File Allocation and Recovery in FAT16 and FAT32

Riya Madaan
Department of Computer Science & Applications ,Kurukshetra University, Kurukshetra-136119
madaanriya92@gmail.com
Rakesh Kumar
Department of Computer Science & Applications, Kurukshetra University, Kurukshetra-136119
rakeshkumar@kuk.ac.in
Girdhar Gopal
Department of Computer Science & Applications, Kurukshetra University, Kurukshetra-136119
girdhar.gopal@kuk.ac.in

----------Abstract----------

The data recovery is the fastest emerging dynamic technology with a huge market in the area of computer security and maintenance. In order to carry out the recoveryone is to be acquainted with the file management systems i.e. FAT, NTFS. FAT is the oldest file system which was used in MSDOS and early versions of Windows. In this paper, an exhaustive study has been performed for the two variants of FAT file systems like FAT16 and FAT32 with respect to data recovery. In addition the main differencesbetween FAT16 and FAT32 are discussed. Recovery issues are also addressed. Some techniques to recover the data that have been deleted accidently or maliciously have also been reviewed.

Keywords- Digital Forensics, File Recovery, FAT, File System, Storage Principle

I. Introduction

Over the years computers have been gradually but unavoidably became record keepers of human activity. This trend enhanced with the advent of PCs, handheld devices such as mobiles, Internet, multimedia and telecommunications. Today's organized world of digital devices grants opportunities and challenges for criminals and investigators, for governments and privacy wanted citizens, for commercial and for other activities. Forensics is the application of science for criminal and civil laws which is used to solve a legal problem. Forensic scientists gather, preserve and evaluate scientific evidence during the course of the investigation. Due to increasing scope of forensics, the categorization was done further and new terms were emerged i.e. computer forensics, digital forensics, Network forensics, OS forensics etc.[1] Digital forensics (DF) is a branch of forensic science which has been grown from a relatively incomprehensible tradecraft to a significant part of many investigations.

The data is very vital in this current world because the data may be vanished either by users own wish to delete it due to some storage issues or by accidentally. In future, if the user needs the same data, it will not be possible at that time to fetch it back; it can only be retrieved if a backup copy was taken. Data recovery for the common people or for the purpose of forensics i.e. digital forensics is an evolving field in computer applications. In the scenarios of real life challenges like cybercrime investigations, a recovery technique would be a boom. Retrieving and analyzing of the records stored in various storage devices becomes an important part of taking evidence from computers. The Windows Operating System (OS) and file system has played a vital role in our life. Retrieving and analyzing the useful data of different file system is playing a very important role in computer forensics. If the file gets deleted, we couldn't directly observe it under Windows Operating System, but it sometimes includes some important evidence of crime. Data recovery varies from data backup in recovering lost

or unreachable data which do not have an identical copy of the original file containing data. To overcome from these issues, a recovery technique will be used for common people or for any person who is not aware about storage media and how the procedure followed to store data on computers, USB, hard drives etc. There are two approaches for performing data recovery and these are - Physical data recovery and Logical data recovery. These approaches are used for corresponding physical failure and logical failure. Physical failure occurs when the storage devices are physically damaged or mechanically and structurally flawed. Logical failure occurs when a file gets contaminated with the virus or deleted[2].

The remainder of this paper is organized in five sections as follows. Section 2 contains the overview of FAT. Section 3 discusses related work. Section 4 discusses the difference between FAT16 and FAT32. The paper concludes in Section 5.

II. Overview of FAT

The file systems are the most vital part of the computer and it is the need of them because it is a method for the durable storage and retrieval of data. File systems offer a mechanism for users to accumulate data in a sort of hierarchy of directories and files. A file system is comprised of structural and user records that are arranged such that the computer will always know where to find them. In most of the cases, the file system is always self-governing from any of the specific computer. These are known as record-storing techniques and every file instance had a unique size. File systems always have a specific structure that is very helpful to store one or thousands of files in the storage array and some of the data needs core structure and the organization inside their file structure.

File allocation table is the oldest file system of windows. The different versions of FAT have their own native structure, but the basis on which they have their structure is same for the FAT file systems. FAT was designed for small disks and simple folder structures. The basic concept of this file system is that it has two important data structures- File allocation table (FAT) and Directory entries. Each file and directory always allocatesa data structure that is known as directory entries. It contains 3 fields-

File Name	Size of file	Starting address

fig 1. Structure of Directory Entries

A FAT file system names the data units as a cluster. A cluster is a collection of consecutive sectors. Each cluster is given an address and the address of first cluster is 2.[3] In FAT12/FAT16 cluster 2 always follows the root directory, but in FAT32 cluster 2 is the first sector of the data area. The cluster size of FAT16 must be a power of 2 between 512 and 65,536 bytes. To know the addresses in terms of clusters and sectors, we need to know how many sectors are there in one cluster .The formula used for calculating the sector address of cluster N is-

(N-2) * (No of sectors per cluster) + (Sector of cluster 2)

The allocation status of any cluster is always determined by FAT. The cluster is marked damaged and cannot be allocated, if the table entry is 0xff7 in Fat12, 0xfff7 in FAT16 and 0xffffff7 in FAT32.

