INTERHACK

Interhack
5 E Long St Ste 1101
Columbus, OH 43215

VOX +1 614 545 HACK
FAX +1 614 545 0076
WEB http://web.interhack.com/

Spying on Spyware

C Matthew Curtin, CISSP

Central Ohio Chapter of ISSA
July 21, 2004

Abstract

Millions of computer users are being watched, not just by employers and auditors, but by the software that they use—frequently without their knowledge or consent. This "spyware" has become the center of the personal privacy debate and threatens to undermine efforts to keep corporate data secured.

What exactly is spyware? How does it work? What is its impact on users—and the businesses that employ them?

Interhack's Internet Privacy Project has been pioneering the dissection and documentation of spyware since 1999.

Id: spyware-wp.tex,v 1.2 2005/11/02 18:54:57 cmcurtin Exp

1 Introduction

Software to observe user behavior to collect information under users' noses is often called *spyware*. These systems have become central to a heated debate regarding online privacy, prompting the U.S. Congress to consider several bills.[1] In addition, the very nature of such systems—the collection of data that would not otherwise be available outside of corporate firewalls—raises questions about how companies can remain compliant with privacy-oriented regulation like the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and the Gramm-Leach-Bliley Financial Modernization Act of 1999 (GLBA).

1.1 What is Spyware?

In its most simple form, spyware is software designed to collect information from computer system users without their knowledge. Typically, spyware can be classified as a type of trojan horse, which is a type of technology-based security incident, allowing for information security policy violation. Figure 1 shows where spyware fits within the broader context of policy enforcement.

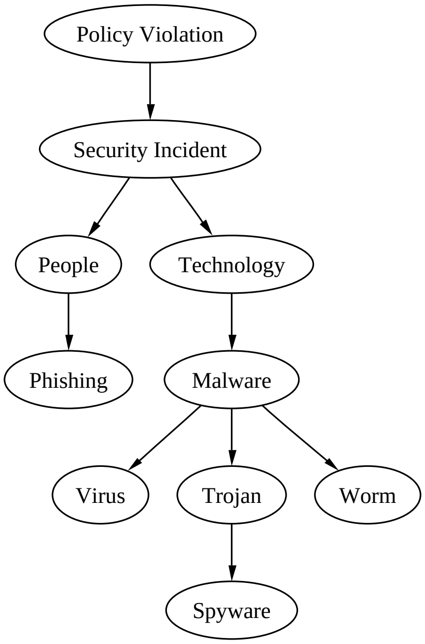

Figure 1: Where Spyware Fits

[1] As of this writing, several bills that have been presented in the 108th Congress that either address directly or mention spyware, including "Safeguard Against Privacy Invasions Act" (H.R. 2929), "Internet Spyware (I-SPY) Prevention Act of 2004" (H.R. 4661), "Piracy Deterrence and Education Act of 2004" (H.R. 4077).

2 How Spyware Works

In this paper, we'll briefly outline two systems that could be classified as spyware to demonstrate different methods for collecting information from users without their knowledge.[2]

In both cases, these systems perform some kind of surreptitious user tracking and then format some part of that data for reporting back to system's operator. It should be noted that there are significantly more egregious cases of spyware in use; we choose these two systems because they represent a Windows-based system that collects and reports information and a Web-based system to do the same. Other cases that we have analyzed include *Spector Professoinal* [5], TheCounter.com [2], Coremetrics [7], DoubleClick [8, 9], and Netscape [6].

2.1 PCFriendly

PCFriendly is an application that shipped on numerous DVD titles between 1996 and 2000. In addition to its stated objective (providing a software-based DVD player for Windows machine), the system collected information about the user and the user's DVD collection, occasionally reporting such things back to InterActual Technologies, the maker of *PCFriendly*.

PCFriendly is a Windows-based application that starts when a DVD is inserted into the system's DVD player. The first time that the application starts, the user is asked for information like name, address, email address, and age. A unique identifier is assigned to the user, and the application appears to track changes over time, for example, additional DVD titles put into the system.

