The 2006 Analysis of Information Remaining on Disks Offered for Sale on the Second Hand Market

Andy Jones

Security Research Center, British Telecommunications and Edith Cowan University andrew.28.jones@bt.com Phone: +44 1473 646133

Fax: +44 1473 644385

Craig ValliEdith Cowan University

Iain Sutherland University of Glamorgan

Paula ThomasUniversity of Glamorgan

ABSTRACT

All organisations, whether in the public or private sector, use computers for the storage and processing of information relating to their business or services, their employees and their customers. A large proportion of families and individuals in their homes now also use personal computers and, both intentionally and inadvertently, often store on those computers personal information. It is clear that most organisations and individuals continue to be unaware of the information that may be stored on the hard disks that the computers contain, and have not considered what may happen to the information after the disposal of the equipment.

In 2005, joint research was carried out by the University of Glamorgan in Wales and Edith Cowan University in Australia to determine whether second hand computer disks that were purchased from a number of sources still contained any information or whether the information had been effectively erased. The research revealed that, for the majority of the disks that were examined, the information had not been effectively removed and as a result, both organisations and individuals were potentially exposed to a range of potential crimes. It is worthy of note that in the disposal of this equipment, the organisations involved had failed to meet their statutory, regulatory and legal obligations.

This paper describes a second research project that was carried out in 2006 which repeated the research carried out the previous year and also extended the scope of the research to include additional countries. The methodology used was the same as that in the previous year and the disks that were used for the research were again supplied blind by a third party. The research involved the forensic imaging of the disks which was followed by an analysis of the disks to determine what information remained and whether it could be easily recovered using publicly available tools and techniques.

Keywords: Computer forensics, disk analysis, data recovery, data disposal.

1. INTRODUCTION

In January 2005, a joint report (Jones, et al., 2005) was published by the University of Glamorgan in Wales and Edith Cowan University in Australia detailing the investigation of a number of second hand hard disks revealed that a large number of the disks examined still contained significant volumes of sensitive information. The research had been undertaken to gain an understanding of the level and type of information that resided on disks that had been offered for sale on the second hand market.

There was a limited amount of evidence from one report (Garfinkel and Shelat, 2003) and a number of commercially sponsored investigations and newspaper reports that there was an ongoing problem with regard to residual data on disks, however, before the 2005 report, there did not appear to have been any significant prior research into the subject.

The report identified that the great majority of the disks that were examined did still contain significant volumes of information that was considered to be sensitive, whether they had originally been used by businesses or in the homes of individuals.

This report contains the initial results of a much wider research project that was again carried out by the University of Glamorgan in Wales and Edith Cowan University in Australia. This project was sponsored by British Telecommunications (BT) and Life Cycle Services (LCS) and was undertaken during 2006 in order to determine whether there had been any change in the situation with regard to the level of information that remained on second hand hard disks in the intervening period and also to gain an understanding of how the results that were obtained compared the situation that was observed in results that were obtained from a number of other countries.

The results for the two sets of research, which were both conducted under the same conditions with tools that had similar capabilities, were then compared and a number of conclusions and recommendations on ways in which the situation could be improved were made.

2. THE RESEARCH

In order to gain realistic results from the research, the number of disks that were purchased in the UK and Australia were the same as or more than in the previous year. Nearly two hundred disks were obtained from the UK and approximately fifty in Australia. In addition, a scientifically sound sample of disks (in the region of 30) was obtained from both North America and Germany. All of the disks obtained were purchased at computer auctions, computer fairs or through eBay in the respective geographic areas. The disks were purchased discretely, either individually or in small batches, to minimise the possibility of the sellers becoming aware of the end use of the disks.

As with the previous research, the disks were supplied 'blind' to the researchers, with the only identifier on the disks being a sequential serial number. That is to say the disks were supplied to the researchers with no indication of where or how they had been obtained. The research was undertaken replicating the earlier research methodology, with each disk being forensically imaged and then placed into secure storage. All further research was carried out on the image of the original disk. While this may not be considered significant, it was thought to be an essential practice, as it allowed all of the research to be carried out in a non intrusive manner that did not affect or change the data that had been collected. This allowed the potential for other investigations to take place in the event of information being discovered during the course of the research that might indicate criminal activity or breaches of security. In the 2006 research this was found to have been an essential precautionary step, as two of the disks were found to contain material that necessitated them being passed to the police for further investigation.

