Deleted and Coroupted File Restoration and Safety

HUZEFA BANDOOKWALA [1], ALIZAIN FAROOQUI [2], MUEEZ ANTULE [3],
MOHD. HUSSAIN DAIRYWALA [4], SONALI BHUTAD [5]
[1,2,3,4,5] *Dept. of Computer Engineering, Shah And Anchor Kuttchi Engineering College*

Abstract -- In this constantly developing world of technology, the number of devices that store transfer and receive digital data is dramatically increasing. Some people use these digital data unethically which results in cyber fraud and cyber-crimes. Nowadays, information is stored digitally rather than keeping offline records, and due to increasing criminal activities carried out using either computers or smart phones, it becomes very important and crucial that digital investigators could conduct their analysis on these data properly. There are many tools available in different forms but not a single multipurpose tool, which can ease investigator's job.

Indexed Terms: file index, restoration, extensions, security, digital forensics

I. INTRODUCTION

Digital devices are seen and used everywhere today, in our modern society. These devices handle huge amounts of data. Tampering, misuse, hiding of such data for criminal intent by some users is certain. An increase in the number of cyber-crimes has been observed in the past few decades, creating a constant need for development of tools that could be useful for investigating such cases of crimes. Digital evidence is essential for solving cyber-crimes and this evidence are also very helpful in courts as they can be credible piece of information. In this report, we will cover how our software will help the investigator to investigate the digital files more easily. Here the investigators should ensure that the digital evidence or assets are not modified or tampered by any unauthorized user. The main goal is to detect any threats and unauthorized modification of the files in the storage media so that the security of the device, system or organization can be improved.

II. EXISTING SYSTEM

There are a number of ways in which a person can work with the data that they are investigating on a storage disk. For instance, if the examiner wants to find the original extension of any file, they can either identify it manually by verifying the header of the file using its hex values, write a code to extract the hex values of the file and then verify the header from there, or they can use a tool such as winhex to get the hex values of the file and cross check the header of the file.

For restoration of a deleted file, an application can be used or a program code can be generated to scan the disk and get the list of all the deleted files and to recover them.

However, all of these tasks are required to be performed using different applications, codes or manually. It would be much better to have a system that provides these features in a single application.

III. CHALLENGES IN CURRENT SYSTEM

When a digital forensics Investigator is called, he needs to download several software's on the victim's PC to do his work. Due to this lot of time is wasted in just installing.

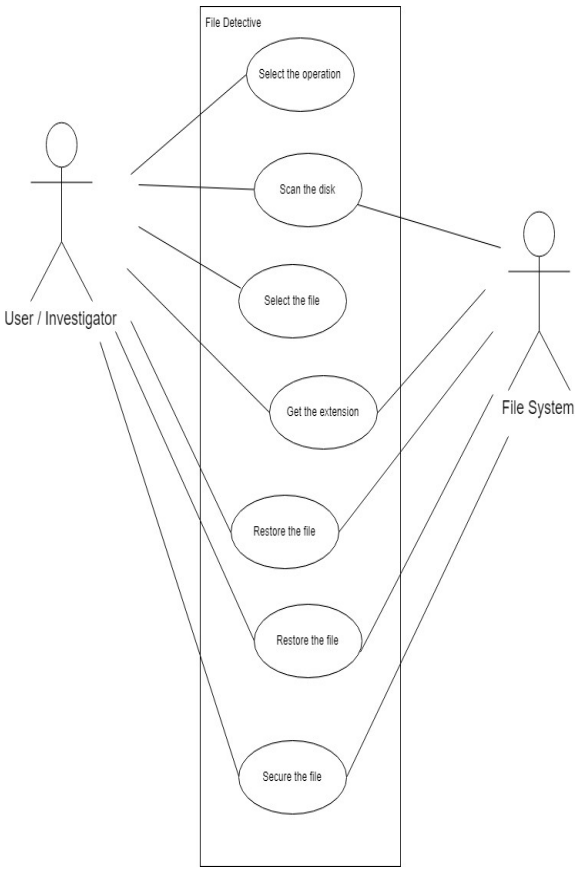
If a file extension is modified by the attacker, it becomes a tedious task for the investigator to get the real file back. It takes lot of time to get the file since it's a manual task and the investigator needs to get the real extension by trial and error method. Since this takes lots of time it could make a huge loss to the victim since it could have some vital confidential information.

Normal users who have no experience in cases where they lose their file intentionally/unintentionally and they want it back for such thing they don't know what to do.

There are catalogues containing several thousand of known file types, without having any global standard for the file types. File type detection methods can be

categorized into three kinds: extension-based, magic bytes-based, and content-based methods, each of them has its own strengths and weaknesses, and none of them are comprehensive or foolproof enough to satisfy all the requirements.

The Use-Case for the system will be as below



IV. LITERATURE SURVEY

- Digital Forensic Investigation on File System And Database Tampering.

