Hard Disk Storage: Firmware Manipulation and Forensic Impact and Current Best Practice

Gareth Davies, Iain Sutherland
Faculty of Advanced Technology
University of Glamorgan
CF37 1DL
+44(0)1443 480480
{gddavies, isutherl @glam.ac.uk}

ABSTRACT

The most common form of storage media utilized in both commercial and domestic systems is the hard disk drive, consequently these devices feature heavily in digital investigations. Hard disk drives are a collection of complex components. These components include hardware and firmware elements that are essential for the effective operation of the drive. There are now a number of devices available, intended for data recovery, which can be used to manipulate the firmware components contained within the drive. It has been previously shown that it is possible to alter firmware for malicious purposes, either to conceal information or to prevent the drive's correct operation. We review the general construction of a hard disk drive. In particular we examine the error handling process present within hard disk drives for dealing with failed or failing sectors and detail how this can be manipulated. The potential forensic impact on an investigation of manipulating firmware is then explored. We propose best practice considerations when analyzing a hard drive where firmware manipulation is suspected and detail a possible method to detect this form of modification.

Keywords: Hard Disk, Steganography, Data Recovery, Firmware.

INTRODUCTION

The hard disk drive remains one of the most common storage devices and therefore commonly features in digital investigations. There are numerous papers discussing best evidential practice and a substantial number of procedures, including the Association of Chief Police Officers Guidelines (ACPO 2008) in the UK and the Department of Justice, Prosecuting Computer Crimes guidelines in the USA (DoJ 2007). These guidelines clearly consider hard drive media and define best practice processes and procedures for the collection and general analysis of digital evidence. However, in specific cases where a technically competent suspect has access to particular data recovery hardware and software, there is the potential for the various hard disk drive firmware implementations to be manipulated for malicious purposes. This can allow the user to have the capacity to conceal information on the drive and place this data beyond forensic recovery using standard tools and techniques. There is also the potential for the drive to be sabotaged by these tools and by possible future forms of malware, prohibiting any form of forensic analysis. Therefore, there is a need for an investigator to understand some of the processes that can be undertaken to recover data from a damaged disk drive and also the potential for these techniques to be misused allowing the concealment of potential evidence. This enables the investigator to comprehend the forensic significance / impact of data recovery techniques.

HARD DISK DRIVE FUNCTIONALITY

A hard disk drive is a complex device composed of platters, voice coils, read / write heads, casing, mountings, a motor and a printed circuit controller board. These are manufactured in a number of form factors, the most common being the 3.5 inch and 2.5 inch disks found in desktop and laptop systems respectively. The data storage area is composed of a stack of metal, ceramic or glass platters coated with a magnetic film. Each disk surface has a separate armature and head assembly. One rotation of the disk at a particular radius is known as a track. For sets of surfaces, a set of tracks at the same radius is known as a cylinder. The sector is the smallest addressable unit, typically containing 512 bytes of data. A specific sector address can be found using the cylinder address (C) the Head (H) and the Sector (S). At a higher level of abstraction the Logical Block Address (LBA) method assigns a sequential number to each sector.

The main hard disk drive manufacturers now offer drives with a maximum capacity of around 2TB. Once the drive has been formatted and contains a file system, the capacity is somewhat reduced. Not all areas of the disk are addressable by the host computers operating system as shown in Figure One. In addition to the user addressable space, there are areas of the drive that are used for the manufacturer to record data. These include the Host Protected Area (HPA) used for holding diagnostics and other utilities required by the PC manufacturer (Gupta et al 2006) and the Device Configuration Overlay (DCO), either or both of which can exist on a hard disk. The Device Configuration Overlay (DCO) is similar to the HPA, but is used by manufacturers to configure drive sizes and may exist at the same time. An excellent overview of the HPA and DCO are provided in Carrier (2005).

