Final Project – Survey Paper:

Deleted Data Recovery Mechanism

Dalton Murray

College of Business & I.T., Lawrence Technological University

INT 6212, MIS Directed Study

Dr. Eralda Caushaj Nazarko

May 2, 2024

**Final Project – Survey Paper: Deleted Data Recovery Mechanism**

**Introduction**

The world has a heavy reliance on digital data, and in order to store this data Hard Disk Drives, HDDs, and Solid State Drives, SSDs, are most commonly used. Today, HDDs are incredibly inexpensive, they store a large amount of data, however, the problem is they are slow and more fragile due to moving parts. SSDs, on the other hand, depending on the quality and amount of storage can be cheap or expensive, they can store a large amount of data, but not more than HDDs, are exceptionally fast, and more durable than HDDs. HDDs rely on using a platter, spindle, actuator, and slider or head. The spindle rotates, or spins, the platters really fast, and the actuator moves the head and writes 1’s and 0’s on the platter. All of these parts are prone to failure, and physically dropping them can break it very easily. SSDs are less prone to failure; however, they have less total amount of times they can be written and read from than HDDs, they can also fail because of excessive heat because we continue to push them further and further, the faster we make them run, the more heat it generates.

As the world continues to operate, and we continue to rely heavier on digital data, the field of digital forensics and data recovery continues to gain more importance. HDDs and SSDs are not immune to failure, in fact, they can fail rather often. Both HDDs and SSDs, however, have methods of retrieving the data from them after they have failed, stopped working, are broken, or unintentionally or intentionally have had data deleted from them.

This survey paper will ignore why and how the data was deleted in the first place and will focus on solely how to recover data, from both HDDs and SSDs. This paper outlines previous studies which have been conducted, will mention the strengths and weaknesses of previous studies, and will offer future proposals on how to conduct data recovery and studies which should be conducted in the future.

**Previous studies**

There have been a large number of studies conducted on data recovery, how to do it, what makes it difficult to do, and how to improve results. One of these studies is titled “Forensic Trails Obfuscation and Preservation via Hard Drive Firmware” (Underhill, P., et al.). In this study, their research aim was to assess the reliability of forensics tools against anti-forensics and to test forensic and recovery software to find its effectiveness by means of manipulating and recovering data from an HDD and SSD. The anti-forensics methods they used were artifact wiping, which is extracting or deleting cryptographic keys from an SSD by placing malware on the device’s firmware that gives it Direct Memory Access, DMA. Their research showed that if you remove these keys, it’s impossible for an analyst to access any data on the SSD (Underhill, P., et al.). They also used data hiding techniques allowing them to hide data into hidden physical sectors (Underhill, P., et al.). They tested trail obfuscation, which has the goal of misleading or confusing a forensic technique or analyst such as deleting logs, IP/MAC spoofing, proxy servers, using zombie accounts and misinformation, changing header file information, change creation, last modified, and recent timestamps, and other means to obfuscate the data (Underhill, P., et al.). Finally, they also used attacks against forensics tools and processes which detects forensics tools and then destroys, hides, or manipulates data (Underhill, P., et al.). They used a black box testing method with six steps: acquisition of software, identification of software functionalities, development of test cases and reference sets, development of result acceptance spectrum, execution of tests and evaluation of results, and release of evaluation results (Underhill, P., et al.). They used a variety of consumer and commercially available software such as: md5checker, ftk imager, recuva, easeus, fex imager, hex editor, hdd hackr, and bootable usb creator (Underhill, P., et al.). They determined that most of the modern anti-forensics methods were not effective at protecting against the software they used, however, many of the methods did also prevent data from being found or fully restored (Underhill, P., et al.).

There are many other studies which have also been conducted which use handcrafted and Free Open Source Software, FOSS, for anti-forensics and data recovery. One of these studies are titled “Comparison of data recovery techniques on master file table between Aho-Corasick and logical data recovery based on efficiency” which compares different data recovery techniques on the master file table in order to determine which has the highest accuracy and efficiency (Sahib, H., et al.). There is even research being conducted specifically on how we can guarantee data deletion and how anti-forensics are made so that we can further data recovery methods (Gnatyuk, S., et al.). Another example of a study that has been conducted is titled “Developing a method for recovering data on storage media” where they develop their own software and method to beat anti-forensics and try to get the highest possible effectiveness and efficiency of data recovery (Zamolotskikh, V., Sidorenko, V.).

***Strengths of Previous Studies***

There are a large number of studies that have been released, are being released, and are planned for which relate to data recovery. The problem with previous studies is that they are often outdated for modern hardware, modern software, and modern firmware. Thankfully, however, it seems that studies based on data recovery seem to be starting to increase rapidly as it becomes a greater concern. One of the strengths of previous studies is that many corporations and consumers are using outdated technology which means that their studies are still relevant today.

Many of the studies which have been conducted are showing new and improved methods on how to recover data after it has been unintentionally or intentionally deleted. They are also showing how they are capable of beating anti-forensics techniques. With the current studies available, there is a large chance for data recovery to be completely or at minimum partially completed. The problem is when anti-forensics techniques start to be deployed in a malicious way. When anti-forensics techniques are deployed the ability to fully recover or recover any data drops significantly. However, there have been strives in closing the gap between being unable to recover data and being able to recover data.

**Future Proposals / Conclusion**

In the future, we need to focus research on data recovery methods that aren’t targeted to specific anti-forensics, specific HDD brands, specific SSD brands, and specific hardware, software, and firmware. We need to develop methods of data recovery that are more adaptive and more efficient and effective. Many studies are still focusing on HDDs when the world is rapidly moving to SSDs. Manufacturers of HDDs and SSDs also need to conduct more research on on-board redundancy of data while still maintaining security and privacy.

**References**

Gnatyuk, S., Sapozhnik, T., Kinzeryavyy, V., Sopilko, I. (2019). *Modern Method and Software Tool for Guaranteed Data Deletion in Advanced Big Data Systems.* Retrieved from: https://www.researchgate.net/publication/333187531\_Modern\_Method\_and\_Software\_Tool\_for\_Guaranteed\_Data\_Deletion\_in\_Advanced\_Big\_Data\_Systems

Sahib, H., Hidayah, Nurul, Al-Qaysi, Attiah A., Mothana. (2021). *Comparison of data recovery techniques on master file table between Aho-Corasick and logical data recovery based on efficiency.*

Underhill, P., Speakman, L., Oyinloye, T., Eze, T. (June, 2022.). *Forensic Trails Obfuscation and Preservation via Hard Drive Firmware*. Retrieved from: <https://www.researchgate.net/publication/361218163_Forensic_Trails_Obfuscation_and_Preservation_via_Hard_Drive_Firmware>

Zamolotskikh, V., Sidorenko, V. (2022). *Developing a method for recovering data on storage media*. Retrieved from: https://www.researchgate.net/publication/359473661\_Developing\_a\_method\_for\_recovering\_data\_on\_storage\_media

I pledge that on all academic work that I submit, I will neither give nor receive unauthorized aid, nor will I present another person's work as my own.

Dalton Murray