As of Interhack's last look at the system (in May 2002), *PCFriendly* was being replaced by a new system, known as *InterActual Player*, written to address privacy concerns, among other things.

Defenses that we identified at the time essentially meant breaking *PCFriendly* functionality:

- Do not watch DVD titles on a computer. Backchannels are easy to implement on systems with Internet connectivity.
- Upgrade to the latest *InterActual Player*. Privacy problems with Web-based content provided by the DVD producers (as opposed to InterActual Technologies) will not be addressed in this case, making this solution incomplete.
- Block all access to `pcfriendly.com` and `interactual.com` domains.

Note that in our second option—using the *InterActual Player*—only meant using newer software designed to give the user a greater number of options for protecting privacy—default behavior was still invasive.

Detailed analysis and discussion is available in the Interhack Research Technical Report, "PCFriendly Enables DVD Backchannels". [4]

[2] It should be observed that these systems might or might not ultimately be considered "spyware" by a legal definition; this is a technical, not a legal white paper.

2.2 Pharmatrak

Pharmatrak was a company that provided a Web site tracking and reporting service to pharmaceutical companies. Its system works much like the Coremetrics system Interhack analyzed in 2000 [7], with two critical differences. Interhack provided forensic analysis to plaintiffs' counsel in the Pharmatrak privacy litigation. Facts regarding the operation of Pharmatrak's service are identified in court documents. [10]

The first difference was that Pharmatrak did not have JavaScript code that was designed to pull users' responses to form data out of the form and to put them into a request for a Web bug. Pharmatrak's entire collection mechanism was predicated on collecting HTTP `Referer` [sic], though it did go to some significant lengths to get the data—including the use of JavaScript (and even a Java applet in the earliest instance of the software).

The Pharmatrak system was designed to collect information about users of pharmaceutical companies Web sites. The users would be pseudonymously tagged, and their activity observed and reported back to the pharmaceutical company. In addition, very high-level information (such as total traffic) would be reported to other pharmaceutical companies that were Pharmatrak clients, allowing each Pharmatrak client to see not only detailed information about its site activity, but high-level information about its competitors' sites. All of this happened with the knowledge and consent of the pharmaceutical companies that hired Pharmatrak to perform the reporting service and implanted the Pharmatrak-supplied code on their sites.

This leads us to our second difference: Pharmatrak was not authorized by its clients to collect personally-identifiable information, and by all appearances, Pharmatrak did not have specific intent to collect such information. (Forensic investigation and analysis showed that they did have detailed personal information on several hundred users.)

Detailed information on the mechanisms for client and server interaction on the Pharmatrak system can be found in court documents. Interaction between Web browsers and clients and how these impact user privacy is described in detail in *Developing Trust: Online Privacy and Security*. [3]

3 Why Spyware Works

Spyware fundamentally requires that a system that the user assumes to be trustworthy is operating under the direction or influence of a third party. We're dealing with the notion of "trusted computing base" here, which essentially means all hardware, software, and procedure used to enforce security policy. One of these components—software—is being subverted to allow the spyware provider to observe the user's behavior surreptitiously.

Note that when we talk about trust in this context, we're not using the same term that is used by some computer manufacturers now, in particular the Trusted Computing Group alliance and related efforts that have been known by such names as NGSCB, Longhorn, and Pallidum. Those systems are designed to make media publishers able to trust your system as a playback device that is under *their* control;

these can actually *break* your security policy. [1]

4 Strategies for Effective Mitigation

There are two primary methods to deal with spyware: the first is to look to the host (computer that could have spyware installed) and the second is to look at the network.

