Combating Information Hiding Using Forensic Methodology

C. Balan, D.S. Vidyadharan, S. Dija and K.L. Thomas

Centre for Development of Advanced Computing, Trivandrum, India
e-mail: cbalan@cdactvm.in, divyasv@cdactvm.in, dija@cdactvm.in, thomaskl@cdactvm.in

Abstract

Advancement in disk technology led to the development of hard disks of terra byte sizes. Users have the option to divide the storage into a number of partitions based on the nature of uses. In case of Master Boot Record partitioning scheme, whenever a partition is created, the complete track containing MBR/EMBR of the storage media is reserved to store boot information and partition table information. But this information requires only the first sector of the track. The remaining sectors in that track cannot be used for any other purpose, as the file system cannot access these free sectors and hence, the chances of overwriting these sectors are very low. So, this area can be used to hide any critical information. The user will get a large amount of storage space for hiding data depending on the number of partitions. This area becomes an important area in Forensics Analysis. In this paper, first we describe the details of data hiding in unused areas, which cannot be easily overwritten by the Operating Systems and how it can be analysed using standard cyber forensics software. Analysing such areas using cyber forensics tools may give lot of valuable information and also will lead to reduce the criminals in hiding information.

Keywords

Data hiding, MBR, EMBR

1. Introduction

Certain kind of users are interested in storing data in a hidden manner. This may be to protect crucial information from intentional access by others. Data can be made hidden in a number of ways without entering any information into the data structures of the file system. For example, data can be hidden inside unused bits of an image file (Hal Berghel, 2007). Data can be stored inside the free area of a cluster used by an ordinary file. But all these methods have its own flaws. For example, the data stored in free space of clusters can be overwritten as more data is added to original file.

In a disk, there is an area that is not considered as part of the partition and hence invisible to the file system. This area constitutes the sectors reserved for storing the master boot record. Boot record needs only one sector space. The remaining sectors in the track can be used for data hiding. The file system and most of the currently available tools are not capable of accessing the information stored in this area. So a

chance of overwriting this area is very less. Thus these free sectors become a very valuable location to keep important information from external access.

2. Disk and Partitions

Hard disk drives are the popular digital storage media used in desktop systems. It comes with different capacities. Nowadays, commonly available capacity ranges from 160 GB to 2 TB. As on July 2010 the highest capacity available is 3 TB. The hard disk drives are divided into tracks and tracks are further divided into sectors (Brian Carrier, 2005). Recently, sector size in a hard disk has changed from 512 bytes to 4096 bytes (Seagate Corporation Web Site, 2011). But the most of the application are still considering a sector as 512 bytes. Built-in emulator inside the hard disk handles the necessary conversions. Here, for explaining data hiding areas, we are taking a sector as a 512 byte sized unit.

Storage space of the disk can be further divided into partitions. Partitions can be either primary or extended. Fig. 1 shows an example of 80GB disk, divided into five partitions with one primary partition. First, the total space is divided into a primary partition and an extended partition. The extended partition is further divided into four logical partitions.

Bytes	Data
0-445	Boot Code
446-509	Partition Information
510-511	Signature (0x55AA)

Table 1:MBR Structure

Boot code and partition information is held in the Master Boot Record (MBR). MBR is the very first sector in the disk and the structure is as shown in Table 1 (Chad Steel, 2006). In a sector of size 512 bytes, the first 446 bytes is the boot code and the next 64 bytes contains information about 4 partitions, each with 16 bytes and ends with signature of 2 bytes. The MBR can hold a maximum of four numbers of partitions, 3 primary and one extended; or four primary partitions. It allows only one extended partition and it can be further divided into logical partitions. The partition table information of 16 bytes length gives the size, starting sector of the partition, number of sectors and the file system of the partition. Fig. 2 shows the boot sector of a disk, with the partition table information highlighted.

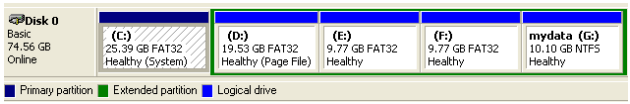

Figure 1: Disk 0 is divided to five partitions

3. Extended Master Boot Record (EMBR)

Extended master boot record (EMBR) is the first sector of the extended partition. The logical partitions in the extended partitions are created as primary partitions in the EMBR. The structure of EMBR is same as MBR but it can hold only one primary and one extended partition. So every logical partition makes a chain of primary and extended partitions.

MBR comes only in the first sector of a track. The remaining sectors in the track containing MBR is free and cannot be used by the partition for any other purpose. So these sectors are completely unused. The EMBR also uses only the first sector from the track assigned to it. So there will be free sectors in each of these logical partitions.

Since the partition cannot start from the second sector of the track, the sectors other than the first sector allocated for MBR/EMBR is not used for any other purpose by the operating system. Normally these sectors in the MBR/EMBR tracks remain unused. Within the EMBR sector, 32 bytes are used to store the information of two partitions and 2 bytes for the signature. The remaining 478 bytes are free since there is not boot code inside the EMBR. So, in the case of EMBRs we have free sectors plus 478 free bytes. This is the area where users can hold or hide secret information of any kind. Again, tools that can retrieve the contents of sectors can show the hidden data. So, the data has to be stored in an encrypted form.

