Information Leakage through Second Hand USB Flash Drives

W.H.Chaerani and N.L.Clarke

Centre for Information Security and Network Research,
University of Plymouth, Plymouth, United Kingdom
e-mail: info@cscan.org

Abstract

USB flash drives are one of the common removable storages used nowadays. As the capacity increase and the price decrease, the content put on these devices become more valuable and vary than before. These USB keys are not only used to store personal files, but also sensitive information for example password list and CVs. Moreover, it is also not impossible for a person to keep information regarding other people. Additionally, since some company still allow their employees to use USB keys on internal network, it is also not impossible for corporate data to be transferred into these devices. This research try to see whether it is possible to obtain information, particularly sensitive ones from second hand USB keys sold on online auction sites. The keys used on this research were obtained randomly. Using digital forensic software, the keys were imaged and examined to obtain any remaining information that may exist.

Keywords

Information leakage, Second hand USB, Digital forensic, USB flash drive

1 Introduction

The United Kingdom has seen several big cases regarding information leakage due to the loss of storage media such as hard disk, laptop, or USB keys (Kennedy, 2008; Oates, 2008; BBC, 2008; BBC, 2009). The leakage itself was possible since most of those devices were not encrypted, making it possible for unauthorised person to easily access the information kept inside which could cause damage to the owner of the device.

Research regarding to information leakage due to storage media have been conducted annually by University of Glamorgan and University of Edith Cowen since 2005 (Jones, 2005; Jones, 2006; Univesity of Glamorgan, 2007; M2 Communications, Ltd, 2008; Jones, 2009). In all of their researches they could find that the majority of end users, particularly corporate ones, are still unaware of how to properly process their unused storage media so that the devices are clean upon leaving their control. Similar results were obtained from studies conducted by O&O Software (Kehrer, 2005; Kehrer, 2007). From their Data Data Everywhere studies, O&O Software could find that in addition to hard disks, which are the common disposed storage media from organizations, removable storage media such as memory cards and USB keys could also be the source of information leakage.

USB flash drives are one of removable storage media commonly used by people ranging from students to company directors. The device is very easy to use, practical and mobile. As the capacity increases and the price decreases, the value of information it kept become more and more sensitive. Obviously, the consequences might be severe if these keys are lost and fall to the wrong hand.

This research aims to see whether it is possible to obtain any sensitive information from second hand USB keys. Prior to selling their unused USB keys, it is only normal that people 'clean' them first. Some attempts people usually use to clean their USB keys include simply deleting the files, formatting the key, or using appropriate disk cleaning software. The problem is the first two attempts do not always clean the keys thoroughly. Deleting files only clear the path to reconstruct the files. The contents of the files itself still exist until it is overwritten by new files (Bunting, 2006). This remaining information may cause some sensitive information to be leaked without the owner's consent.

2 Methodology

In order to do the research, twenty second hand USB keys were bought randomly from online auction site and were examined forensically to extract any information left on them. The USB keys were picked randomly from different sellers with different range of capacity. In order to maintain the chain of custody, before any investigation can begin, all of the keys were imaged using digital forensic software. All of the investigations are then done on these images. The original keys are kept individually and then labelled along with the packages or envelopes that came with them. This is done so that the original keys will remain unchanged throughout the research. This also helps in maintaining the chain of custody, since the researcher does not know beforehand what kind of information will be found from the key. In case an evidence of a crime is found on the image, the investigation will be stopped and the key along with the packaging will be sent straight to the law enforcement agencies in its original condition.

This research utilised two digital forensic software and two additional file recovery software. The two digital forensic software are Encase 5 and Forensic Tool Kit 1.81.3. However, Encase was the mainly used because Forensic Tool Kit was a trial version and has limitation of number of files that can be investigated. The other two file recovery software was added to obtain more results. Recuva from Piriform was very useful in differentiating between recoverable and unrecoverable files and Zero Assumption Digital Image Recovery was very useful in recovering images. These software are easy to obtain and the use is quite straightforward, making it possible for novice users to utilize them in their own convenient time. Moreover, the last two software mentioned are available for free.