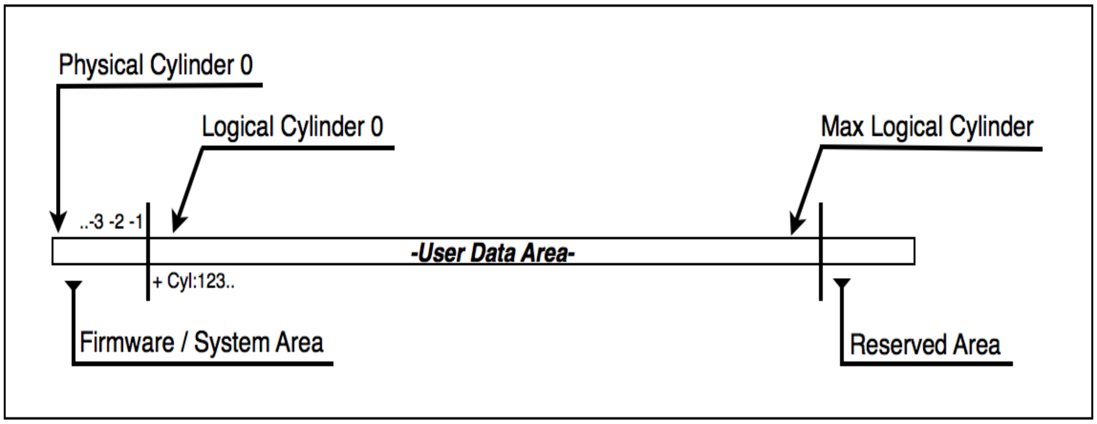

Figure One: Overview of Disk Data Storage Areas

The firmware area / system area of the drive is not accessible during the normal operation of the drive and subsequently is not addressable by the average user or the operating system. An initial portion of the drive firmware is present on the PCB controller board. This is then responsible for loading the platter resident firmware / system area which facilitates full operation. Disk firmware controls all aspects of the internal hard drive operation. The firmware controls the disk startup / self-check sequence when the system is powered on, placing the drive in a ready state that allows the host computer to load an operating system. During operation the firmware ensures the correct operation of the hard drive, allowing it to correctly interact with other components on the system (e.g. the operating system).

FIRMWARE OPERATIONS

The firmware in the majority of drives consists of a series of modules; P-list, G-list, SMART Attributes and U-List (Firmware Zone Translator). Each of these performs a key function. An example function is defect control, no disk is manufactured without flaws and there will be some sectors on the drive that cannot be used. At the time of production these flaws are recorded in the disk firmware as the ‘P’ (permanent / primary / production) list. As the disk ages and through wear & tear other sectors may fail; this is recorded in the ‘G’ (growth) list. Reads and writes are automatically redirected (remapped) to spare sectors, see Figure Two below.

P-list and G-list sectors are automatically bypassed by the drive electronics and so do not slow down drive sector access times. By adding or removing a sector from the P-list and/or G-list, we have the ability to hide/make-visible data on the hard-drive.