The file and the directory content are stored in the form of clusters. A file can have more than one

cluster and it is always found by the FAT structure and also the allocation status of clusters is also checked. The FAT file system is divided into different areas- Boot block, FAT, Root directory, File data area.

The physical layout of the FAT is having 3 areas- Reserved area, FAT area and Data area.

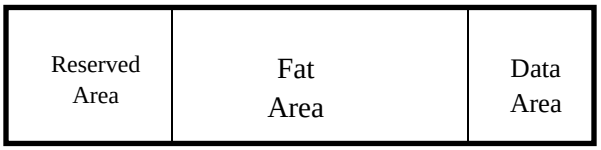

fig 2. Physical layout of FAT

Reserved area- This is the very first area and it is usually 1 sector and the size is decided by the boot sector.

FAT area- This second sectioncontains two copies, one main structure and one backup FAT structure. There are mostly multiple FATs to increase fault tolerance on floppy disks because of the higher possibilities of errors. The data area of the disk is divided into the form of clusters. The important point to keep in mind is the cluster numbering always starts from 2 as cluster 0 and 1 does not exist. The first byte of the first entry is a replica of the media descriptorbyte, and the second byte is set to 0xff. Both bytes in the second entry are set to 0xff. There is one entry in every cluster of the disk; we can say there is a 1:1 mapping. The information the cluster contains is the no of successor cluster so that the data can be retrieved by following this chain and the last cluster contains the entry 0xffff to indicate the end of data cluster, this process is known as cluster chaining.

Data area- This is the last area which contains the clusters that will be allocated to store the file and directory content.

Boot block- This occupies the first block of any disk and for loading the OS into memory a special Bootstrap program is used. It is about 512 bytes and in FAT12 and FAT16 the 62 to 509 bytes and in FAT32 the 90 to 509 bytes are not used because it contains the boot code. It also contains some vital area which describes the rest of the file system. The first three bytes of the boot block have the jump instruction that causes the CPU to jump to rest of the boot code. The contents of the boot block are given below-

Table 1. Description of Boot block

Offset from start	Length	Description
0x00	3 bytes	Instruction of boot code
0x03	8 bytes	OEM name.
0x0b	2 bytes	Number of bytes per sector
0x0d	1 byte	Sectors per cluster
0x0e	2 bytes	Size in sectors of reserved area
0x10	1 byte	Number of FAT's
0x11	2 bytes	Number of root directory entries
0x13	2 bytes	No of files in root directory
0x15	1 byte	Media Descriptor
0x16	2 bytes	No of sectors in FAT
0x18	2 bytes	Sectors per track of storage device
0x1a	2 bytes	Number of heads
0x1c	4 bytes	Number of sectors before

		start of partition
0x20	4 bytes	Total number of sectors in the file system
0x24	2 bytes	Physical drive number
0x26	1 byte	Extended Boot Record Signature
0x27	4 bytes	Volume Serial Number
0x2b	11 bytes	Volume Label
0x36	8 bytes	File system type label
0x3e	0x1c0 bytes	The remainder of the bootstrap program.
0x1fe	2 bytes	Boot block 'signature' value

Root directory- The root directory is 32 bytes entry and contains an entry for each file whose name appears at root of file system. The following table illustrations a summary of root directory; note that the offset is merely from the start of the particular entry, not from the beginning of the block.

Table 2. Description of root directory

Offset	Length	Description
0x00	8 bytes	Filename
0x08	3 bytes	Filename extension
0x0b	1 byte	File attributes
0x0c	10 bytes	Reserved
0x16	2 bytes	Time created or last updated
0x18	2 bytes	Date created or last updated
0x1a	2 bytes	Starting cluster number for file
0x1c	4 bytes	File size in bytes

File name- The first 8 bytes contains the filename because the standard directory contains only eight characters in the name and three characters in the extension. If the name doesn't contain only eight characters then unused bytes then those bytes are filled with ASCII value for space i.e. 0x20.The first byte is vital because it tells the allocation status, if set to 0x00 or 0xe5 then we can conclude the directory entry is unallocated.

III. Related work

The data recovery for the data under FAT file system was performed manually using disk analysis tool i.e. Winhex by Yao Qingshan[4]in 2010. The basic concepts for data recovery comprises of some basic concept of file recovery, the basic cataloging, and the causes and indications of the loss of data. The recovery of FAT32 file system data is done by first analyzing the main boot sector and then by analyzing the root directory structure. The recovery was done very efficiently and easily but there were some demerits while doing recovery with the use of Winhex[5]. If it is a text file then it can be easily accessed but if we talk about the file containing an image, restoring them is troublesome. The deepen knowledge to recover data in a FAT32 system needs the knowledge of assembly language for detailed calculation process and for a groundwork of future in-depth study.