Once the keys were forensically imaged, the software was used to see whether any information was left on the key. Deleted folders were recovered and the unallocated clusters were checked to see remnants of the files. Each file were also analysed to see if it has any security measure applied such as password protection or encryption. Keyword search technique was also used to help in searching specific information. Encase provided templates for keyword search which include email addresses, IP

addresses, web addresses, phone numbers, dates with 4 digits year, and credit card numbers. Basic keywords added in the research were credential keyword such as user, username, uname, password, login, pass, and pwd. Additional keyword can be added in the search based on the findings on each key. For example, if the name of the owner can be found, then the name will be added in the keyword search to find any information related to the owner.

3 Result

A total of 36,136 files could be retrieved from the keys, dominated with documents such as Microsoft Words, spreadsheets, and PDFs. Overall, the information found on the key were ranging from individual to confidential corporate files. Individual information is the kind that relate to a person, either the owner of the key or other person whose data for some reason may reside in the key. This type of information is then divided into personal and private information. Personal information is the type of information that describes the person and the person does not bother to share them, such as names, addresses, and emails. Private information are the type of information that should be kept private and should not be used without the owner's consent, such as bank details, national insurance numbers, date of birth, mother's name, online credentials, etc. The corporate information found during the research includes confidential company reports, client's financial information, meeting notes, internal and external correspondences, etc. Apparently both individual and corporate information could be found on most of the keys.

<table>
<tr><th colspan="2">Category</th><th colspan="2">Number of Keys (out of 20)</th><th colspan="2">Percentage</th></tr>
<tr><td colspan="2">Empty</td><td colspan="2">4</td><td colspan="2">20%</td></tr>
<tr><td colspan="2">Corporate</td><td colspan="2">6</td><td colspan="2">30%</td></tr>
<tr><td rowspan="2">Individual</td><td>Personal</td><td rowspan="2">14</td><td>11</td><td rowspan="2">70%</td><td>55%</td></tr>
<tr><td>Private</td><td>14</td><td>70%</td></tr>
</table>

Table 1: Nature of information found on the keys

After examining the total of twenty keys, the following are found:

3.1 Blank keys

Out of twenty USB keys examined in the research, only four or 20% of the keys were totally blank. The unallocated clusters of these keys were mostly blank with no meaningful information can be obtained.

3.2 Identifiable to previous owners

65% of the keys examined, were identifiable to the previous owner. This means that at least the name of the owner is known. This information obtained from document signatures, CVs, invoices, and correspondences. Additional information found about the owner were ranging from email addresses, contact numbers, addresses, online credentials, personal pictures, bank details and even national insurance numbers. One

of the key even contained medical record and history of the owner. Another key identified to once belong to an insurance company in London and it contains enormous amount of corporate information. The other three keys were not blank and contain information but they were not identifiable to the previous owner.

3.3 Financial Information

15% of the keys examined contain bank details such as account number, sort code, bank name, and bank address. One of the key contains more than 30 bank details belong to companies. The financial information found did not only belong to individuals but also company. In one of the key, reports regarding profit and loss, balance sheet, customer list, invoices, and sales forecast could be recovered. Some of these documents were clearly labelled as 'confidential' or 'for internal use only'.

3.4 Identification Number

40% of the keys examined, contained identification numbers. These IDs can be in the form of passport numbers, national identification numbers, national insurance numbers, driver license number, or license plate number. This information was obtained from documents or scanned pictures. In one case, a document was found and it listed names, addresses, and passport numbers of students in one college.

3.5 Curriculum Vitae

15% of the keys examined contain curriculum vitae of the owner or of the owner's acquaintances. These CVs provide detail of a person starting from names, addresses, contact numbers, and date of births, to passport numbers and national identification numbers. Other information provided includes education and work experience detail. By gathering information from these CVs, impersonator could have enough information about the victim's profile.

3.6 Network Information

One key examined contain full details of the owner's network configuration. This information found from keyword search 'user' and it reveals the network user credentials, the wireless LAN key and phrase, the manual configuration of the network and also the router configuration. By gathering information from the user manual and executables found, the hardware details can also be deduced. Another key examined contain a text file containing a company's VLAN password.

3.7 Medical Information

One key examined contain full medical information of the owner. This information is in the form of medical strip which reveals basically everything one needs to know about the owner's health including what kind of allergy he had, blood transfusion type, main medication, daily medication he had to take, patient number, the doctor's name and contact number, and sickness history. The owner's full health history could also be deduced from the correspondences found on the key.