4.1 Host-Based Solutions

The host-based solution will provide several valuable options. The best of which is *prevention*. By using systems that are not vulnerable to the kinds of attacks that spyware—particularly the nasty variety not discussed here—one will gain a measure of protection; the vulnerabilities in ActiveX, for example, that enable such problems are simply not present in other operating systems like MacOS, Linux, and FreeBSD. Note that not all spyware works by ActiveX controls, however—the Pharmatrak system worked for any Web-based system; users of these systems would (and, indeed, did) have information about them collected.

Another host-based option is to create a standard "build" of the desktop system for users that includes not only the operating system and applications, but also defense mechanisms such as anti-malware packages.

A significantly less effective mechanism is spyware "removal." While this might appear to be a more attractive solution than prevention in some cases (because there is no need to justify the expense of an anti-spyware package on the grounds that such a threat might materialize in the future), it should be noted that any software running on a system that has been compromised might not be able to behave as advertised. In particular, malware that changes operating system libraries could cause a "removal" program to do more damage than harm to the system in question. The safest option in the event of a system compromise is to throw away the compromised installation and to replace it with one that can be trusted—which takes us back to the standard build option mentioned earlier.

4.2 Network-Based Solutions

Another option is to take a network-based view of the system. That is, to configure intrusion detection systems, firewalls, and other policy enforcement mechanisms to prevent spyware packages from working.

The first means of doing this would be to identify unsafe content (e.g., ActiveX controls) flowing from an untrusted zone (e.g., the Internet) into a trusted zone (e.g., an internal network) and blocking the download. Another means would be to identify attempts of spyware to "phone home," effectively preventing them from being able to report their activity, but not preventing the spyware from hitting the user's system in the first place. A third mechanism would be to enforce a policy that refuses connectivity from trusted systems to unknown sites or to allow downloads of unidentifiable content.

5 Conclusions

Spyware, though not a particularly new problem when defined generally, remains a problem that is difficult to manage. While there is no silver bullet to solve all of these problems, there is hope. Like other security incidents, the problem can be managed effectively with a comprehensive definition of the trusted computing base and a program to maintain it. With the right support from policy and technology, malware, including spyware, can be defeated. □

To see how Interhack can help you to define and to enforce security policy, please visit us online at web.interhack.com or call us at +1 614 545 HACK.

References

[1] Ross Anderson. 'trusted computing' frequently asked questions. Technical report, University of Cambridge, 2003. [online] http://www.cl.cam.ac.uk/~rja14/tcpa-faq.html.

[2] Matt Curtin. A failure to communicate: When a privacy seal doesn't help. Technical report, Interhack Corporation, August 2000.

[3] Matt Curtin. *Developing Trust: Online Privacy and Security*. Apress, November 2001.

[4] Matt Curtin. Pcfriendly enables dvd backchannels. Technical report, Interhack Corporation, February 2002. [online] http://www.interhack.net/pubs/pcfriendly/.

[5] Matt Curtin. Spector pro review and commentary. Technical report, Interhack Corporation, May 2002. [online] http://www.interhack.net/pubs/spector/.

[6] Matt Curtin, Gary Ellison, and Doug Monroe. "What's Related?" Everything But Your Privacy. Technical report, The Ohio State University, Department of Computer and Information Science, October 1998.

[7] Matt Curtin, Paul Graves, and Shaun Rowland. Getting to know you (intimately): Surreptitious privacy invasion on the e-commerce web. Technical report, Interhack Corporation, July 2000.

[8] Gary Ellison, Matt Curtin, and Doug Monroe. DoubleClick Opt Out Protocol Failure == Opt In. Technical report, Interhack Corporation, May 2000.

[9] Gary Ellison, Matt Curtin, and Doug Monroe. Opting In, By Accident. Technical report, Interhack Corporation, May 2000.

[10] United States Court of Appeals For the First Circuit. IN RE PHARMATRAK, INC. PRIVACY LITIGATION, May 2003. [online] http://www.ca1.uscourts.gov/cgi-bin/getopn.pl?OPINION=02-2138.01A.