Secure Data Deletion for File System by Automated Process

AJAY BALAJI D

Abstract:

The permanent removal of data from filesystem is a major area of concern mainly because of the misconception that once a file is deleted or storage media is formatted, it cannot be recovered. In this paper we will be focusing on NTFS filesystem. If we delete files on a NTFS partition, we can recover them shortly with a very high probability of success. When a file is deleted, only the file system entry is deleted, thus the content of the file will remain intact for a period of time, depending on the disk activity afterwards. We have number of ways proposed over the period to make data unrecoverable like DoD 5220.22-M, Gutmann method, British HMG IS5 etc. This paper will propose a way to make our file system more secure by making the permanently deleted file unrecoverable after certain amount of time automatically.

1. Introduction

With the rapid development of information technology, computers play an increasingly important role in people's work and life, and computer information security issues are becoming more and more concerned. A large amount of information is stored in the form of data on computer file systems. Most users believe that such information is erased as soon as they delete a file. Even those who know that this is not true often ignore the issue. Most users want deleted files to be permanently erased, and they believe that deleted files are physically erased. Technically, when a file is deleted, the operating system only removes pointers to the deleted data. The information remains on the hard disk until another file overwrites it.

One third of hard drive resold on eBay contains confidential information such as credit card or medical records. That is why government agencies, and some businesses require proper sanitation of the physical media that was used to store sensitive information. However, as in the case of many security-related problems, security must be balanced with convenience and performance. There has been the development of both commercial and freeware data erasing tools, which all claim complete file or disk erasure. We can select any one free software and carry out our data recovery. Any person will zero knowledge about file system can do this.

At present, the market share of Windows operating system exceeds 92%. FAT32 and NTFS are the popular file systems on Windows operating system. A file system is the main mechanism responsible for the management of data on a storage medium. Data recovery techniques are mostly used for digital forensics, Accidental deletion of files or folders, Logical damage to the file system, etc. But now a days, illegal access to computer systems is increasing rapidly, that's why need of computer security has been increased.

In computing, file system controls how data is stored and retrieved. In other words, it is the method and data structure that an operating system uses to keep track of files on a disk or partition. It separates the data we put in computer into pieces and gives each piece a name, so the data is easily isolated and identified. There are five types of Windows file system, such as FAT12, FAT16, FAT32, NTFS and exFAT. This paper is to focus on the NTFS filesystem. Its partition structure is shown in [Figure 1].

Figure 1. File system structure on NTFS

The organization of this paper is as follows, Section 2 describes the related research and limitation of current method. In Section 3 we describe our hypothesis and calculated metric to prove why our approach is better. We conclude in section 4 and discuss future work.

2. Related Research

The study, by N. Zhang, Y. Jiang and J. Wang., involved research on NTFS System and its data recovery [1]. This paper analyzes the structure of NTFS file system, and the file records and key attributes. By comparing with FAT32, the NTFS system is found to be easier to find files and recover files due to its advantages in structure and file storage mechanism.

Over the period of time many data recovery software's are developed, and many approaches were developed to erase data permanently. The study by Jones, Andrew and Afrifa, Isaac showed how different data erasing techniques works and which will be efficient in current circumstance [5]. There are numerous software-based data erasing standards currently being used. These include the Peter Gutmann's Algorithm, Bruce Schneier's Algorithm, the US Department of Defense (DoD) 5220.22-M standard, Secure Erase, Random Data, Write Zero, the Russian GOST R 50739-95, the German VSITR method and the British HMG IS5, both Baseline and Enhanced.

[Figure 2] will show different methods of data sanitization and its effectiveness with respect to security and time. User tends to select the method that provides best security in the shortest time. Explanation of each method from [Figure 2] will be shown in the [Table 1].

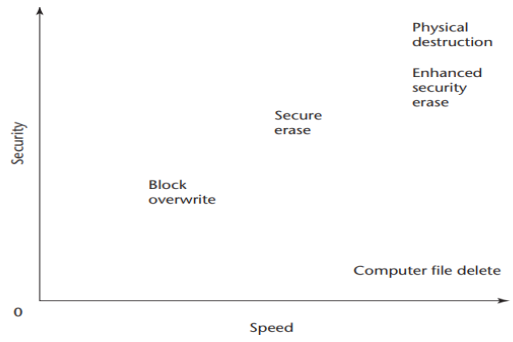

Figure 2. Security vs. speed for data sanitization methods

The above-mentioned papers showed what is the importance of data sanitization and what is the importance of data recovery but in the near future we don't see any mechanism proposed to erase the data permanently from file system (automatically) after some point in time when it was not recovered. As the scope of this research, we will only be concentrating on how we can erase the data from NTFS file system effectively and we will be comparing it results will other standard approaches.