3.8 Information about other people

Over half of the keys examined contain information about other people. This information include name (or full name), picture, national insurance number, policy insurance number, online account credential, bank detail, and full curriculum vitae where you can get more detailed information such as date of birth, address, contact numbers, address, education history and work experience.

3.9 Corporate Information

20% of the keys examined contain corporate information. This information include company logo, letter or form templates, meeting notes, financial report, list of board of directors along with their contact addresses, contact numbers and date of births. There were also turnover analysis, sales forecasts, signature scans, and companies' bank details. One of the highlight of the research is the last key investigated, which consists of corporate data of an insurance company in London. Around 3000 still intact documents could be recovered from the key and 91% of those files contain sensitive corporate information, not only to the insurance company but also to its partners and clients. It is such a pity that out of these files, only one was password protected, leaving the rest easy to be accessed and analysed. Almost half of the sensitive documents found were explicitly labelled as 'private and confidential' or 'for internal use only'. The rest which were not labelled as 'confidential' also contained sensitive information to the company such as clients' financial reports, bank details, invoices, sales forecast, financial analysis, and board of directors' details. Other than corporate information, the key also contained a collection of personal folders of the staffs. From these folders, a bank detail and signature scan of one of the staff can be obtained. The staff also kept a confidential letter made by his spouse that contained the spouse's sensitive details such as national insurance number, bank details, and holiday dates.

4 Consequences

The richness of information obtained from just recovering files from second hand USB keys was alarming. Moreover, some of the information was simply obtained by recovering files using free file recovery software available online. These software are relatively easy to use and quite straightforward, making it possible for novice users to use and conduct file recovery on these keys. Obviously, the consequences will be severe for the owner (or other people that has their information stored on the keys) if the key ever fall to the wrong hand. Based on the information gathered during the research, it can be deduced that more than half of the keys could cause identity theft if it is loss or stolen. Obviously, this is not the only threat.

Threats	Identity Theft	Fraud	Industrial Espionage	Blackmail	Hacking / Network Intrusion	Robbery
Number of Keys	11	9	2	5	4	1

Table 2: Summary of possible threats to the keys

4.1 Identity theft

The information found on this research were enough to conduct an identity theft which can cost the victim not only money, but life. In extreme case, the owner who kept his medical records on the key may be the victim of medical identity theft, which means it is possible in the future when he need to get blood transfusion, he get injected by the wrong blood type.

4.2 Industrial Espionage

The corporate financial information found on this research was quite recent. Information such as turnover analysis, sales forecasts, and financial reports can be used by competitors in industrial espionage. Not only financial information, documents revealing a company investment strategy could be found and also could be used in industrial espionage which can cause loss in terms of money and reputation.

4.3 Fraud

From a key that contain the owner's request for a web domain, his signature scan could be recovered. This piece of information can be used to falsify document which can lead to fraud. Moreover, company logos, form and letter templates, insurance policy wordings gathered from the key that belong to a company would make a convincing fake document. Even information as simple as scan of license plate number can be used to falsify a stolen car's license plate which could lead to other fraudulent activities.

4.4 Hacking or Network Intrusion

Full network configuration which detailing the W-LAN key and phrase could cause the owner to suffer from network intrusion. Moreover, online credentials such as eBay and PayPal that could be found by recovering encrypted file can be used to access the service without the consent of the owner. Another case is the online credentials found for a company online service. The credential that belongs to one of the staff would make it possible for intruders to login and access the internal resources of the company.

4.5 Blackmail

Illicit material could be found from one of the key, mainly in the form of videos. This might be incriminating to the owner if such information is exposed. Blackmail can also be done to the company whose financial information was kept in one of the key examined.

4.6 Robbery

The financial information found throughout this research was sufficient enough for a criminal to commit theft from the bank account found. In addition, the information

about holiday dates also enables criminals to indicate when the victim will not be home.

5 Reason of Incidents and Preventions

Unfortunately, it seems that most people do not know that simply formatting or deleting files will not clean the devices completely, as the remnants of those files can still be detected and recovered using appropriate software. This lack of knowledge make it possible for someone to buy a second hand USB flash drive for less than £5 and obtain someone else's bank detail instead.