The tools used to carry out the disk analysis performed similar functions to the Windows Unformat and Undelete commands and a hex editor (which can be used to view any information that existed in the unallocated portion of the disk) and the freely available tool Autopsy in the Linux based Helix Version 1.7 disk. These types of tool are available to anyone who could obtain the disks.

The first objective of the research was to determine whether there was any information on the disk that was easily visible or recoverable with the tools identified above. The second stage of the research was to look for specific elements of information that would allow for the identification of the organisation or individual that had used the disk and, if possible, further information such as the usernames, email addresses or documents, spreadsheets and databases. The purpose of this phase of the research was to determine the proportion of the disks that could be traced to an organisation or an individual.

The research is currently on-going; however, the initial results indicate that in both the UK and Australia little has changed in the level of information that remains on disks that have been released by organisations.

3. SUMMARY OF RESULTS FROM 2005 RESEARCH

In summary, the 2005 research reported that, of the ninety-two disks that were obtained in the UK that could be accessed, the following were revealed:

- 17 percent of the disks were totally blank and contained no file structure or data. Information that was made available after the analysis revealed that of these, 12 of the 16 disks that contained no data had been seeded by the recycling company (Life Cycle Services) that had sponsored the purchase of the disks for the research, to test its own procedures. This meant that only 4 (4%) of the 92 disks that were obtained on the second hand market were totally blank.
- In 48 percent of the disks that were examined, some attempt had been made to remove data. These efforts ranged from the simple deletion of files, to a formatting of the disk, to a fresh installation of an operating system over the existing operating system.
- 57 percent of the disks examined contained sufficient information to enable the organisation from which they had originated to be identified.
- 51 percent of the disks contained details from which individuals could be identified.
- 20 percent of the disks contained financial information relating to the organisations, including staff salary details, sales receipts and profit and loss reports.
- 8 percent of the disks contained details of the network infrastructure of the organisations from which they had originated, including server names, proxy numbers and other IP numbers.
- 4 percent of the disks contained potentially illicit (pornographic) data. In the form of either illicit photographs or references to sites that appeared to contain illicit material.

The results of research conducted in Australia in 2004 and 2005 (Valli 2004; Jenkins 2005) showed a situation very similar to that which was revealed by the research in the UK. Of the 23 disks examined in the Australian research, information on 21 of the disks was easily recoverable, one hard disk had been erased and one hard disk had mechanical failure. It became apparent that some of the disks had come from systems belonging to critical infrastructure providers in the form of power generation, water and telecommunications utilities; a particular concern as the files systems were intact and required a simple power on in compatible hardware.

4. RESULTS OF 2006 RESEARCH

The 2006 research was undertaken to determine whether the situation had changed in the intervening period and whether organisations and individuals had taken steps to address the issues that were revealed during the earlier report.

The initial 2006 research has revealed that, while there have been some changes, the results, as in the previous research, show that a significant number of disks still contain information that could be considered sensitive. This would appear to highlight two separate points. The first is that the situation with regard to the disposal of hard disks in the countries that had not been previously been surveyed is producing results that are very similar to those that have been observed in the UK and Australia. The second is that, despite an increasing awareness as a result of publicity from a number of information losses and incidents of identity theft and the results of a number of surveys (Synovate 2003; Price Waterhouse 2006; Johannes 2006) and increasing level of regulation, organisations are not modifying their procedures to ensure that information is effectively removed before the disposal of the computer hardware.

The initial results from the disks analysed to date in the 2006 research have revealed a small increase in the number of disks that have been effectively wiped, but that 44 percent of the disks have recoverable data in the slack space and 24 percent of the disks still have file structures present from which the data could be easily recovered.