Sanitization type	Average time for 100 Gbytes	Security	Comment
OS file deletion	Minutes	Low	Deletes only file pointers, not actual data; data recovery possible through common commercial data recovery software.
Block overwrite (DoD 52203)	Up to a half-day	Medium	Now obsolete; needs three writes and one verify; might not erase reassigned blocks; multiple cycles leave other vulnerabilities.
NIST 800-88 secure erase (SE)	30 minutes to three hours	Medium	Performs in-drive erase of all user-accessible blocks.
Enhanced SE	Milliseconds	High	Executes SE of in-drive encryption key
Physical or magnetic destruction	Seconds to minutes	Highest	Required for top secret data and above

Table 1. Comparison of data sanitization methods.

3. Hypothesis and Empirical Evidence

From [Figure 2] [Table 1], we can see Block overwrite method will be more efficient with respect to both time and speed. We will be using block overwrite method to automate our approach.

In this research, when a data is deleted from NTFS file system, we will keep track of the deleted file in some temporary file. A job will be scheduled to run every 24 hours from task scheduler. This job will trigger a script and this script will verify the temporary file and find the locations of the deleted data. After finding the location of the deleted files we will overwrite the file location with random data. When overwrite is done then the file will become unrecoverable. In this approach we will give user 24 hours window to recover the file if needed. If not recovered, we will delete the data permanently. Below mentioned python script [Figure 3] will be used for data sanitization,

```
import os

def secure_delete(path, passes=1):
    length = os.path.getsize(path)
    with open(path, "br+", buffering=-1) as f:
        for i in range(passes):
            f.seek(0)
            f.write(os.urandom(length))
        f.close()
```

Figure 3. Secure data deletion script

[Table 2] will shows the different data erasing techniques with respect to block overwrite. All the below methods are not with respect to specific data or location. In our method as we are tracking the deleted file and refreshing it every 24 hours the run time, CPU time and memory usage will reduce drastically comparing to the below mentioned data erasing techniques.

Erasing Tool	Features	Run Time (mm:ss)	CPU Time (mm:ss)	Memory Usage
Remo Drive Wipe	Supports US DoD 5220.22-M,Mtmann's Algorithm, SITR	35.17	6.24	34624 K
Active KillDisk	DoD 5220.22-M,DoD 5220.22-M (ECE), DoE M205.1-2,German VSITR, British HMG IS5 Baseline, British HMG IS5 Enchancedtmann	28.43	0.51	115632 K
Puran Wipe Disk	1pass,3passes,7passes	28.51	0.07	39424 K
SuperFile Shredder	Simple One Pass, DoD 5220.22-M,Secure erasing (7 Passes), Gutmann's Algorithm	24.33	1.19	31992 K
Eraser	Pseudorandom data, BritishHMG IS5, DoD 5220.22-M,German VSITR, Bruce Schneier's Algorithm	28.29	4.03	84004 K
Disk Wipe	DoD 5220.22-Mtmann's Algorithm, One Pass Zeros,Pseudorandom data	65.62	1.25	23216 K
Macrorit Data Wiper	Pseudorandom data, DoD 5220.22-M, Gutmann's Algorithm	28.41	0.13	38152 K
Hard Wipe	Pseudorandom data, write zero, DoD 5220.22-Mtmann's Algorithm, Bruchneier's Algorithm	28.37	0.15	37920 K

Table 2. Result Analysis for Historical methods

4. Conclusion and Future work

We have restricted the scope of this research to only analyze the secure way of deleting the file permanently. For automation to be completed we need to track the deletion process in the file system using a temporary file. This work can be extended to find the mechanism to track the deletion in filesystem. Completely automating the process via proposed method will definitely help in making the system more secure and will reduce the deletion tool runtime.

5. References

[1] N. Zhang, Y. Jiang and J. Wang, "The Research of Data Recovery on Windows File Systems," 2020 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), 2020, pp. 644-647, doi: 10.1109/ICITBS49701.2020.00141.

[2] Kaustubh Aggarwal , Dr. Shravan Kumar Garg. "Computer Forensics: Data Recovery Perspective over Windows and Unix ". International Journal of Innovative Science and Research Technology, 2021, ISSN No:-2456-2165.

[3] G. H. Carlton, "A Critical Evaluation of the Treatment of Deleted Files in Microsoft Windows Operation Systems," Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 2005, pp. 310c-310c, doi: 10.1109/HICSS.2005.8.

[4] Nikolai Joukov and Erez Zadok, "Adding Secure Deletion to Your Favorite File System", Proceedings of the Third IEEE International Security in Storage Workshop (SISW'05) 0-7695-2537-7/05 © 2005 IEEE

[5] Jones, Andrew and Afrifa, Isaac (2020) "An Evaluation Of Data Erasing Tools," Journal of Digital Forensics, Security and Law: Vol. 15 : No. 1 , Article 2.