled Credit Cards. In *Proceedings of the 11th International Conference on Financial Cryptography and Data Security*, February 2007.

[30] K. E. Himma. *Internet Security: Hacking, Counterhacking, and Society*. Jones & Bartlett Publishers, 2007.

[31] 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.

[32] 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.

[33] ICRC. The Geneva Conventions: the core of international humanitarian law, January 2006. http://www.icrc.org/Web/Eng/siteeng0.nsf/htmlall/genevaconventions.

[34] IEEE Board of Directors. IEEE Code of Ethics, Feb. 2006. http://www.ieee.org/portal/pages/iportals/aboutus/ethics/code.html.

[35] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer. Social phishing. *Communications of the ACM*, 50(10):94–100, 2007.

[36] M. Jakobsson, N. Johnson, and P. Finn. Why and how to perform fraud experiments. *IEEE Security and Privacy*, 6(2):66–68, 2008.

[37] V. Jayaswal, W. Yurcik, and D. Doss. Internet hack back: counter attacks as self-defense or vigilantism? In *2002 International Symposium on Technology and Society*, pages 380–386, 2002.

[38] J. P. John, A. Moshchuk, S. D. Gribble, and A. Krishnamurthy. Studying Spamming Botnets Using Botlab. To appear in Proceedings of the 6th USENIX Symposium on Networked Systems Design and Implementation (NSDI '09), Apr. 2009.

[39] D. G. Johnson and K. W. Miller, editors. *Computers Ethics*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2009.

[40] D. Kaminsky. Black Ops 2008 – It's The End Of The Cache As We Know It. In *Black Hat Briefings USA 08*, Las Vegas, Nevada, USA, July 2008.

[41] 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.

[42] C. Kanich, K. Levchenko, B. Enright, G. M. Voelker, and S. Savage. The Heisenbot Uncertainty Problem: Challenges in Separating Bots from Chaff. In *LEET'08: First USENIX Workshop on Large-Scale Exploits and Emergent Threats*, April 2008.

[43] kc claffy. Ten things lawyers should know about the internet. CAIDA Web site, 2008. http://www.caida.org/publications/papers/2008/lawyers_top_ten/.

[44] S. Kelly. BBC team exposes cyber crime risk, Mar. 2009. http://news.bbc.co.uk/2/hi/programmes/click_online/7932816.stm.

[45] D. McCoy, K. Bauer, D. Grunwald, T. Kohno, and D. Sicker. Shining light in dark places: Understanding the Tor network. In *The 8th Privacy Enhancing Technologies Symposium*, pages 63–76, 2008.

[46] J. H. Moor. What is computer ethics? *Metaphilosophy*, 16(4):266–275, 1985.

[47] T. Mullen. The right to defend. http://www.securityfocus.com/columnists/98, July 2002.

[48] R. Naraine. Kraken botnet infiltration triggers ethics debate, May 2008. http://www.eweek.com/c/a/Security/

Kraken-Botnet-Infiltration-Triggers-Ethics-Debate/.

[49] B. Prince. BBC responds to botnet controversy, Mar. 2009. http://www.eweek.com/c/a/Security/BBC-Responds-to-Botnet-Controversy/.

[50] W. Rehnquist. Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993), June 1993. http://laws.findlaw.com/us/509/579.html.

[51] Several. Results of the Distributed-Systems Intruder Tools Workshop. CERT/CC, Dec. 1999.

[52] D. C. Sicker, P. Ohm, and D. Grunwald. Legal issues surrounding monitoring during network research. In *IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pages 141–148, 2007.

[53] 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.

[54] W. G. S. Sr. *CyberSpace and the Use of Force*. Aegis Research Corporation, 1999.

[55] 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, Aug. 5–9 2002.

[56] R. Vogt, J. Aycock, and M. J. J. Jr. Army of botnets. In *Proceedings of the 14th Annual Network and Distributed System Security Symposium (NDSS 2007)*, pages 111–123, February 2007.

[57] P. Wang, S. Sparks, and C. C. Zou. An advanced hybrid peer-to-peer botnet. In *HotBots'07: Proceedings of the First USENIX Workshop on Hot Topics in Understanding Botnets*, 2007.

[58] T. C. Wingfield. *The Law of Information Conflict*. Aegis Research Corporation, 2000.

[59] N. Wyler. *Aggressive Network Self Defense*. Syngress, 2005.

[60] W. Yurcik. Information warfare: Legal & ethical challenges of the next global battleground. In *Proceedings of the 2nd Annual Ethics and Technology Conference (Ethics'97)*. Loyola University Chicago, June 1997.

[61] W. Yurcik. Information warfare survivability: Is the best defense a good offense? In *Proceedings of the 5th Annual Ethics and Technology Conference*. Loyola University Chicago, July 2000.