# DUMPSTER DIVING: A STUDY ON DATA RECOVERY AND EXPLOITATION

***Robert M. Weaver, Joseph A. Cazier***[1]

Appalachian State University

## Abstract

*Social Engineering is a 'low tech' method of attack that involves obtaining personal information and using it to hack into a system [1]. Social Engineering takes on many forms, including dumpster diving. Dumpster diving is where one goes through trashcans and dumpsters looking for information such as IP addresses, usernames, passwords, and other information [2, p. 63]. Individuals who set their old computers on the curb for trash pickup, companies who simply toss them in the dumpster or entities that donate their old computers to charity should think about what information is available on the hard drives and take every step possible to eliminate that information before disposal of the computer. Increased awareness of identity theft and corporate stealing has brought the issue of protecting one's personal information. However, do most people take a real effort to destroy data remaining on their hard drives? This research in progress study will attempt to identify the types of information found and/or recovered from hard drives in computers donated to charity and/or set out for trash pickup. The results of this study will be provided at the conference. In addition to the results, the authors will also make recommendations as to what the average individual can do to protect their personal data as well as changes to corporate security policies to ensure protection of corporate data and personal information of their employees.*

## INTRODUCTION

Today, many large companies either wipe their hard drives in house or they contract out to computer recycling companies when disposing of their old computers. However, it is incredibly easy for a computer to slip through the cracks, or for a hard drive not to be wiped sufficiently. Twenty-five percent of systems audited by Redemtech, a leading recycler of PC and IT products, still have data on them even though IT personnel thought the system had been wiped clean [4].

It is not known exactly what the majority of small businesses do with their old hard drives. One assumes that because they do not have access to the same resources as larger corporations, that they probably do not wipe their hard drives sufficiently clean, if at all, before disposing of them. In the consumer sector, often times, when consumers dispose of their personal computers, they have not even been an attempt to remove data from hard disks. There are several ways to destroy the data on hard disks, some are simple, and others are more complicated.

[1] Corresponding author

When data is written to hard disks, the disk drive head is smaller than the track which it must write, so when data is erased, similar to the furrows left when plowing a field, the drive head spews the "erased" data to either side of the head, often intact and recoverable. A single pass with a program designed to wipe a hard disk clean of any data will often result in data left on the disk. To be effective it takes between 3 and 7 passes with the program to consider a drive fully erased. The more passes one makes with the program, the less of a furrowing effect is left on the drive, thus effectively eliminating more of the recoverable data with each pass of the drive head [4].

Many hard drives built after 2001 have a built-in program for securely erasing data, Secure Erase. It is accessed through a series of commands embedded in the drive, however, before it can be used, it has to be enabled in the motherboard BIOS, as most of the time it is disabled by default. This program works by overwriting every track on the hard drive. Most areas not touched by a simple deletion include "bad blocks", directory structure, tracks not touched by the operating system, and unformatted sections of the disk, all of which can be touched by this embedded hard disk utility. There is external block overwriting software that can be purchased, there is now a free open source version called "Boot and Nuke" as well as a freeware version of Secure Erase [3].

As has become apparent post-September 11$^{th}$, recovered data, especially that from government agencies, officials, and employees, as well as from government contractors and their employees, can jeopardize our national security and freedom if it is obtained by terrorist organizations. Many agencies in the federal government such as the FBI and the Department of Defense simply remove the hard drives and drill holes through them, effectively rendering the drive inoperable and shredding the data. Physical destruction of the disk is the ultimate surefire way to render data irrecoverable, but also destroy a potentially useful asset and prevent the computer from being sold or donated to those in need.

When computers have been disposed of by companies, or individuals, it can be relatively easy to recover the data off the disk. In many cases it simply requires booting the computer and getting around the security, if any, to browse the files on the disk. Other times it requires that special data recovery programs and/or hardware be used.

It is not just data containing personal and/or corporate information that can be exploited. Files containing video and audio footage, pornography, blogs and diaries, e-mails, and instant messenger conversations can prove to be equally damaging and more easily exploited, especially if it can be linked to an individual.

In some cases hard drives that have been insufficiently erased to ensure irrecoverable data are a violation of federal laws and/or regulations. For example, the healthcare industry is largely governed by the Health Insurance Portability and Accountability Act of 1996 (HIPAA). HIPAA states that hospitals and other "covered" healthcare entities must ensure that patient information is available only to the patient and those designated in writing by the patient [5]. Any entity found in violation of HIPAA can, depending on the violation, face fines up to $250,000 and up to ten years of in prison per violation [6].

## METHODOLOGY

Hard drives will be collected from thrift stores, goodwill, yard sales, and curbside trash piles in residential neighborhoods during the summer of 2007. The drives will then be analyzed first by simply browsing the programs and files on the disk and recording the types of data found. Once the first analysis is complete, file recovery software will be run on the disk to see what remains of previously deleted files.

Following a pilot analysis, types of information expected to be found during analysis includes social security numbers, names, addresses, credit card numbers, usernames, and ultimately passwords. Other items expected to be found include, but is not limited to e-mails, blogs, and pornography.

All identifying information is necessary for the research and will remain confidential, known only to the authors. The author(s) have an ethical responsibility to turn over to the proper authorities any information found regarding illegal activities.

Results, future discussion, recommendations and future research will be discussed at the time of presentation.

## PRELIMINARY CONCLUSION

Some of the data found in the pilot analyses included full names, usernames, and e-mail addresses of the individual(s) who used the computer, as well as personal documents, websites, and files of a pornographic nature which were linked to individual e-mail addresses containing the individual's full name. All of the information listed would be useful to a hacker in some way. The e-mail addresses could be used to initiate phishing and spamming attacks, the full names in addition to the usernames could aid in identity theft, and the pornography could be used as leverage against the individual linked to it, as it contained embarrassing content.

In Today's age of ID theft, cyber crime, and terrorism, too many people are leaving data unguarded. This paper will raise awareness of the massive security issues surrounding unguarded data left on computers disposed of by consumers, corporations, and governments. In some cases some methods of data elimination may prove inadequate. This paper will make recommendations to ensure that there is minimal, if any, security risk when the computer is ready for disposal.

Complete data and discussion will be provided at the conference.

## REFERENCES


[1] Botelho, C. M. & Cazier, J. A. (2007) Social Engineering's Threat to Public Privacy. In *Assuring Business processes*, Proc. of the 6th Annual Security Conference, April 11-12. Ed. G. Dhillon. Washington DC: Global Publishing, USA.

[2] Easttom, Chuck. Computer Security Fundamentals. Security Ser 1. Upper Saddle River, NJ: Pearson Education, Inc., 2006.

[3] Harris, Robert. Blog Entry. 1 May 2007. 23 May 2007 <http://blogs.zdnet.com/storage/?=129&tag=nl.e622>.

[4] Mitchell, Robert L. "Dawn of the Undead Data." Computerworld 15 Dec. 2003: 36. ABI/INFORM Global. ProQuest. Western North Carolina Library Network, Boone, NC. 31 May. 2007 <http://0-www.proquest.com.wncln.wncln.org:80/>.

[5] Weeks, Katie. "Hospitals Protect Data By Erasing Old Hard Drives." San Diego Business Journal 27 Mar. 2006: 6. ABI/INFORM Dateline. ProQuest. Western North Carolina Library Network, Boone, NC. 31 May. 2007 <http://0-www.proquest.com.wncln.wncln.org:80/>.

[6] United States. Dept. of Health and Human Services. Office for Civil Rights Privacy Brief. Summary of the HIPAA Privacy Rule. Washington: GPO 2003.