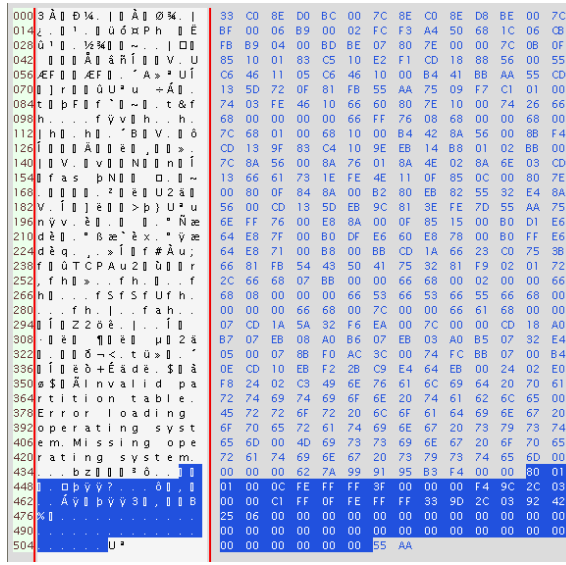

Figure 2: The Boot sector with Partition information highlighted.

It is clear that as the number of partitions in the hard disk increases, the percentage of unused sectors in the disk also grows. This wasted area can be used for storing any important information, which is to be hidden because of any reason.

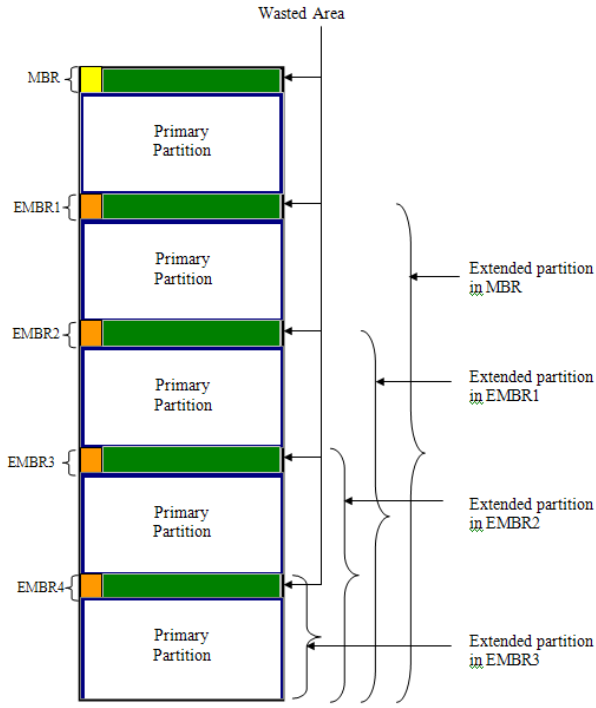

Figure 3: EMBR and the slack space inside a single track

Since this space cannot be accessed by the file system, there is no way to recover the hidden data from these sectors. Fig. 3 shows MBR and EMBR along with the wasted space inside track allocated for storing boot record.

4. Other Slacks

The unused space inside the track allocated for the MBR is called MBR slack and the wasted space inside the EMBR is called EMBR slack. Other than these MBR and EMBR slacks, there are a number of unused spaces or slacks in a hard disk.

4.1. Disk Slack

The sectors those are left after dividing the total space in to a number of partitions are called disk slack. So these are the unpartitioned sectors inside the disk. Normally file system cannot access this area and so this area can be used for information hiding.

4.2. Partition Slack

Cluster is the basic data allocation unit within a partition. The partition size may not be a multiple of the cluster size. So there will be a number of free sectors within the partition. These sectors are known as partition slack. Data hidden in this area will never be overwritten until formatting or repartitioning is done.

4.3. File Slack

In windows file systems, a file or folder takes at least one cluster for storing the data. Cluster size can vary from 512 bytes to 64 KB. File size can be of any size limited by the file system attributes. Since the file size cannot be multiples of cluster size always, there may be free space in the last cluster of the file. This is named as a file slack. Briefly it can be defined as space from the end of the file to the end of the last sector allocated to that cluster. Space from end of the file to the end of containing sector is called RAM Slack. RAM Slack will always be less than the size of a sector.

The partition slack and disk slack can be used for information hiding where there is less chance of overwriting the data without formatting the disk. But in the case file slack and RAM slack, the space may get overwritten at any time by modifying the file content.

5. Writing Data

Before writing the data that is to be hidden, first calculate the sectors or bytes that are free in the disk. Fig. 4 shows a flow chart for computing the MBR/EMBR slack inside a hard disk. After this, we can use a hexadecimal editor tool for writing the required content to these sectors. The advantage here is that no file system or forensic tool will show this as a file while inspecting the disk.

