Ethical Guidelines for Computer Security Researchers: "Be Reasonable"

Len Sassaman*

Katholieke Universiteit Leuven
len.sassaman@esat.kuleuven.be

1 Introduction

For most of its existence, the field of computer science has been lucky enough to avoid ethical dilemmas by virtue of its relatively benign nature. The sub-disciplines of programming methodology research, microprocessor design, and so forth have little room for the greater questions of human harm. Other, more recently developed sub-disciplines, such as data mining, social network analysis, behavioral profiling, and general computer security, however, open the door to abuse of users by practitioners and researchers. It is therefore the duty of the men and women who chart the course of these fields to set rules for themselves regarding what sorts of actions on their part are to be considered acceptable and what should be avoided or handled with caution out of ethical concerns. This paper deals solely with the issues faced by computer security researchers, be they vulnerability analysts, privacy system designers, malware experts, or reverse engineers.

The computer security researcher can do well to take a cue from the Hippocratic oath: first, do no harm. Of course, were the matter as simple as that, we would not as a community have discussed these issues for the past twenty-five years.

2 What is Harm?

Questions surrounding the nature of acceptable research activities in the computer security field have been a persistent feature of the field almost since its inception. Examples of well-intentioned breaches of ethics can be found throughout the early literature, perhaps none more notorious than the Morris worm of the 1980s [14]. While it is easy to see the lack of malice intended by Morris with his experiment, modern researchers would almost unanimously agree that releasing malware onto the public network, intended to exploit unpatched security vulnerabilities on systems not controlled by the experimenter, without the authorization or knowledge of the affected network operators, is itself unethical.

* Supported in part by the Concerted Research Action (GOA) Ambiorics 2005/11 of the Flemish Government, by the IBBT (Flemish Government), and the IAP Programme P6/26 BCRYPT of the Belgian State (Belgian Science Policy).

Despite that, there remains a significant, though minority, group of opinion-makers who propose effectively the same sort of conduct when it comes to fighting the present scourge of botnets, arguing that the Internet would benefit from the release of "benign worms" to patch the security holes being exploited by botnet authors; or from uninstall instructions issued from botnet command and control servers that have been seized by network administrators. When researcher Julia Wolf gained control of the command-and-control server domains for the largest spam botnet of 2008, Srizbi, many people asked why she did not use the botnet for "good" purposes [16]. Most commonly proposed uses involved removing the Srizbi bot from the affected clients with an update and patching the vulnerabilities that allowed those machines to become infected in the first place [15].

Similarly, when the Code Red worm infected over 359,000 machines on the Internet in 2001 [9], it was a matter of days before a programmer released a software package called "Code Green" – intended to "infect" already-infected Code Red nodes with a program that removed Code Red and patched the exploit [5]. Many network operators and security experts were quick to decry this approach both on ethical grounds, for it involved the installation of outside software without the consent of the administrators of the target machines, and for practical concerns, since it would necessitate a reboot of potentially critical systems with no means for the Code Green operators to address errors should they occur (or even identify "critical machines" so that they could avoid targeting them). Nevertheless, the idea of "killer worms" persists on certain Internet mailing lists and in academic communities. A paper on this topic by Wu et al. of Zhongshan University supporting this approach dismisses concerns about the potential criminal nature of such a worm by proposing a "centralized administrative power" with the authority to unleash "helpful" worms on systems without the necessity of consent [17]. The potential for unexpected behavior to cause damage is not acknowledged or discussed in that paper.

The Western legal doctrine of property rights has come to play a key role in guiding the actions of computer security professionals in cases like this — network operators are free to cease routing packets from worm-infested machines, dropping them from the net, but once an individual takes it upon himself to modify the restricted-access state of another party's system, he is in breach of ethical standards, as well as cyber-crime laws in many jurisdictions.

3 Software Vulnerability Analysis

For over a decade, a debate in the computer security community raged: what was the ethical course of action for a vulnerability researcher to take upon discovering an exploitable flaw in a piece of software? Many researchers espoused the doctrine of "full disclosure" — publication of their findings, including the necessary information to reproduce them, in accordance with the traditions of scientific research. Other parties, notably software vendors whose code these researchers were evaluating, argued that dissemination of knowledge on performing exploits put the general public at risk and helped facilitate criminal action.[1]

Our field has, in recent years, come to a happy medium, where responsible disclosure is considered to be the publication of reproducible vulnerability analysis after a "vendor notification period", sufficient to allow the affected software author time to verify, patch, and test fixes to their software's flaws. Details of exploits are then released after a patch has become available, and serve to encourage adoption of the bugfix.

The relatively straightforward process of responsible full disclosure has become more complicated in recent years, however. Web-based applications are more prevalent, meaning the act of vulnerability discovery is no longer conducted against an instance of code contained on a machine owned by the researcher, but often performed against a live server operated by a web service provider. In cases such as this, the researcher must tread lightly, for the simple act of discovering an exploit may mean performing that exploit on another party's system.

