Hard-drive Disposal and Identity Fraud

Paula Thomas and Theodore Tryfonas
Information Security Research Group
Faculty of Advanced Technology
University of Glamorgan
Wales, UK
pthomas@glam.ac.uk, ttryfona@glam.ac.uk

Abstract. A personal computer is often used to store personal information about the user. This information may be intentionally kept by the user or information maybe automatically stored as the result of the user's activities. In this paper we investigate whether it is possible for identity fraud to occur as a result of post-disposal access to the residual data stored on a personal computer's hard drive. We provide indicative types of information required to commit an identify fraud and examine the personal information contained in a series of second-hand personal computer hard disk drives, purchased as part of a wider research study.

1 Introduction

A large amount of personal information can be found available on the Internet today. Public access documents found online, such as the voting register, used by themselves or in combination with other documents, can give an amount of information that may be more than enough for a fraudster to commit various criminal activities [1]. This availability of personal information is one of the main enablers for the crime that can be identified as Identity (ID) Fraud. The US Federal Trade Commission defines ID fraud as "a fraud committed or attempted using the identifying information of another person without authority" [2].

This can be accomplished in a number of ways depending on the information available to the perpetrator and the objective of the ID fraud. From the perspective of the perpetrator, identity fraud has to provide some form of gain that is usually, but not always, financial. Therefore, the identity fraud perpetrator is interested in acquiring specific pieces of personal information, some of which may need to be combined with other personal information in order to be of value. Some of the most common types of personal information that may be useful and the purpose for its collection are indicatively highlighted in Table 1 below.

Please use the following format when citing this chapter:

Thomas, P. and Tryfonas, T., 2007, in IFIP International Federation for Information Processing, Volume 232, New Approaches for Security, Privacy and Trust in Complex Environments, eds. Venter, H., Eloff, M., Labuschagne, L., Eloff, J., von Solms, R., (Boston: Springer), pp. 461–466.

Table 1. Examples of the use of various types of personal information for identity fraud

Type of personal info use	Information required	Impact on the individual
Registering for mail / on-line services	Name (surname, first names, possibly maiden / previous names), Address / postcode	Reputation
Phishing	Email Address, Possible passwords	Financial
Taking control of existing bank accounts / Obtaining Loans	Name (surname, first names, possibly maiden / previous names), Address / postcode, Other family details e.g. mother's maiden name, photographs, Bank Account Information (Bank, Account Number, etc)	Financial
Opening New Bank accounts Requesting loans	Name (surname, first names, possibly maiden / previous names), Address / postcode	Financial

Personal information could be used in a number of ways to commit different forms of identity fraud and could range from taking over an existing bank account to opening a new account or obtaining a loan in the victim's name. For example, in order to apply for a loan in a U.K. context, the ID fraud perpetrator would require someone's full name, other family details, a bank account number, branch details etc.

Information items listed in Table 1 may well exist in electronic form within a personal computer. Computer security countermeasures are deployed exactly in order to protect those, amongst other information assets. Despite a lot can be done to secure this information in its early and mid life cycle stages (e.g. password protection, encryption), it is interesting to examine if there are enough provisions in place for protecting such information at the end of its lifecycle.

Such measures can be of significant importance in the light of the growing popularity of e-commerce sites that facilitate the exchange of data storage, e.g. used disks and flash memories. Electronic marketplaces or on-line auction sites provide the means for a profitable disposal of unwanted memory capacities. By analysing the disk drive of a personal computer (PC) acquired in such a manner, it is possible to identify personal and sensitive data that could be used for a potential criminal activity. The personal data that is left behind on a PC's hard drive seems to pose a very realistic risk of identity fraud.

The purpose of this paper is to illustrate the level of risk associated with the disposal of media that could contain personal data and to bring to a wider audience the discussion on the necessary provisions that would need to be in place for its successful mitigation.

2 Second-hand Hard Disk Acquisition and Recovery of Personal Data Remaining on Disposed of Drives

One of the most common ways of acquiring storage that possibly contains personal identifiable information is via the purchase of second hand equipment. Especially if

this happens via an on-line facility (e.g. via an Internet e-auction site), as it provides anonymity and introduces complexity in the tracing of a potential fraudster. An article reported e.g., that a research group had purchased a hard disk for £5 through an auction site that contained a customer database as well as the current access codes to what was supposed to be a secure Intranet of a large European financial services group [3]. This fact illustrates how improperly recycled equipment that is intended to be reused, needs to be considered within a safe recycling strategy.

Another way for storage media acquisition is through conventional dumpster diving. Equipment that is not intended to be reused, often ends up unprocessed in bin bags at the doorsteps of individuals or organisations. Such disposal may occur for hard discs that are believed to be old or faulty, optical discs that may contain user backups etc. all of which may provide information useful to perform an ID fraud. Social engineering may be employed to acquire the disk as well; the fraudster may foil a scenario for the collection of unwanted discs for a supposedly charity or other fictional occasion.