Detailed below are the initial results of the 2006 research, reported by country. For the disks analysed in the UK:

- Of 200 disks obtained, 82 (41%) were unreadable and failed to spin up.
- 37 (31% of the readable disks) were totally blank and no data could be recovered from them.
- 25 (21% of the readable disks) contained sufficient information for the organisation that they had come from to be identified.
- 24 (20% of the readable disks) contained sufficient information for individuals to be identified.
- 18 (15% of the readable disks) contained information on personnel and of a personal nature.
- 12 (10% of the readable disks) contained financial information on the organisation or individual from which they had originated, including:

- An Automotive Company
- A Child Day care centre
- Local Government
- A Financial services organisation
- A retail store
- 1 (1% of the readable disks) contained information of a paedophile nature and has been reported to the police.
- 12 (10% of the readable disks) contained information that might be considered as illicit.

For the 24 disks that were obtained from North America:

- 12 (50% of the disks) were physically damaged and could not be accessed
- 1 (8% of the readable disks) had been wiped and contained no data.
- Of the remaining 11 readable disks,
 - 5 (42% of the readable disks)had come from commercial organisations and
 - 6 (50% of the readable disks) appeared to be from individuals.
- 3 (25% of the readable disks) contained illicit material.

For the 40 disks that were obtained from Germany:

- 29 (72% of the disks) were not in working order and could not be accessed.
- 5 (45% of the readable disks) had been wiped and contained no data.
- Of the remaining 6 (55% of the readable disks), 2 had come from commercial organisations and 2 appeared to be from individuals. The source of the other 2 disks has not yet been determined.

For the 53 disks that were obtained from Australia:

- 3 (6%) of the disks were physically damaged and could not be accessed
- 18 (36% of the disks) had been wiped and contained no data.
- Of the remaining 32 (60%),
 - 14 (43%) could be identified as having come from commercial organisations and

- 9 (28%) appeared to be from individuals.
- 2 (6%)of the disks contained illicit material

The table below shows a comparison of the results of the 2005 and the 2006 disk surveys.

	2005	2006
Total Number of Disks UK and Australia	116	253
Faulty/Unreadable	13 (13 %)	90 (36 %)
Wiped	17 (16 % ¹)	73 (45 % ¹)
Commercial Data Present	$60 (70 \%^2)$	$42 (47 \%^2)$
Individual data present	51 (59 % ²)	$44 (49 \%^2)$

¹ Percentages are of the readable disks

Table 1: Comparison of results of all of the disks in the 2005 and 2006 surveys

These raw results appear to indicate that there has been a significant change between the results obtained during 2005 and 2006, with a significant increase in the proportion of disks that were faulty and also a increase in the number of disks that had been wiped effectively. This has had the effect of reducing the proportion of disks that contain either commercial or individual data. Table 2 shows a comparison of the results from the disks obtained in the different regions.

	UK	Australia	Germany	North America
Total Number of Disks	200	53	40	24
Analysed				
Faulty/Unreadable	87 (43%)	3 (6%)	30 (75%)	12 (50%)
Wiped	55 (49% ¹)	18 (36% ¹)	4 (40%1)	1 (8% ¹)
Commercial data present	28 (35% ²)	14 (44% ²)	4 (66% ²)	$6(50\%^2)$
Individual data present	35 (44% ²)	$9(28\%^2)$	$3(50\%^2)$	7 (58 %²)

¹ Percentages are of the readable disks

Table 2: A comparison of the results from the disks obtained in the different regions

The results from the individual countries reveal what appear to be two significant disparities. The first is that the number of disks that were unreadable varied from 6 percent in Australia to 75 percent in Germany. The number of disks that were obtained in these two countries was not significantly different and there is no obvious reason for the large difference in findings. The second is that in the UK in 2006, the percentage of disks that had been wiped had increased to 49 percent from 16 percent in 2005, but the first results from the USA in 2006, with only 8 percent of the disks wiped are significantly different.

² Percentage of readable disks that had not been wiped

² Percentage of readable disks that had not been wiped

These are the first two sets of results from the UK and Australia that are available and the first from the USA and Germany and because of this, they may not be fully representative of the overall situation. It will be necessary to monitor the situation over a period of at least three more years to determine whether there are any significant patterns developing.

The type of material that was recovered from the disks that originated in commercial and academic establishments included:

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. In addition 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.

