INT301 CA-3

Open-Source Technologies

A Training Report

Submitted in partial fulfillment of the requirements for the award of degree of

Bachelor of Technology (Computer Science and Engineering)

Submitted to

LOVELY PROFESSIONAL UNIVERSITY

PHAGWARA, PUNJAB

From 31/03/2023 to 10/04/2023

Submitted By-

Lavish Gupta (11917631)

Roll No: - 73

Section: - KE023
B. Tech Computer Science and Engineering
Lovely Professional University, Phagwara

Introduction

In the field of digital forensics investigations, it is essential to have tools that can extract data from various sources. One such tool is Bulk Extractor, which is a command-line tool that can extract data from disk images, network traffic captures, and memory dumps. This report will discuss the features of Bulk Extractor and its applications in digital forensics investigations.

Features of Bulk Extractor

Bulk Extractor is a powerful tool that can extract data from various sources. The tool supports multiple scanners, which can be customized to extract specific types of data. Some of the key features of Bulk Extractor are:

Disk Imaging:
Bulk Extractor can extract data from disk images, which are digital copies of a storage device. The tool can scan the disk image for various types of data, such as email addresses, credit card numbers, and phone numbers.

Network Traffic Analysis:
Bulk Extractor can analyse network traffic captures, which are records of data sent and received over a network. The tool can extract data from the captures, such as URLs, email addresses, and credit card numbers.

Memory Analysis:
Bulk Extractor can analyse memory dumps, which are copies of a computer's RAM. The tool can extract data from the dumps, such as passwords, encryption keys, and malware.

Customizable Scanners:
Bulk Extractor supports multiple scanners, which can be customized to extract specific types of data. The scanners can be configured to search for specific patterns, such as email addresses, phone numbers, and credit card numbers.

Applications of Bulk Extractor

Bulk Extractor has various applications in digital forensics investigations. Some of the key applications of the tool are:

Recovering Deleted Data:
Bulk Extractor can recover deleted data from disk images. The tool can scan the disk image for unallocated space and try to recover any deleted files. This feature is useful in cases where important data has been accidentally deleted.

Password Cracking:
Bulk Extractor can create a wordlist for cracking encryption. The tool can generate a JSON file containing all the words found on the disk image. This file can then be used as a password dictionary for cracking encryption.

Malware Analysis:
Bulk Extractor can analyse memory dumps for malware. The tool can extract data from the dumps, such as malware signatures and encryption keys. This feature is useful in cases where malware has been used to compromise a system.

Incident Response:
Bulk Extractor can be used in incident response investigations. The tool can extract data from disk images and network traffic captures to determine the scope of an incident. This information can be used to identify the source of the incident and prevent future attacks.

Steps for the Task Assigned

Obtain a Disk Image of the Target System or Device:
The first step is to obtain a disk image of the target system or device. This can be done using various tools, such as FTK Imager or dd.

Install and Configure Bulk Extractor on a Forensic Workstation:
The next step is to install and configure Bulk Extractor on a forensic workstation. The tool can be downloaded from the official website, and the installation process is straightforward. Once installed, the tool can be configured by modifying the configuration file, which is in the Bulk Extractor directory.

Extract Data from the Disk Image:
Once the tool is installed and configured, the next step is to extract data from the disk image. This can be done by running the following command in the command prompt:

```
bulk_extractor -o output_folder -R disk_image_file
```

This command will extract data from the disk image and store it in the specified output folder. The -R option tells Bulk Extractor to recursively search for data in the disk image.

Locate Potentially Deleted Emails:
To locate potentially deleted emails, use the following command:

```
bulk_extractor -o output_folder -R -E email disk_image_file
```