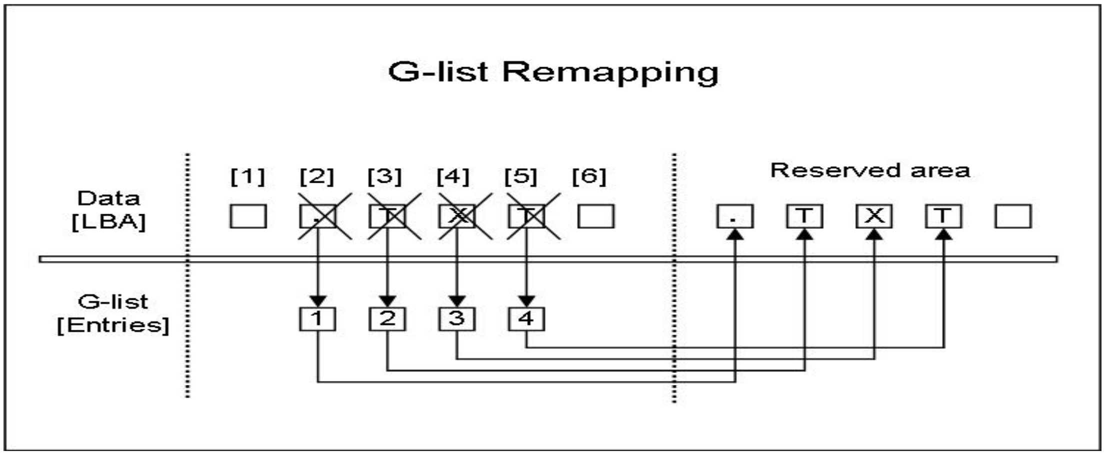

Figure Two: G-list Remapping

This process is transparently handled by the disk and occurs ‘beneath’ the operating system via the two lists, P-list and G-list (Blyth et al 2008). The firmware in these disks may fail. The G-list may become full on some disk models and as a result the disk may stop working. An error in the firmware can prevent the disk being accessed while still physically healthy and all user data remaining intact.

FORENSIC IMPACT

In a previous paper (Sutherland et al 2009) we have examined the possibility of steganography and data hiding via the manipulation of disk firmware. This paper addresses the issues of the possible forensics impact of these techniques on forensic practice and procedure. The possibility of attacking systems via firmware distribution and the use of microcode exploits has also been discussed by Zhou et al (2009).

To date there are only a limited number tools available tools to perform repair or modifications on firmware. There are a number of free / shareware tools that claim to read some portions of the firmware, usually disk model / serial number (Browsedata 2004). But these tools do not facilitate sufficient control over the firmware to make repairs or exploits possible. In terms of commercial products the authors are aware of two systems available for this type of analysis and repair. Both systems comprise a combination of hardware and software tools. One particularly sophisticated tool originates in Russia and costs in the region of $4000. The full Russian tool suite includes the ability to extract data and work with some solid-state devices and SCSI disks is in the region of $15,000. A more readily available device is offered from China and can be obtained via resellers in Europe for around $350 per disk manufacturer. Either of these tools would enable a competent user to manipulate firmware to conceal data or code from the hard disk dive itself.

There are a number of possible scenarios where this technology could be misused. An individual may use disk firmware steganography to conceal information within the drive, either by using the firmware defect control system or by the manipulation of bad sectors (Blyth et al 2008, Sutherland et al 2009). Another possibility is the use of malware on the drive to prevent the disk ever operating correctly again by attacking unique critical elements of the firmware, denying a user or forensic investigator access to any data. This kind of exploit would be developed and targeted at particular disks and systems and would act as a sophisticated method of sabotage that could render the drive contents irrecoverable.

Firmware manipulation can have a significant impact on the forensic process. Data that has been hidden using firmware steganography techniques will not appear for analysis in a traditional forensic image. In the event of malware targeting and corrupting the firmware, this will prohibit any forensic images to be acquired from the hard disk drive. As previously outlined, this type of manipulation is not detected by the current suite of forensics tools and so requires both specialised tools and training to detect. However, an investigator having been trained in this area and in possession of the correct equipment is still faced with a number of problems.

In the case of determining if the disk firmware has been tampered with, either to conceal information or as a result of malware targeting and corrupting the firmware, the investigator would need to be able to assess the validity of the firmware. This process has a number of challenges: each family of drives is unique so the investigator would need to evaluate the drive against a comparable disk and perhaps use key firmware modules, even hardware components, from the donor drive to verify the firmware is valid or to enable the repair of the original drive to a fully functional state to allow forensic analysis. The ability to do this work would require a substantial library of firmware / donor drives for comparison and replacement of failed mechanical parts. This library would be difficult to build, as often donor drives are difficult to obtain and match due to strict compatibility criteria, which changes dramatically by manufacturer.

In a previous paper (Sutherland et al 2009), the authors examined and highlighted the possibility to manipulate the firmware, via the defect control system i.e. the error lists; the P-list and G-list. If a malicious attempt has been made to manipulate the error lists then an LBA has been remapped to a reserve location on the drive and subsequently altered. This presents a situation where the original ‘bad’ sector contains one piece of information and the remapped sector may contain different information, this maybe due to time factors i.e. the original file copied has been overwritten and replaced with another file, or, direct editing of the information contained at those remapped LBA addresses has been done purposely. This can be achieved via a hex editor by zeroing out the hex values that correspond to the LBA number or saving an alternate file to that LBA location, overwriting the copy of the original data (see Figure Three below).