In order to discuss on the level of perceived risk against the individual, we shall examine in this section the results from Jones et al. [4, 5] study on the forensic analysis of residual data, in relation to the method of examination used and the effort allocated to achieve those. In their study, Jones et al. analysed a large number of second hand discs. In terms of the methodology used, as far as the sample is concerned, those were purchased at computer auctions, computer fairs or on-line and were supplied 'blindly' to a research laboratory for analysis. The study looked at the residual data on the disk drives and determined whether there was any information on the disk that was easily recoverable and possibly might allow corporations or individuals to be identified.

The research involved the discs' forensic imaging, which was followed by an analysis aiming at determining what information remained and whether it could be easily recovered. Forensic imaging is the making of an exact duplicate of the entire hard drive and there are a number of proprietary tools that can perform this procedure. This copy of the hard drive or image is then used for analysis, ensuring that the original drive is not altered. The tools used to carry out the disc analysis included similar functions to the MS Windows Unformat and Undelete commands and to a hex editor, which can be used to view any information that existed in the unallocated portion of the disk. Besides commercial data recovery and forensic analysis suites, there are also freely available applications, such as the open source Sleuth Kit, as well as freeware toolsets such as the Windows Forensic Toolchest [6]. These types of tools are in essence available to anyone who could obtain the disks.

Fragkos et al. [7] have proposed an empirical methodology as a result of their work on Jones et al.'s study. Their proposed methodology is concerned with the identification and extraction of the most important repositories for information on a disk drive. The methodology starts with the exclusion of faulty disks then an automated procedure is used to check for wiped drives ie disk drives that do not contain any data. The next step is to locate all image files on the disk drive and to extract all the thumb.db files which are then used to extract the actual images. The thumb.db files store a backup/preview of image files even if that image file has been deleted or wiped. The next stage of the proposed methodology suggests that the directory names be extracted from index.dat file in order to access the cache of the

user's activity. The final stage is to extract the registry file of the user in order to identify information that may help to profile the user.

Of the disks analysed, over 40% contained a range of personal information including addresses, phone numbers, bank accounts and credit cards, other personal details, photographs, email and various on-line discussions.

3 Level of Exposure, Risk Mitigation and Challenges

Examples of the recovered data that could relate to acquisition of enough personal information to commit an identity fraud include the following [5]. From a major automotive company, there was payroll information, internal telephone contact details including mobile phone numbers, details of the internal network configuration and copies of invoices and orders. There were also emails between the company and its customers, meeting minutes and communications that were intended as written warnings to staff relating to poor performance. Other data from a disk recovered from an academic institute, included the names and web surfing habits of the users, the names of teachers, some confidential emails etc. There was also data that is thought to be test scores for individual students that was held in a database.

Similar data may well be used to commit an ID fraud; payroll records and personal contact information such as mobile phone numbers may provide enough data for a fraudster to attempt a request for a personal loan, on the financial details of an existing employee, under their own address. The availability of both employee and customer contacts may provide the grounds for social engineering attacks that will damage a company's reputation, if, for example, an attacker contacts customers using the credentials of real, existing employees.

An analysis of residual data on the grounds discussed in the previous section requires for the time being a significant amount of effort and expertise. However the tools required to perform such a task are freely accessible and know-how on their use is readily available over the Internet. And as on-line defences mature and become stronger (e.g. the move towards two-factor authentication systems driven by the banking sector), potential perpetrators of ID fraud may viably resort in off-line ways for committing an attack.

Users of personal computers do not necessarily have the knowledge, or the tools, available to forensically wipe their hard disk drives prior their disposal. Failure to adequately remove such data may result in the personal data being exploited for criminal gain in the form of identity fraud.

Therefore there is a clear requirement for the *education and training* of the relevant staff within organisations and of home users to inform them of the potential problems that arise from the failure to properly remove the information from disks and systems that are leaving their control. When organisations dispose of obsolete computers and hard disks, they must ensure that, whether they are handled by internal resources or through a third party contractor, adequate procedures are in place to destroy any data and also to check that the procedures that are in place are effective.

In this direction future releases of Apple's operating systems for example, will embody functions for secure erasure of the disposable medium. The recent AppleMac OS X operating system includes a file erasure facility that immediately overwrites the file with erroneous data, so that the file disappears and cannot be reconstructed.