  By, Sindhu. K. K, Shweta Tripathi, Dr. B. B. Meshram

Digital forensics is the identification, extraction, analysis and documentation of digital evidence from storage media. It is relatively new technology which is increasingly becoming important as the criminals aggressively expand the use of technology. Digital information is fragile and it can be easily modified or destroyed like File system and Database tampering. In the course of the investigation, the investigator should assure that digital evidences are not modified unauthorized and authenticate submission in the court of law. This paper explained forensic investigation procedures using a WinHex tool. Main focus of this paper is digital forensic investigation of different locations of windows file. This will help us in knowing the file system that is basics for all our modules.

- Computer Hacking Forensic Investigator [2]

  By, EC-Council

This book contains all the information about forensics and how to perform it.This is the official guide provided to all the forensics student so that they can give the exam for becoming the forensics expert. The book is written by the experts of digital forensics that contains various chapters that contains all the information that is need to prepare the algorithms for our modules. The book is used for all the information and reference material that can be used for forensics detail of file system and deleted file recovering process.

- A Comparative Study of File-Type Identification Techniques [3]

  By, William H. Allen, Nasser S. Alamri

Research in file-type identification has employed a number of different approaches to classify unknown files according to their actual file type. In this paper, a comparison of five common file-type identification approaches, along with the parameters used to perform the comparisons. All approaches were evaluated with the same dataset which was drawn from public or widely available sources. The results of this paper show that each approach can produce good results with 88% to 97% classification rates, but achieving these results requires “tuning” the parameters of the inputs to the classifiers.

- An Approach for the Validation of File Recovery Functions in Digital Forensics’ Software Tools [4]

  By, Sultan Al Sharif1, Mohamed Majed Al Ali, Naser Salem, Farkhund Iqbal, May El Barachi, and Omar Alfandi

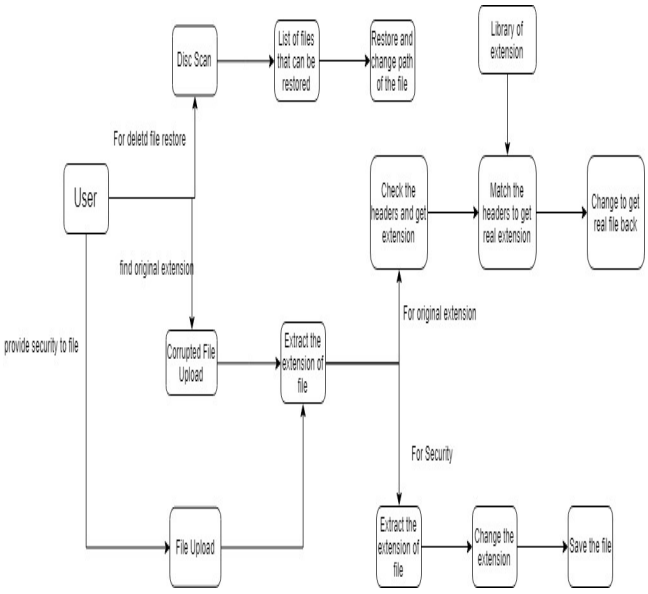
Recovering lost and deleted information from computer storage media for the purpose of forensic investigation is one of the essential steps in digital

forensics. There are several dozens of commercial and open source digital analysis tools dedicated for this purpose. The challenge is to identify the tool that best fits in a specific case of investigation. To measure the file recovering functionality, we have developed a validation approach for comparing five popular forensic tools: Encase, Recover my files, Recover, Blade, and FTK. These tools were examined in a fixed scenario to show the differences and capabilities in recovering files after deletion, quick format and full format of a USB stick. Experimental results on selected commercial and open source tools demonstrate effectiveness of proposed approach. By getting all this information one better software can be created.

V. PROPOSED SOLUTION

In our project, we propose a system m that would allow the user to perform different tasks such as identifying the file extensions, recovering deleted data from the storage media and encrypting any file that needs to be secured by the user.

The Block Diagram for our proposed solution is as



We aim to build a forensic tool that would help the investigators and ease their task by automating most of the processes in the system. This would also save some of their time as they would just have to select files to be identified, restored or encrypted.

- Hex View:

The name 'hex' comes from 'hexadecimal': a standard numerical format for representing binary data. A typical computer file occupies multiple areas on the platter(s) of a disk drive, whose contents are combined to form the file. Hex editors that are designed to parse and edit sector data from the physical segments of floppy or hard disks are sometimes called sector editors or disk editors.

A user can see or edit the raw and exact contents of a file, as opposed to the interpretation of the same content that other, higher level application software may associate with the file format. For example, this could be raw image data, in contrast to the way image editing software would interpret and show the same file. Hex view can also interpret and view contents of documents (.txt, .docx etc.). The data of the computer file is represented as hexadecimal values grouped in 4 groups of 4 bytes (or two groups of 8 bytes), followed by one group of 16 printable ASCII characters which correspond to each pair of hex values (each byte). Non-printable ASCII characters (e.g., Bell) and characters that would take more than one-character space (e.g., tab) are typically represented by a dot (".") in the following ASCII field.

- To Find Hidden Data:

One method is the file deletion technique, here when a file is deleted; the record of the file is removed from the table so it appears, as the file no longer exists. Although the record is gone, the file may still exist on the cluster. Partition deletion is similar to file deletion where the record may be gone but the file still resides on the hard disk.