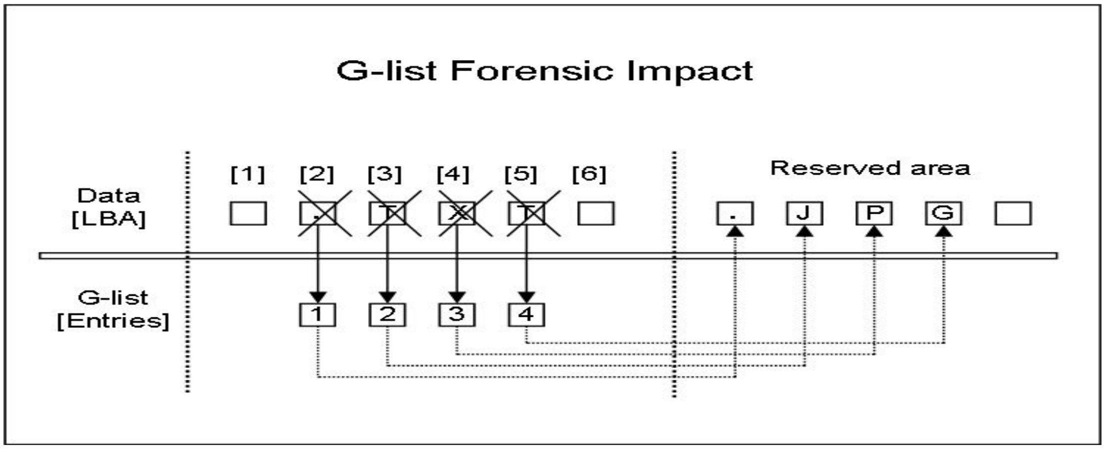

Figure Three: G-list Forensic Impact

In this situation the user and most forensic tools would see and access the data at the remapped sectors, not the original contents. It is possible using advanced data recovery tools and techniques for the investigator to still access the original location and data, regardless of whether it was a genuine system shift or the cause of system manipulation. In this case, in attempting to access these areas the investigator will have to work on, and alter, the original media, which would require changes to best practice procedure.

Detecting this form of misuse is potentially difficult. The major problem lies in the ability to obtain and verify the error lists. The error lists and certain portions of the data contained in the firmware /

system area are unique and disk model specific and therefore cannot be compared to another version of the disk. A donor drive with matching firmware is required and this can still differ in parts somewhat. However, it is possible to return the hard disk drive back to it's original 'factory settings' with regards to the original list of all available physical sectors that can be used by the disk, whilst still retaining user data. This would enable the investigator to access and forensically analyse all the original sectors. Advanced data recovery tools would be necessary to facilitate the manipulation of the defect control system and also adjust the hard disk drive read / write parameters, enabling the retrieval of data from all sectors regardless of whether they are legitimate bad sectors, as these sectors have not yet ultimately failed but have been marked as bad by the system and discontinued from service.

To access all original CHS locations that were available at the time of manufacture, the G-list module for the defect management system would have to be fully cleared. Carrying out this action would allow the retrieval of any data that may have been maliciously hidden via the firmware steganographic technique previously devised. However, all data that has been saved to the reserved sector area will have been lost at this point, which would result in the loss of potential evidence data.

FORENSIC BEST PRACTICE

These types of malicious techniques have the potential to impact upon forensic best practice. There are two possible alternatives when proposing best practice for this type of analysis. The first is to assume that this type of analysis is required in each case, clearly this would be prohibitively expensive, time consuming and unnecessary in the vast majority of cases. The second option is to apply this type of analysis when the evidence indicates the possibility of firmware tampering. The latter would have to be indicated by a combination of the suspect's technical expertise, the presence of certain hardware and software tools at the scene and suspected incomplete or missing evidential material.