To overcome the shortage of FAT file system i.e. low speed H.Zhao, Chang & Zangproposed a new way to improve it in their paper[6]in 2015.The use of map management and B+ tree improved the speed of query under FAT file system. The storage and management was performed with the help of B+-tree index using map. The advancement can be done by optimizing the B + tree indexing such that the

number of tasks at the same time visit B+ tree hotspot where the temporary copies are stored.

IV. Differences in FAT16 and FAT32

The recovery procedure for Fat16 and Fat32 is same but they differ only in terms of the size like FAT16 has 16 bit entries and FAT32 has 32 bit entries, data structures for Boot sector, Root directory etc. There is no single field that identifies a FAT file system as FAT12, FAT16, or FAT32 [7].They have different numbers of clusters and the method which calculates the type is by calculating the number of sectors in the root directory by the given formula –

((NUM_ENTRIES *32) +(BYTES_PER_SECTOR -1)) / (BYTES_PER_SECTOR

- The foremost difference among FAT16 and FAT32 is the logical partition size.[3] FAT32 do not have the limitation of the 2-GB logical drive because it can extend a single logical drive with a capacity of at least 127 GB.

- FAT32 always uses 4 bytes per cluster in the file allocation table. This varies from FAT16 as it uses 2 bytes per cluster within the file allocation table.

- The FAT16 has the root directory at first sectors of the data area and the FAT32 root directory is dynamic in the data area.

- FAT32 has the main feature that it allows the root directory to grow as FAT16 holds a maximum of 512 entries, and due to use of long file names in the root directory the limit can be even lower.

- The root directory folder on a FAT32 drive is a normal cluster chain, so it can be found anywhere on the volume. For this reason, FAT32 does not constrain the number of entries in the root folder.
- FAT32 uses small clusters (4 KB for drives up to 8 GB), causing 10 to 15 percent additional efficient use of disk space compared to large FAT16 drives. FAT32 also decreases the resources essential for the computer to activate.

The value for the number of entries in the root directory is mentioned in the boot sector, this value for FAT32 is 0. To determine the number of sectors that are allocated in clusters because these values correspond to the maximum number of clusters that a FAT can store is given by the following formula-

((TOTAL_SECTORS-RESERVED_SIZE – NUM_FAT) * FAT_SIZE – ROOTDIR_SIZE)

- If the number of sectors that are allocated in clusters is less than 4,085, the file system should be FAT12; if the value is 4,085 or greater and less than 65,525, the file system should be FAT16 and any size greater than or equal to 65,525 should be FAT32.

V. Conclusion

This paper aimed at providing the basics of FAT file system to set a background for understanding the storage principle used by FAT file system. The FAT16 and FAT32 have no difference in terms of

recovery as both follows the same strategy, but they differ in terms of number of bits, cluster size and location of the root directory. To determine the type of FAT file system calculations are performed to determine the number of sectors in the root directory and number of sectors that are allocated to clusters because these values correspond to the maximum number of clusters that a FAT can store. Data recovery expertise is the fastest emerging and best dynamic technology, with a huge market and progress predictions in the area of computer security and maintenance. The data recovery approaches used for FAT file system can recover documents, images and media depending on different file systems. Further studies would involve developing a complete data recovery tool for FAT file systems which can recover both FAT16 and FAT32 formatted devices using indexing by B+ trees, parsing techniques used by the object oriented method, user friendly GUI frameworks and which will be scalable for bigger file systems.

REFERENCES

[1] Simson L. Garfinkel, "Digital forensics research: The next 10 years," Elsevier, pp. 64-73, 2010.

[2] R. Kalal Soumya , V. Mandal P. Ravindra, "Logical Data Recovery Technique for USB Devices," in International Conference on Emerging Trends in Communication, Control, Signal Processing & Computing Applications (C2SPCA), Bangalore, 2013, pp. 1-6.

[3] (2016) Microsoft TechNet. [Online]. https://technet.microsoft.com/en-us/library/cc938438.aspx

[4] Gu Chunying Yao Qingshan, "Research and Implementation of Data Recovery Technology Based on," in International Conference on Machine Vision and Human-Machine Interface (MVHI), China, 2010, pp. 549-552.

[5] Shen Yong LI Chunwang, "FAT data recovery following the devastation of the table," in Technical Education Journal, 2003, pp. 26 – 31.

[6] X. Li H. Zhao, L. Chang, and X. Zang, "Fat File System Design and Research," in International Conference on Network and Information Systems for Computers (ICNISC), Wuhan, 2015, pp. 568-571.

[7] Brian Carrier, File System Forensic Analysis.: Addison-Wesley, 2005.