The finding of corporate USB in this research also indicates that there are still companies that do not aware the dangerous risk these tiny devices have. Based on BERR research, 67% of companies do nothing to prevent data from leaving their company premises through removable media such as USB flash drives (BERR, 2008). Another reason is simply ignorance. In accordance to lack of knowledge, people seem to think that it is impossible someone could recover files which are deleted long ago.

In order to prevent this kind of leakage from happening, several preventions can be done:

1. Encrypt all removable storage media
2. Protect devices from unauthorised access by utilising password protection or biometric authentication
3. Use specialised erasure software to clean all storage media that are no longer used.
4. Destroy physically all devices at the end of their lifetime.
5. Companies should have strict rules and guidelines regarding the usage of removable storage media. This rules need to be audited periodically to assess its effectiveness.
6. Government and/or academic communities should give education or public awareness regarding this issue.

6 Conclusion and Future Work

Based on the results found in this research, it can be concluded that it is possible to obtain sensitive information from second hand USB flash drives. It is inevitable that people keep sensitive information on their USB keys. However the problem is when the USB keys are no longer used. Proper erasing method needs to be done in order to erase the content completely and remove any remaining information.

To enlarge the scope of the research, the number of keys should be increased. Thus, more conclusions can be extracted and more variables can be measured. Such as, does the capacity of the key relate to the value of information it keeps, or is there any relation between the origin countries of the keys with the richness of information found.

7 References

BBC, 2008. *More secret files found on train*. [Online] Available at: http://news.bbc.co.uk/1/hi/uk/7455084.stm [Accessed 13 August 2009].

BBC, 2009. *Previous cases of missing data*. [Online] Available at: http://news.bbc.co.uk/1/hi/uk/7449927.stm [Accessed 13 August 2009].

BERR, 2008. *BERR Information Security Breaches Survey*. [Online] Available at: http://www.security-survey.gov.uk [Accessed 15 November 2008].

Bunting, S., 2006. *The Official EnCe: Encase Certified Examiner*. [Online] Available at: http://i.techrepublic.com.com/downloads/home/0782144352_chapter_1.pdf [Accessed 9 April 2009].

Jones, A., 2005. How much information do organizations throw away? *Computer Fraud & Security*, 2005(3), pp.4-9. DOI: 10.1016/S1361-3723(05)70170-6 [Accessed 31 July 2009].

Jones, A., 2006. Cradle to grave - security failure to the very end. *Computer Fraud & Security*, 2006(9), pp.4-8. DOI: 10.1016/S1361-3723(06)70418-3 [Accessed 13 August 2009].

Jones, A., 2009. Lessons not learned on data disposal. *Digital Investigation*, pp.1-5. DOI: 10.1016/j.diin.2009.06.017 [Accessed 15 July 2009].

Kehrer, O., 2005. *Data Data Everywhere 2005*. [Online] O&O Software Available at: http://www.oo-software.com/en/study/study_ddd2005_en.pdf [Accessed 29 May 2009].

Kehrer, O., 2007. *Data Data Everywhere*. [Online] Available at: http://www.oo-software.com/en/docs/ddd2007_en.pdf [Accessed 3 December 2008].

Kennedy, J., 2008. *Another data security breach reported at Bank of Ireland*. [Online] Available at: http://www.siliconrepublic.com/news/article/11718/cio/another-data-security-breach-reported-at-bank-of-ireland [Accessed 13 December 2008].

M2 Communications, Ltd, 2008. *BT: Research Reveals At Least One In Five Second-Hand Mobile Devices Still Contain Sensitive Information; Today's Sophisticated Devices Exacerbating The Problem Of Keeping Sensitive Information Safe*. [Online] Available at: http://www.tmcnet.com/usubmit/2008/09/25/3670481.htm [Accessed 6 February 2009].

Oates, J., 2008. *Million bank details sold on eBay and a few more gone AWOL*. [Online] Available at: http://www.theregister.co.uk/2008/08/26/more_details_lost/ [Accessed 14 July 2009].

Univesity of Glamorgan, 2007. *Discarded Hard Disk Hold Sensitive Data*. [Online] Available at: http://news.glam.ac.uk/news/en/2007/sep/19/discarded-hard-disk-hold-sensitive-data/ [Accessed 6 February 2009].