From a North American ship building company, there was information on the names of some employees, photos of employees, photos of the company site, and of ships (including the name of one vessel) and company emails, referring to ship re-fits and visiting guests. The recovered information also included details of the visit of a senior Local Government official and his wife regarding the re-dedication of a Merchant Ship and recovered emails included bids for contracts for Naval work which required TOP SECRET security clearances to work on the contract. There was also personal information recovered that would have been embarrassing for the individual to which it referred and a quantity of pornographic material that may have created the potential for blackmail.

From a British construction company, the information that was recovered contained internal memos, user names and telephone numbers. Emails were present and some personal information such as a Curriculum Vitae (résumé). In addition, information was recovered on the network infrastructure which contained some IP addresses and network information referring to the company data centre.

Data from a disk recovered from an academic institute, a community college, included the names and web surfing habits of the users, the names of teachers, some confidential emails and possible information from the school SIMS system. There was also data that is thought to be test scores for individual children that was held in a database.

Data from a disk that appears to have originated in a local government housing office made it possible to identify the main user, and contained information on complaints relating to the housing office, lists of repairs, contacts, lists of information for different sections of the council in addition to email. It also included personal information for the main user including a Curriculum Vitae, a wedding invitation, and a surprise birthday

party for the user's partner.

The type of material that was recovered from the disks that originated in commercial and academic establishments included:

Illicit information including downloaded audio and video files. A number of cases of pornographic material were encountered and also one case of paedophile material (this has been referred to the police for further investigation). There were also cases of family financial details including bank accounts and credit cards, details of personal habits and email and ICQ discussions. Additionally, there were also details of family names, addresses and photos, details of a family's holiday itinerary and information on a music festival. In one case there was evidence that one of the users had subscribed to online newsgroups such as alt.binaries.erotica.teen.female.newsgroups.

In another case there were details of a family (a mother, father and 2 daughters) that ran a restaurant. The information recovered included menus for the restaurant and staff names. The information recovered also included the daughters' school work and numerous photos of friends and boyfriends (?) and a number of emails relating to a Canadian holiday and the visit of the Canadian friends to the UK.

The disposal of disks that have not been wiped effectively and that may still contain significant information is not a new issue. As early as 1993, an article in the Canadian Globe and Mail (Canadian Globe and Mail 1993) reported that a second hand hard disk that had been sold to an Edmonton man that contained confidential personnel files on 166 employees of an Alberta organisation. A paper presented at a 1996 conference (Gutmann 1996) discussed the subject of the secure deletion of data. A report (Cullen 2000) appeared in 'The Register', regarding the disposal by the Morgan Grenfell Asset Management merchant bank of a disk that contained details of Sir Paul McCartney's bank account. A second paper by (Gutmann 2001) examined Data Remanence in Semiconductor Devices and then another paper (Garfinkel and Shelat 2003) examined ways in which disks could be 'sanitized'. In 2004, an article (Levden 2004) published in the Register reported that a mobile security group called Pointsec Mobile Technologies had purchased a hard disk for £5 through eBay that contained a customer database and the current access codes to what was supposed to be a secure Intranet of a large European financial services group. Finally in 2005, there were a number of reports including:

One in May (TechWeb News 2005) indicating that 70% of second hand hard disks still contained data and one in September (BBC News), which reported the information found by Disklabs on 100 disks that they had obtained.

Given the level of exposure that the subject has received over a considerable

period, the availability of suitable tools¹ to ensure the safe disposal of information, increasing legislative pressure and the increasing literacy of computer users over the period, there is now little excuse in the public and business sectors for ensuring that disks are effectively cleaned before disposal.

In the case of disks that have come from the home environment it is, perhaps, easier to understand as home users will not necessarily have the knowledge or tools available to clean a disk before disposal. On the other hand, with the increasing publicity with regard to identity theft and fraud, it would seem to be reasonable to expect that they would be very cautious and take significant steps to ensure that their personal information was removed before they disposed of the disks.

The implications of these failures to adequately remove the data are that, as a result of a breakdown in their corporate or individual security or inadequate management of the procedures for the disposal of equipment, information continues to be available to people who may seek to exploit it.