One of the first widely publicized cases of this problem involved a man popularly referred to as "the homeless hacker", Adrian Lamo, who took it upon himself to penetrate the computer systems of high-profile websites and companies, with the stated intent of improving their security [12]. By performing basic SQL injection attacks and other relatively simple techniques against websites such as that of the New York Times and Microsoft, he was able to gain access to information such as Social Security numbers and other confidential data. Lamo was initially unapologetic for his actions, insisting he was performing a public service. Popular opinion was mixed, with some members of the press going so far as to defend his theft of multiple Lexis/Nexis database logins, which he used to perform vanity searches [11]. The opinion of this author, however, is that Lamo knowingly and with premeditation conducted unethical and illegal intrusions into private computer systems, stole the resources of other parties for his own personal gain, and regardless of any security enhancements he may have performed on the systems he attacked, his actions did not render a service to his victims, who were left to audit the compromised systems, pay the usage fees he accrued, and engage in time-consuming legal proceedings[2] resulting as a direct consequence of Lamo's crooked moral compass. Regardless of motive, behavior such as this should be condemned.

Numerous individuals have run afoul of the law during the course of their investigations of website security. The technology media routinely reports stories of well-intentioned individuals stumbling upon (or seeking out) a security flaw in a website, reporting the problem to its administrators, and then facing prosecution for bypassing the website's security. Similarly, research into physical security that relies on computer science has its risks; after publishing a proof-of-concept

[1] As recently as 2003, a Microsoft representative took the stance that vulnerability disclosure was reckless and irresponsible [3].

[2] As Lamo's actions violated criminal statutes, the victims were not in a position to choose whether to sue or not — they could simply have been called as witnesses for the prosecution, leaving them little option but to comply.

that demonstrated weaknesses in airport checkpoint security, graduate student Christopher Soghoian had his house raided by the FBI, his computers seized, and was subject of an investigation after a congressman called for his arrest [6, 2]. While that case was dropped, this and similar cases illustrate that caution when conducting evaluations of information security is not only important from an ethical standpoint, but also prudent from a legal perspective.

4 Responsible Data Handling

The simple line between "my system" and "another party's system" is not the end of the matter, however. Operators of services that process user data, or have the ability to collect information on usage patterns for specific users, have an obligation to those users. Examples of well-intentioned mismanagement of private user data include the notorious "Netflix Challenge", where online video rental site Netflix published a large data set containing user commerce logs with the intent that researchers would develop a better recommendation system for the service. The company removed the obvious identifying information prior to releasing the data, but researchers were able to cross-correlate the user preferences in that dataset with other information and deanonymize the user data [10]. AOL made a similar mistake when it released logs of queries to its search engine [4].

Other recent incidents involve the willful violation of user privacy in the name of research. Swedish independent researcher Dan Egerstad operated a server for the popular anonymity network Tor during 2007. His intent, rather than to provide a service to his users, as would be reasonable to expect, was instead to use his server as a surveillance mechanism. He published user login credentials he gathered by listening to the cleartext network traffic exiting from his server. His experiment added little to the literature, for the attack he performed was generally common sense, had already been published [13], and his publication of the actual login information (rather than, for example, collated statistics on usage of insecure protocols over Tor, without identifying the victims of his attack or their passwords) was highly ethically unsound.

5 Summary

In this paper, we have discussed some of the ethical conflicts researchers and other members of the software and network security communities have encountered in recent years. We have drawn attention to specific instances of ethical failings, as a means of illustrating the problems that can arise when the ethical course of action is not perfectly clear, or the community neglects its responsibility to enforce ethical standards for itself and its members. Despite the obvious impropriety of some of the examples given, it is almost certainly the case that the people involved in ethically improper actions, or who have proposed ethically questionable schemes, were not engaging in a conscious, willful breach of ethics when doing so. It is essential that we as a community hold debates and discussions concerning the ethics of the choices before us, for they are rarely black and white, and it is through such critical examination that we develop the standards by which we hold ourselves and each other accountable to the collective conscience of our field. Dilemmas by their nature are not easily resolved, and thus such self-examination is a necessary part of our development as a community respectful of others. In her 1988 paper, Campbell suggests that a major cause of ethical lapses is a lack of this type of introspection [1]. Whether researchers choose to adhere to a specific set of ethical guidelines set forth by their peers, institutions, or professional organizations[3], or simply look inward to answer the question "Is what I am doing *good?"*, the question must be asked, if for no other reason than the sake of the questioner's own happiness. To paraphrase Leibniz, perhaps the first computer scientist to ponder questions of ethics, the more a man desires to know virtue in his quest for knowledge, and the more inspired he is to incorporate virtue in his life, the happier his life will be.[4]

As computer security researchers, we have a duty to advance the state of the art of secure systems, encourage their adoption, and identify weaknesses in currently deployed software, protocols and systems. We must do so in a manner that balances overall improvement in system security, correction of specific security concerns, and advances in the general foundation of our discipline without endangering current users or putting existing deployed systems at unnecessary risk of attack. Sometimes it is not clear where to strike such a balance, but a responsible researcher will respect the need for open discussion of security issues (including offensive techniques) while attempting to accommodate the needs of at-risk services and vulnerable software vendors, as well as the users reliant upon their software for their security, privacy, industry, and peace of mind.

