

MIT: Discarded hard drives yield private info

By Paul Roberts, IDG News Service
JANUARY 16, 2003

Planning to auction off that old computer on eBay? You might want to take a closer look at what's on your disk drive first.

According to a new study by two MIT graduate students, companies and individuals are frequently selling or giving away old computer disk drives with sensitive information still on them.

The study, which is detailed in a report called "A Remembrance of Data Passed: A Study of Disk Sanitization Practices," analyzed 158 disk drives purchased through eBay Inc.'s online auction site eBay.com, at computer stores, salvage companies and swap meets.

The study found that 117 (74%) of the drives contained old data that could be recovered and read. Twenty eight of the drives (17%) contained fully installed and functional operating systems with user data that required no particular effort to recover. And while 57 (36%) of the drives had been freshly formatted, they still contained old data that could be recovered, according to the report.

Only 12 disk drives (9%) had been properly cleaned (or "sanitized") before being purchased by the students; 29 of the 158 drives purchased didn't work at all.

Among the sensitive information retrieved from the disk drives were detailed personal and corporate financial records, medical records, love letters and gigabytes worth of personal e-mail and pornography, according to a statement released by MIT.

Financial log files on one drive yielded what appeared to be 2,868 credit card numbers in addition to bank account numbers, dates of transactions and account balances. The students believe that the drive came from an ATM in Illinois and that no effort was made to remove any of the financial information on the drive prior to resale, according to the report.

Another drive that had been reformatted still contained 3,722 credit card numbers in what appeared to be a log file, while other drives yielded financial information that had been stored in cached Web pages that were recovered, the report said.

Other findings struck a more personal chord.

The students recovered an anguished letter from a father to his son's doctor expressing dissatisfaction over the course of the son's cancer treatment. The letter was stored on a disk drive that originally belonged to the father's employer, a now-defunct Seattle software company, according to Simson Garfinkel, a graduate student at MIT's Laboratory for Computer Science, who conducted the study along with classmate Abhi Shelat.

"Some of the data we got was so disturbing, I just stopped wanting to look at the data on the drives," Garfinkel said.

The recovered data problem stems from failures on the part of vendors and consumers alike, Garfinkel said.

Companies such as Microsoft Corp. that make operating systems are guilty of misrepresenting their products' "file delete" and "disk format" features, according to Garfinkel.

Casual computer users often assume that, when used, such features permanently delete the data stored in a file from the computer's disk drive. Instead, most operating systems simply change the data to indicate that the file has been deleted, then mark the areas of the hard disk that contain the "deleted" data as being available for reuse by other programs.

Assuming that data isn't overwritten by another program, it remains there undisturbed and can be retrieved and read using a variety of techniques, ranging from simple Unix commands to free and commercial forensic software tools, Garfinkel said.

Operating system vendors should include software-based tools that securely delete files and sanitize the disk space they leave behind, the report said.

The commonly used Microsoft format commands such as fdisk, for example, verify the integrity of the disk drive blocks but don't erase files, Garfinkel said.

The manufacturers of disk drives are also to blame for not embracing existing technologies such as cryptographic subsystems, which encrypt information using a secret key as it is written to the hard disk, then decrypt it when it needs to be viewed, according to the report's authors.

Data on such drives could be quickly rendered unintelligible by securely erasing the secret key.

Garfinkel admitted that he may have a personal stake in such changes, however. He has filed for a patent on an encryption technique that could be used to protect data stored on disk drives.

Finally, consumers need to be better educated about the proper techniques for erasing the data stored on their computer disk drives, while companies and organizations should develop and deploy policies for sanitizing storage media that are sold, donated or reused, the report said.

The potential financial and legal ramifications of old data stored on unused areas or "blocks" of the disk drive could be huge, Garfinkel said.

"Let's say I buy a formatted drive. It comes up, and I don't see any files. I run [the DOS format command] myself and don't see any bad blocks. I could use that drive for years and never know that it has child pornography on it because that data wouldn't be part of blocks that I was looking at. But if somebody came in and used forensic tools, that data is still there," he said.

And with the growth of the secondary computer hardware market, the handling of old data may be an area that requires government oversight, Garfinkel said.

While financial institutions and health care organizations are required to follow certain procedures in disposing of customer and patient data, no such laws apply to companies that don't do business in those areas, let alone to individuals.

National security could be at risk as well, Garfinkel said.

"All I can say is that when I bought [used] drives on eBay, there was always somebody bidding against me. Now I don't know who's buying them or why, but if I were a foreign government interested in doing economic espionage on the U.S., it seems to me that buying disk drives would be a cheap way to get a lot of good stuff," he said.

The study, which will be published in the January/February 2003 inaugural issue of IEEE Security and Privacy, identifies a number of tools that individuals can use to clean files from their disk drives. Those tools range from free programs available on the Internet to complex forensic tools that sell for more than \$2,000.

One inexpensive method for sanitizing drives is a free program called Autoclave, according to Garfinkel. Available on the Internet, Autoclave can be copied to a compact disk or floppy disk and inserted into a computer drive. When the machine is rebooted, Autoclave erases the entire disk, overwriting the disk contents with zeroes, according to the report.

In the meantime, with more than 150 million disk drives retired from their primary use in 2002 alone, it's likely that a large quantity of potentially sensitive data is floating, unprotected, in the secondary disk drive market.

A quick poll of individuals selling disk drives on eBay revealed that most had formatted or planned to format their drives before shipping them out using tools such as those provided by Microsoft in DOS and Windows XP, but didn't mention other strategies for erasing the disks' contents.