5. CONCLUSIONS

While there has clearly been an improvement since the previous research in 2005 in the proportion of disks that have been effectively wiped (up from 17% in 2005 to 31% in 2006 in the UK and up from 4% in 2005 (1 in 23) to 36% in 2006 in Australia) this still represents less than half of the disks that could be read. The increased number of disks that had been effectively wiped has reduced the proportion of the disks that contain sensitive organisational or personal information, but this type of information was still found on more than one in four of the disks that could be accessed.

Perhaps the most significant change was in the proportion of disks that were purchased that could not be accessed, with a huge 36 percent of the disks being faulty. Whilst it is accepted that the purchase of second hand disks is normally on an 'as seen' basis, this high a level of faulty disks seems inordinately high and cannot be attributed to damage in transit or any other single cause. It was notable that the overwhelming majority of the disks obtained in Germany were faulty.

There remains 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 and to make them aware of the potential costs to their organisations and the individuals of this failure. When organisations dispose of surplus and 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.

_

¹ Such as the Blancco tool which is approved for use by the UK Government.

6. RECOMMENDATIONS

The type of measures that organisations and individuals could be taken to reduce the volume of information that is inadvertently released when equipment is disposed of includes:

- 1. A general public awareness campaign and the wide publication of information on the risks of data leakage and measures that can be taken to prevent it. This could be carried out by the Government, the media, industry or academia.
- 2. Organisations should carry out a risk assessment to determine the level of assurance that is required that all information has been removed from the disks.
- 3. Organisations should introduce procedures to ensure that computer systems and computer hard disks that are disposed of are properly cleaned and that all data is erased to an appropriate standard. This could include steps such as a process to validate the effectiveness of internal data removal procedures to an identified standard or checks to ensure that any third party contractor has cleaned the disks to an adequate standard.
- 4. An alternative step that organisations could be introduced is the physical destruction of the disks.
- 5. The information and communications technology industry could make available the tools and facilities to enable individuals to effectively remove the information from their computers before they are recycled.
- 6. Individuals and organisations should consider the full encryption of hard disks so that if the disk is lost or the data is not effectively removed, it will not be easily recoverable. This would provide adequate protection in most situations.

7. CONTRIBUTING ORGANISATIONS

British Telecommunications (BT). BT is one of the world's leading providers of communications solutions serving customers in Europe, the Americas and Asia Pacific. Its principal activities include networked IT services, local, national and international telecommunications services, and higher-value broadband and internet products and services. In the UK, BT serves more than 20 million business and residential customers with more than 30 million exchange lines, as well as providing network services to other licensed operators.

<u>Life Cycle Services (LCS).</u> LCS is the UK's leading specialists in the asset refurbishment, redeployment and secure disposal of information and communication technology.

Edith Cowan University (ECU). The Security and Intelligence research cluster of the School of Computing and Information Science at ECU conducts research into all aspects of Information Operations from the technological aspects of computers and network security to the 'softer' side involving issues such as perception management and information policy. At present, its research theme is 'deception'. The group has numerous doctoral, masters and honours candidates. Its main areas of interest are information operations, computer/network forensics, RFID security, mobile computing security, honeypots and the use of deception in security.

<u>University of Glamorgan (UoG)</u>. The Information Security Research Unit in the School of Computing at the UoG has a strong and well established theme in the areas of Network Security, computer forensics and the threats to information systems. The Unit is focused on the issues associated with the design and development of early warning systems that are capable of detecting and responding to a variety of cyber based attacks, and on the issues associated with computer forensic science. In particular, they are developing technologies such as secure XML, threat assessment methods, vulnerability management, IDS data integration, data mining, data fusion and secure wireless mobile computing. The unit has two specialised laboratories, the Network Security Laboratory and the Computer Forensics Laboratory.

8. ACKNOWLEDGEMENTS

In addition to the individuals named as authors for this paper, we would like to acknowledge the people who assisted in the disk imaging of more than 200 disks (a non trivial task). The people involved were:

Andrew Blyth, Theodore Tryfonas, Konstantinos Xynos, Grigorios Fragkos, Michael Pilgermann, Vivienne Mee, Stilianos Vidalis, Huw Read, Olga Angelpoulou, and Abdulrazaq Al-Murjan.