This discourse on ethical conduct in computer security research is critical, and must be revisited frequently as threats change and technology advances. We must answer our calling as scientists, to pursue knowledge for its own sake, which is justification enough to seek new methods and techniques for uncovering attacks on computer systems, or applying known attacks to systems whose flaws have not yet been fully excavated, upon which the light of understanding has not yet shined brightly enough to illuminate all of their unintended operation potential.

As scientists, we have a duty to preserve academic freedom — but with care, we can exercise that freedom in a responsible manner, and (if we believe Leibniz) by conducting ourselves virtuously, find happiness through our quest for knowledge in this tiny slice of all that which awaits our knowing it.

[3] Examples of formal ethical guidelines can be found in [1]; additional such treatises have been authored in the twenty-two years since the publication of that paper, but they share a common goal: the codification of standards of conduct compatible with ethical action for the purpose of informing and guiding the individual's conscience.

[4] "Il faut tenir pour asseuré que plus un esprit desire de connoitre l'ordre, la raison, la beauté des choses que Dieu a produites et plus il est porté à imiter cet ordre dans les choses que Dieu a abandonnées à sa conduite, plus il sera heureux." [7]

6 Acknowledgements

The author would like to thank the people, too numerous to list, who have helped shape his views on ethical conduct, both in the field of computer security, and as a human being. *Duplicatur autem jucunditas reflexione, qvoties contemplamur pulchritudinem ipsi nostram, qvod fit conscientia tacita virtutis nostrae. Sed qvemadmodum duplex in visu refractio contingere potest, altera in lente oculi, altera in lente tubi, qvarum haec illam auget, ita duplex in cogitando reflexio est, cum enim omnis mens habeat speculi instar, alterum erit in mente nostra, alterum in aliena, et si plura sint specula, id est plures mentes bonorum nostrorum agnitrices, major lux erit, miscentisbus speculis non tantum in oculo lucem, sed et inter se, splendor collectus gloriam facit* [8].

References

1. Marlene Campbell. Ethics and Computer Security: Cause and Effect. In *CSC '88: Proceedings of the 1988 ACM Sixteenth Annual Conference on Computer Science*, pages 384–390, New York, NY, USA, 1988. ACM.
2. Jennifer Carmack. IU Student, Focus of FBI Probe, Speaks Out, 2006.
3. Stanford Cyberlaw Clinic. CyberSecurity, Research and Disclosure, 2003.
4. Katie Hafner. Researchers Yearn to Use AOL Logs, but They Hesitate. *The New York Times*, 2006.
5. Herbert HexXer. CodeGreen Beta Release, September 2001. http://archives.neohapsis.com/archives/vuln-dev/2001-q3/0575.html.
6. Brian Krebs. Student Unleashes Uproar with Bogus Airline Boarding Passes. *The Washington Post*, 2006.
7. G. W. Leibniz. *Textes Inédits D'après les Manuscrits de la Bibliothèque Provinciale de Hanovre*, chapter La Félicité. Presses Universitaires de France, 1948.
8. G. W. Leibniz. *Philosophische Schriften*, chapter Elementa Juris Naturalis. Akademie Verlag GmbH, 2006.
9. David Moore and Colleen Shannon. The Spread of the Code Red worm, 2008. http://www.caida.org/research/security/code-red/coderedv2_analysis.xml.
10. Arvind Narayanan and Vitaly Shmatikov. How To Break Anonymity of the Netflix Prize Dataset, 2006.
11. Annalee Newitz. TECHSPLOITATION: Subpoena Me, Too! *San Francisco Bay Guardian*, October 2003.
12. Kevin Poulsen. Feds say Lamo Inspired Other Hackers. *The Register*, 2004. http://www.theregister.co.uk/2004/09/16/feds_on_lamo/.
13. Len Sassaman. The Faithless Endpoint: How Tor puts certain users at greater risk. Technical Report ESAT-COSIC 2007-003, Katholieke Universiteit Leuven, 2007.
14. Eugene H. Spafford. The Internet Worm Program: an Analysis. *SIGCOMM Computer Communication Review*, 19(1):17–57, 1989.
15. Julia Wolf. Technical Details of Srizbi's Domain Generation Algorithm, 2008. http://blog.fireeye.com/research/2008/11/technical-details-of-srizbis-domain-generation-algorithm.html.
16. Julia Wolf and Len Sassaman. Unpublished manuscript, December 2008.
17. Dan Wu, Dongyang Long, Changji Wang, and Zhanpeng Guan. Modeling and Analysis of Worm and Killer-Worm Propagation Using the Divide-and-Conquer Strategy. In Michael Hobbs, Andrzej M. Goscinski, and Wanlei Zhou, editors, *ICA3PP*, volume 3719 of *Lecture Notes in Computer Science*. Springer, 2005.