A file extension or file name extension is the ending of a file that helps identify the type of file in operating systems, such as Microsoft Windows. In Microsoft Windows, the file name extension is a period that is often followed by three characters, but may also be one, two, or four characters long.

Sometimes these extensions can be changed by a user to rename it. In such scenario the os will identify the file and will try to read or execute that file in other program. Due to this the file won't open and user

might think it is corrupted. The context of the file remains the same.

In such cases, we have to manually check the hex value of the file in order to verify the file type. Hex values of the files are used as they store all the information of the file including the file type information, which shows the extension, in which the file should be stored.

- Hidden File Extension:

In such cases, the investigators have to manually check the hex value of the file in order to verify the file type. Hex values of the files are used as they store all the information of the file including the file type information, which shows the extension, in which the file should be stored.

- File Restoration:

Files can be recovered by calculating the starting and the end of the file in hex format and copying it into a text file and saving it into the appropriate extension. The common problem with deleting files is accidental removal of information that later proves to be important. One way to deal with this is to back up files regularly. Erroneously deleted files may then be found in archives, but what if the file was not backed up? Although many times deleted file are present in recycle bin but we may also delete the file permanently which cannot be recovered in recycle bin or from temp folder.

Operating system uses various File system such as FAT 16, NTFS etc. to store and read data. These file systems manage a Storage table which uses indexing to store and point towards the next address. When file is deleted its data is not deleted. Instead its entry from file table is wiped out. We can use enumeration and base address to calculate hidden address and restore the files. File can only be restored as long as they aren't overwritten by another file. In such case it becomes very difficult to capture and restore such files.

- 1.3.4 Data Security:

Securing or hiding the data of the user using the methods of the attacker itself. These methods consist of data hiding in the hidden compartments of another file, changing the extension of the file to be secured.

Flow Chart and the Screenshots of the solution

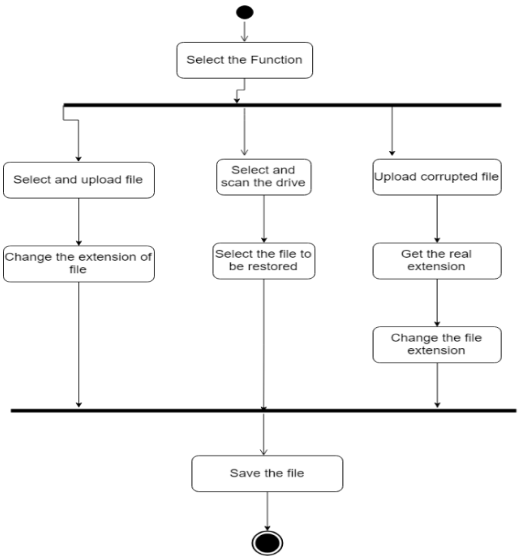
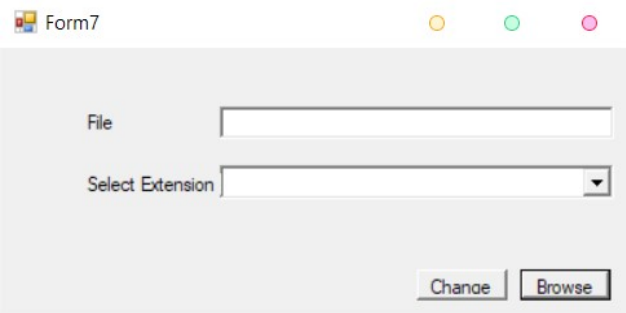


Chart:



VI. ACKNOWLEDGEMENT

We wish to express our profound gratitude to our principal Dr. Bhavesh Patel for allowing us to go ahead with this project and giving us the opportunity to explore this domain. We would also like to thank our Head of Department Prof. Uday Bhave for our constant encouragement and support towards achieving this goal. We would also like to thank the Review Committee for their invaluable suggestions and feedback without whom our work would have been very difficult. We take this opportunity to express our profound gratitude and deep regards to our guide Prof. Sonali Bhutad for his exemplary guidance, monitoring and constant encouragement throughout the course of this project. The blessing, help and guidance given by him from time to time shall carry us a long way in the journey of life on which we are about to embark. No project is ever complete without the guidelines of these experts who have already established a mark on this path before and have become masters of it. So, we would like to take this opportunity to thank all those who have helped us in implementing this project.

REFERENCES

[1] Research Journal "Digital Forensic Investigation on File System and Database Tampering, By Sindhu. K. K, Shweta Tripathi, Dr. B. B. Meshram",

[2] Book "Computer Hacking Forensic Investigator, By EC-Council",

[3] Research Paper "A Comparative Study of File-Type Identification Techniques By William H. Allen, Nasser S. Alamri",

[4] Research paper "An Approach for the Validation of File Recovery Functions in Digital Forensics' Software Tools, By Sultan Al Sharif1, Mohamed Majed Al Ali, Naser Salem, Farkhund Iqbal, May El Barachi, and Omar Alfandi".