9. REFERENCES

- BBC News (2005), Data dangers dog hard drive sales, BBC, 12 September 2005
- Canadian Globe and Mail (1993), Disk Slipped Into Wrong Hands, Canadian Globe and Mail, 2nd August 1993.
- Cullen D. (2000), Paul McCartney account details leaked on second user PC, The Register, 9th February 2000.
- Garfinkel S.L, Shelat A. (2003), Remembrance of Data Passed: A Study of Disk Sanitization Practices. IEEE Security & Privacy, Vol. 1, No. 1, 2003.
- Gutmann, P. (1996), Secure Deletion of Data from Magnetic and Solid-State Memory, Sixth USENIX Security Symposium Proceedings, San Jose, California, July 22-25, 1996.

- Gutmann, P. (2001), Data Remanence in Semiconductor Devices, 10th USENIX Security Symposium, Washington, D.C., August 13-17, 2001.
- Jenkins, C. (2005), Govt data sent to auction. The Australian, 2nd August 2005.
- Johannes, R. (2006), The Demographics of Identity Fraud: Through education and vigilance, banks can prepare and protect those most vulnerable, Javelin Research,
 - http://www.javelinstrategy.com/uploads/607.R_2006_IDF_Demographics.pdf, Aug 2006.
- Jones, A., Mee, V., Meyler, C., and Gooch, J., (2005), Analysis of Data Recovered From Computer Disks released for sale by organisations, Journal of Information Warfare, (2005) 4 (2), 45-53.
- Leyden, J. (2004), Oops! Firm accidentally eBays customer database, The Register, 7 June 2004.
- Price Waterhouse Cooper (2006), DTI Information security breaches survey 2006, http://www.dti.gov.uk/industries/information_security Sept 2006.
- Synovate, (2003), Federal Trade Commission Identity Theft Survey Report, Federal Trade Commission, June 2006.
- TechWeb, (2005), Seven-In-Ten Second-hand Hard Drives Still Have Data, TechWeb News, 31 May 2005.
- Valli, C. (2004), Throwing out the Enterprise with the Hard Disk, In 2nd Australian Computer, Information and Network Forensics Conference, We-BCentre.COM, Fremantle Western Australia.

AUTHOR BIOGRAPHIES

Dr. Andy Jones is the Head of Security Technology Research a the Security Research at the Security Research Centre at British Telecommunications (BT) where he leads the research into the security of Virtual Intranets and the use of Trusted Computing concepts and tools and the development of a risk management methodology. In addition he holds a post as a visiting adjunct at Edith Cowan University in Australia, where significant research is being carried out into wireless networking, RFID vulnerabilities and computer forensics.

Dr. Craig Valli is a Senior Lecturer (Network and Computer Security) within the School of Computer and Information Science. He has 20 years experience in the IT Industry and consults to industry on network security issues. He is the Chair of the Australian Computer, Information and Network Forensics Conference and Co-Chair of the Australian Information Security Management Conference. Craig is also a Co-Editor of the Journal of Information Warfare. He has over 30 publications to his name on security related topics. His research interests include Network Security, Honeypots, Intrusion Detection Systems, Compute Clustering, Computer Forensics, Wireless and SCADA Security.

Dr. Iain Sutherland is a Senior Lecturer at the School of Computing in the University of Glamorgan. He has been involved in a variety of research projects in the area of information security including secure XML transactions, and reverse engineering metrics. Dr. Sutherland's main field of interest is computer forensics; he has acted as a consultant on commercial cases in the University's Forensics Computing Laboratory. In addition to being actively involved in research in this area and supervising a number of Ph.D. students, Dr. Sutherland teaches computer forensics at both undergraduate and postgraduate level.

Paula Thomas holds a BSc (Hons) in Computer Studies and is currently employed as a Senior Lecturer in the School of Computing at the University of Glamorgan. She is an active member of the Information Security Research Group (ISRG) and has a number of publications on network and wireless security. The ISRG has a fully equipped Computer Security and Forensic Analysis laboratory and conducts many projects for law enforcement agencies and industry. Paula manages the consultancy activities for the School of Computing and also participates in projects in the area of computer security and forensics.