Where there are grounds for suspecting that a suspect may have modified drive firmware, then there are a number of actions that could be constituted as best practice in this type of case. At the crime scene this impacts the seizing of any computer equipment. Any products that have data recovery branding or related documentation need to be seized. The functionality of the device needs to be examined to determine if it has the capability to perform such exploits, informing the subsequent direction of the investigation. If data recovery products are recovered then the following recommendations are made:

Firstly, provided the hard disk drive is functioning correctly and allows access to the user data, a standard image of the drive should be acquired as a baseline for further detailed investigation. All hard disk drives will develop legitimate bad sectors due to natural wear and tear, these bad sectors will invoke reserved space into service and data will be remapped to this area as previously discussed. Over time it is likely the data at those LBA locations will change, so an alternative copy cannot be guaranteed. For this reason it is necessary to obtain a baseline image first with which to work from, otherwise, further firmware modifications will prevent access to this reserve area and its data resulting in the loss of potential evidence.

Secondly, once a baseline image has been created, it is suggested that the investigator should use data recovery tools and techniques to reset the dynamic defect list, the G-list, clearing its contents, which would re-align all of the original LBA numbers to their CHS counterparts. After the realignment has been successfully completed it is crucial that the hard disk drive read parameters are altered, allowing the drive more time to read from all sectors. If this is not altered, the disk may encounter problems when reading the sector and re-mark the sector as failed, subsequently not acquiring any data. This is important as malicious users may take the extra step to not only hide data in physical sectors not accessible by the disk itself, but also to hide data in legitimate failing locations, so that if the defect G list was re-set, the data would still not be fully obtainable due to default drive read time configuration.

The proposed method would allow a comparison of data between the two images. The first forensic image of the drive would contain all data residing in the user data area and in the reserved area; the second would contain all data from the user area. This would contain all the data in the original physical sectors of the drive that were mapped out due to natural wear and tear, and or through suspected manipulation of the firmware to conceal data. Best practise would then be to separate the duplicated data as much as possible from the two images, highlighting the differences. This could be achieved by creating a known file filter from the first image. This could be facilitated via MD5 or SHA-1 hash recognition systems, and could then be used to disregard all known files from the second image, leaving the variations to be analysed in a separate environment from that of the first original forensic image. Both images together would then make up a complete copy of all data contained on the drive, allowing a comprehensive forensic investigation to be performed.

Where the suspect may be in the possession of data recovery hardware and software tools and in the event of the drive not responding to any forensic imaging attempts, there is the possibility that the drive has been purposely sabotaged via firmware corruption methods. The drive may possibly be repaired; it would depend upon the type of corruption and the model of hard disk drive itself, as all have differing firmware implementations. Best practice would be to use advanced data recovery tools and techniques to firstly diagnose the exact form of firmware or hardware corruption. This will provide an indication as to how the drive was damaged and also how to proceed with the recovery. It is possible that some repairs can be made with data recovery tools without the use of donor disk drives. The drive can then be imaged via standard practise, taking into account the above mentioned current best practise method to gain access to all areas of data.

In other cases where donor drives are required to enable repair, the investigator would have to source a suitable donor hard disk drive and begin the recovery using the original diagnosis as a starting point. If full recovery of the disk and the file system cannot be achieved, it is possible to recover some data through standard data carving techniques e.g. Scalpel. (Richard III & Roussev 2005) Advanced data carving techniques would enable the investigator to carve out tangible user and or system files from the raw data available. The proposed two part imaging method outlined above would be have to be implemented to facilitate carving from all data areas.

Malware engineered to target hard disk firmware can in theory render a full acquisition of the data contained on the drive very difficult to achieve even with sophisticated data recovery tools and donor parts. This could be due to the malware targeting disk specific critical subsystems contained in the firmware, damaging the drive. This is an area for further research. The drive in this scenario would not be able to reach a ‘ready’ state and the firmware / system area would be inaccessible for diagnosis and repair. In this case the investigator should attempt to use advanced data recovery techniques, which may be able to achieve a ‘ready’ state. One example of such a technique would be to emulate the service area of a matched donor disk on to the hard disk needing to be recovered, in some cases enabling access to the firmware so repairs could be made, again, if possible.