6. Advantages

The main advantage of using MBR Slack, EMBR Slack, Partition Slack and Disk slack for information hiding is that, nobody can easily overwrite the content since these sectors are not accessible by the file system. In the case of RAM Slack or file slack the data may be simply overwritten while deleting the files or modifying the files. L. Shu-fen et al. use a mechanism to mark the sectors that are used for data hiding as Bad Sectors (Shu-fen L. et al., 2009). But here, there is no point in marking the sectors as bad sectors as the file system will never access these sectors. So this information hiding methodology offers high flexibility and low risk. The MBR slack data cannot be over written even by re-partitioning.

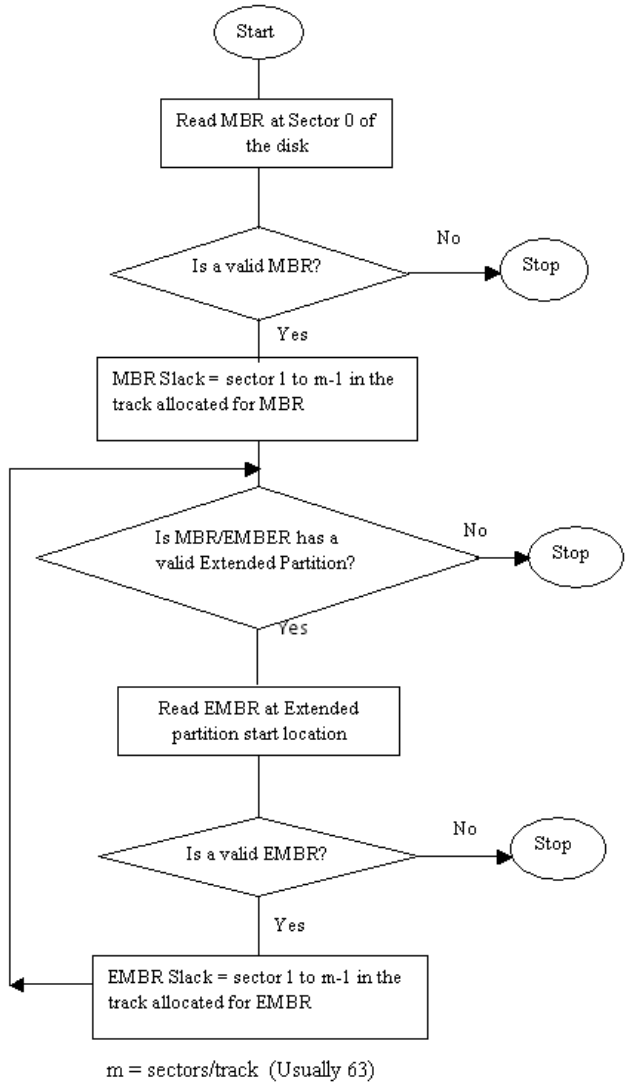

Figure 4: Flowchart for computing MBR/EMBR slack.

So we can use the MBR/EMBR slack for hiding important data for a long term. Also, this method is secure and the complexity is low. Another major advantage is that, we can hide large chunk of data, that requires more than one EMBR slack, can be stored as fragmented in different EMBR slacks in the same hard disk. It enables the user to hide big files inside the hard disk without much risk. If we are storing the data in an encrypted form then it will not be directly readable using a hex editor. These features make this method of data hiding very useful for a wide range of information hiding applications.

7. Hidden Data Recovery

We have seen different areas within a hard disk that can be utilized for intentional data hiding by criminals. Cyber crime analysts are interested in finding hidden data as they may give very crucial hints for proving crimes. Cyber forensic analysis starts with the acquisition of digital evidence without the support of operating system and file system. In the case of hard disk, a bit stream copy commonly known as disk image of the hard disk is acquired and further analysis is carried out in this copy. In this disk image, all the sectors are available and so the analyst can access every location including MBR/EMBR slack. The analyst should give special attention to the contents residing in the slack.

8. Conclusion

We have discussed how criminals can hide data inside a hard disk without the intervention of file system. As this is a safe method of hiding data without the fear of being overwritten by other files, there are chances that this fact can be misused by anti-social elements. So depending on the purpose for which the hidden data is stored, we have to devise methods for quick recovery of data stored in MBR/EMBR slack of large hard disks with numerous partitions. Standard Cyber Forensics Tools can analyse data hidden in these areas and can provide protection from these antisocial activities.

9. References

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

Steel, C. (2006), *Windows Forensics: The Field Guide for Corporate Computer Investigations*, John Wiley & Sons, Indiana.

Berghel, H. (April 2007), "Hiding Data, Forensics and Anti-Forensics", *Communications of the ACM*, vol. 50, No.4, pp15-20.

Shu-fen L., Sheng P., Xing-yan H., Lu T. (2009), "File hiding based on FAT file system", *IT in Medicine & Education*, vol. 1, pp1198-1201.

Seagate Corporation Web Site (2011), "The Transition to Advanced Format 4K Sector Hard Drives", www.seagate.com/docs/pdf/whitepaper/tp613_transition_to_4k_sectors.pdf, (Accessed 31 May 2011)