Individuals and organisations should also consider the full encryption of hard disks so that if the disk is lost or the data is not effectively removed on disposal, it will not be easily recoverable. This would provide adequate protection in most situations. In this respect Microsoft planned the next generation of their operating systems (Windows Vista) to provide capabilities for full disk encryption. The new Windows Vista operating system has improved support for data protection at all levels with a full volume encryption of the system volume, including Windows system files and the hibernation file, which helps protect data from being compromised on a lost or stolen machine. The Encrypting File System has also been enhanced to allow storage of encryption keys on smart cards, providing better protection of encryption keys.

Both of these accommodations (disc erasure and encryption) are necessary to mitigate the level of anticipated risk, however they do not come without controversy. Law enforcement and the regulatory environment in countries where full encryption is not permitted in an unrestricted fashion (e.g. the French crypto legislation or the U.S. export controls) will need to reflect on these new technological developments.

4 Conclusions

Much of the focus on the protection from ID fraud has been given to its on-line form (e.g. [8]). However, there is another electronic-based form that fraudsters may employ, as more sophisticated on-line defences are used: the off-line route. In this paper we outlined the nature of personal data which may be found on personal computer hard drives that, once it has been uncovered may be used to commit an identity theft.

In the light of the growing popularity of electronic marketplaces and auction sites, the exchange of technology items that contain some form of storage capacity is booming. But such exchange could lead to exposure of personal information to potential fraudsters. Another off-line route may be the acquisition of improperly disposed of hardware. Indeed, there seems to be limited awareness in relation to the protection of information assets at the later stages of their life cycle.

The findings of relevant empirical studies [9, 10] demonstrate the feasibility of such off-line attacks and the methods and tools used provide an estimate of the effort required and the profile of a potential perpetrator. At the minute, considerable expertise and effort is required to perform such a task, however in the light of the availability of toolsets for forensic analysis and data recovery and the documentation of the research experiences and the knowledge in the public domain, this method poses as a significant emergent threat of the near future.

Therefore, there is a realistic risk associated with the incorrect recycling or disposal of personal information contained on computer hard disk drives. The recent

European Waste from Electrical and Electronic (WEEE) directive requires computer manufacturers to have recycling and refurbishment programmes in place [11]. There are many specialist waste disposal companies who will handle electronics and computer disposal on behalf of the manufacturers who need to comply with the EU directive. In the future, the recycling and disposal of computer hard disk drives should become less of a risk as this EU directive should go some way to preventing the current ad-hoc disposal of disk drives and thus will reduce the risk of theft of personal information from such drives.

Computer manufacturers and software vendors are incorporating enhanced security features in their products. The new Operating System countermeasures have been strategically positioned within the future releases of popular products, however it will be interesting to examine their impact in terms of effectiveness, user acceptance and the reaction of communities of interest such as law enforcement.

References

1. Tryfonas T., Thomas P., Owen P. (2006), "ID Theft: Fraudsters' techniques for Personal Data Collection, the Related digital Evidence and Investigation Issues", *Information Systems Control Journal,* (JOnline) Vol. 1.
2. Federal Trade Commission (2003), Fair and Accurate Credit Transactions Act of 2003 Revision, www.ftc.gov/os/2004/10/041029idtheftdefsfrsm.pdf
3. Leyden, J., "Oops! Firm accidentally eBays customer database", The Register, 7 June 2004.
4. Jones A., Mee V., Meyler C., Gooch J., "Analysis of Data Recovered from Computer Disks released for Resale by Organisations", *Journal of Information Warfare,* 2005.
5. Jones A., Valli C., Sutherland I., Thomas P. (2006), "An Analysis of Information Remaining on Disks offered for sale on the second hand market", *Journal of Digital Security, Forensics & Law,* Vol. 1 No 3.
6. Windows Forensic Toolchest, freeware tool available for download at http://www.foolmoon.net/security/wft/ (last accessed January 2007).
7. Fragkos, G., et al. (2006), "An empirical methodology derived from the analysis of information remaining on second hand hard disks", in Blyth, A., Sutherland, I., *WDFIA 2006. Proceedings of the First Workshop in Digital Forensics and Incident Analysis.*
8. Marshall, A.M. and Tompsett, B.C. (2005), "Identity theft in an online world", *Computer Law & Security Report,* Vol. 2005, Issue 2, Page 128 – 137.
9. Garfinkel, S.L., Shelat A., "Remembrance of Data Passed: A Study of Disk Sanitization Practices", *IEEE Security & Privacy,* Vol. 1, No 1, 2003.
10. Valli, C. (2004), "Throwing out the Enterprise with the Hard Disk", In *Proceedings of 2nd Australian Computer, Information and Network Forensics Conference,* We-BCentre.COM, Fremantle Western Australia.
11. Department for Trade and Industry, on-line (latest accessed 20-01-2007) http://www.dti.gov.uk/innovation/sustainability/weee/page30269.html.