SUMMARY AND CONCLUSIONS

There are a number of potential problems relating to the forensic analysis of malicious hard disk firmware modification. Without the correct knowledge of the systems it can be very difficult to find or reverse this type of modification. Hardware and software costs supporting this type of analysis are significant. The correct training is not widely available and is expensive. While this remains unlikely to impact the vast majority of forensic cases, the increasing availability of the data recovery tools used to carry out this work makes it a possible area for future concern. In this paper the forensic impacts of such hard disk firmware exploits have been discussed and suggested current best practice has been put forward for the correct handling of such cases.

ACKNOWLEDGEMENTS

The authors would like to acknowledge members of the Information Security Reach Group (ISRG) at the University of Glamorgan in particular Prof. Andrew Blyth and Mr. Nick Pringle. And EBS for the proof read.

BIOGRAPHY

Mr. Gareth D. O. Davies is a Ph.D. Student in the Faculty of Advanced Technology at the University of Glamorgan. The main focus of his research is the security and forensic analysis of hard disk technology. He is a part-time lecturer on the Computer Forensics undergraduate degree at Glamorgan University and has been involved in a variety of other research projects in the area of Computer Forensics and Information Security. Mr. Davies has also acted as a consultant and investigator on forensic and disk recovery technology cases in the University's Computing Forensics Laboratory.

Dr. Sutherland is Reader of Computer Forensics in the Faculty of Advanced Technology at the University of Glamorgan. His main field of interest is computer forensics, he maintains the University's Computing Forensics Laboratory. Dr. Sutherland has acted as an investigator and consultant on both criminal and civil cases. 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 on the university's computer forensics degree schemes.

REFERENCES

ACPO (2008) Association of Chief Police Officers Good Practice Guide for Computer based Electronic Evidence, Version 4.0
www.7safe.com/electronic_evidence/ACPO_guidelines_computer_evidence.pdf
Accessed 9th April 2010

Badtrk (ADM) Documentation http://docsrv.caldera.com:507/en/man/html.ADM/badtrk.ADM.html

Accessed 11th March 2009

Blyth A.J.C., Sutherland I, Pringle N., (2008) *Tools and Techniques for Steganography and Data Insertion onto Computer Hard-Drives*, 8th Annual Program Manager's Anti-Tamper Workshop, Sponsored by US DoD Anti-Tamper Executive Agent SAF/AQL and Department of the Army, Redstone Arsenal, Huntsville, AL, USA.

Carrier B, (2005) *Forensic File System Analysis*, Addison Wesley.

DoJ (2007) Department of Justice, Computer Crimes and Intellectual Property Section Prosecuting Computer Crimes
http://www.justice.gov/criminal/cybercrime/ccmanual/index.html
Accessed 9th April 2010

Gupta M.R., Hoeschele, M.D., Marcus K. Rogers M.K., (2006) *Hidden Disk Areas: HPA and DCO.* International Journal of Digital Evidence, Fall 2006, Volume 5, Issue 1

Browsedata (2004) HDD Firmware Serial Number Source Code 1.01 Free Download http://www.softlow.com/windows/development-tools/debugging/shareware/hdd-firmware-serial-number-source-code.html
Accessed 11th March 2009

Sutherland I. & Davies G. (2009), *The Impact of Hard Disk Firmware Steganography on Computer Forensics*, Proceedings of the Conference on Digital Forensics, Security and Law, 20-22nd May 2009.

Richard III G.,G., & Roussev V. (2005) *Scalpel: A Frugal, High Performance File Carver,* Digital Forensic Research Workshop (DFRWS) New Orleans, Louisiana, August 17-19, 2005. Tool available at: http://www.digitalforensicssolutions.com/Scalpel/
Accessed 10th April 2010

Zhenliu Zhou1, Jiapeng Fan1, Nan Zhang1, and Rongsheng Xu2 (2009) *Advance and Development of Computer Firmware Security Research*, Proceedings of the 2009 International Symposium on Information Processing (ISIP'09) *Huangshan, P. R. China*, 21-23 August 2009.