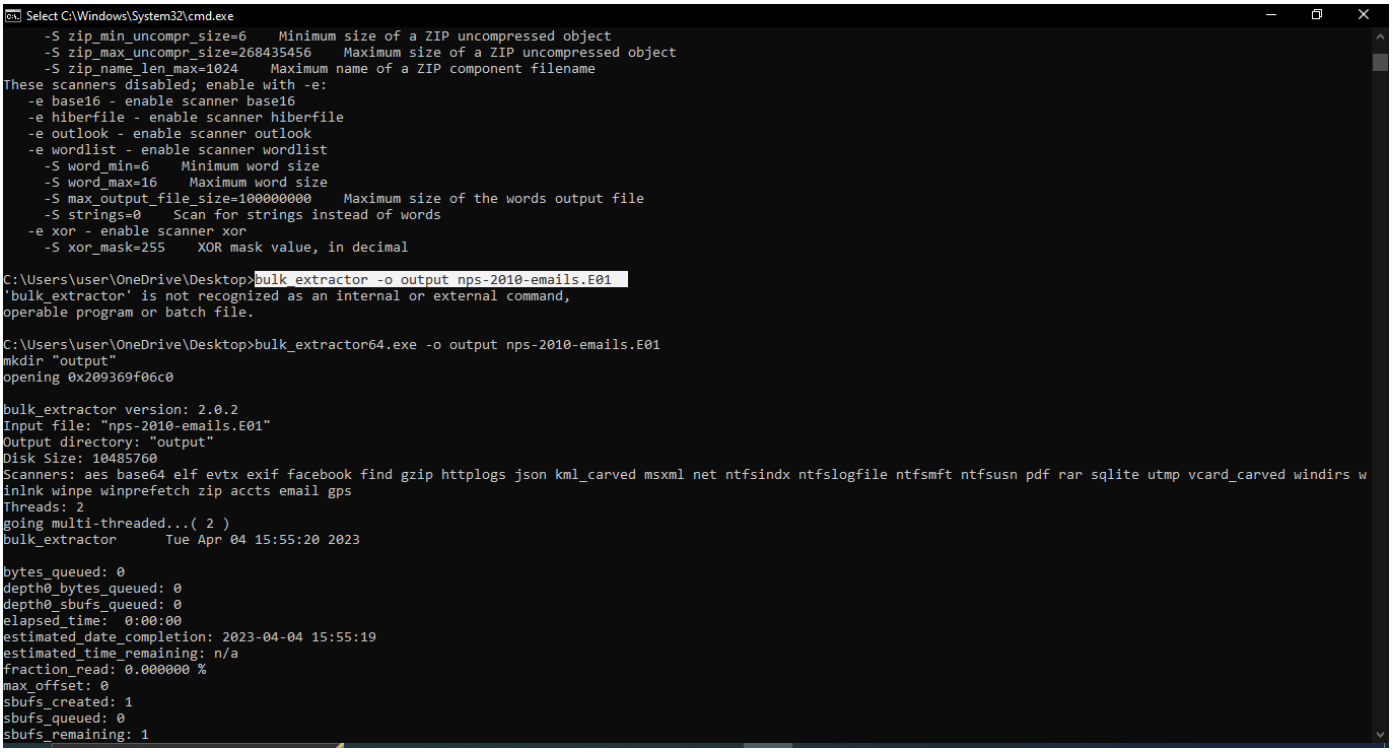
This command will extract all email addresses from the disk image, including those that may have been deleted. The -E option tells Bulk Extractor to extract only email addresses, and the output will be stored in the specified output folder.

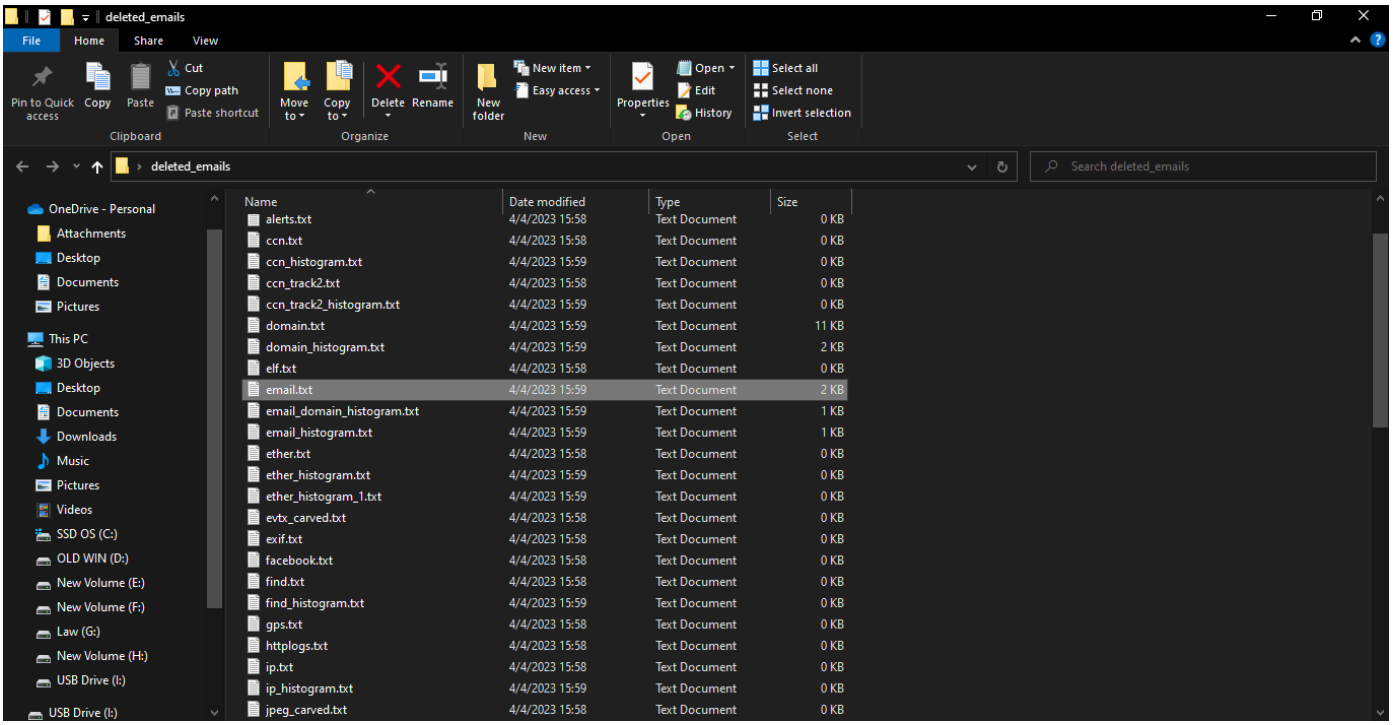
Create a Wordlist for Password Cracking:
To create a wordlist for password cracking, use the following command:

```
bulk_extractor -o output_folder -R -e wordlist disk_image_file
```

This command will extract all words from the disk image and store them in a JSON file. This file can be used as a password dictionary for cracking encryption.

Images of the task

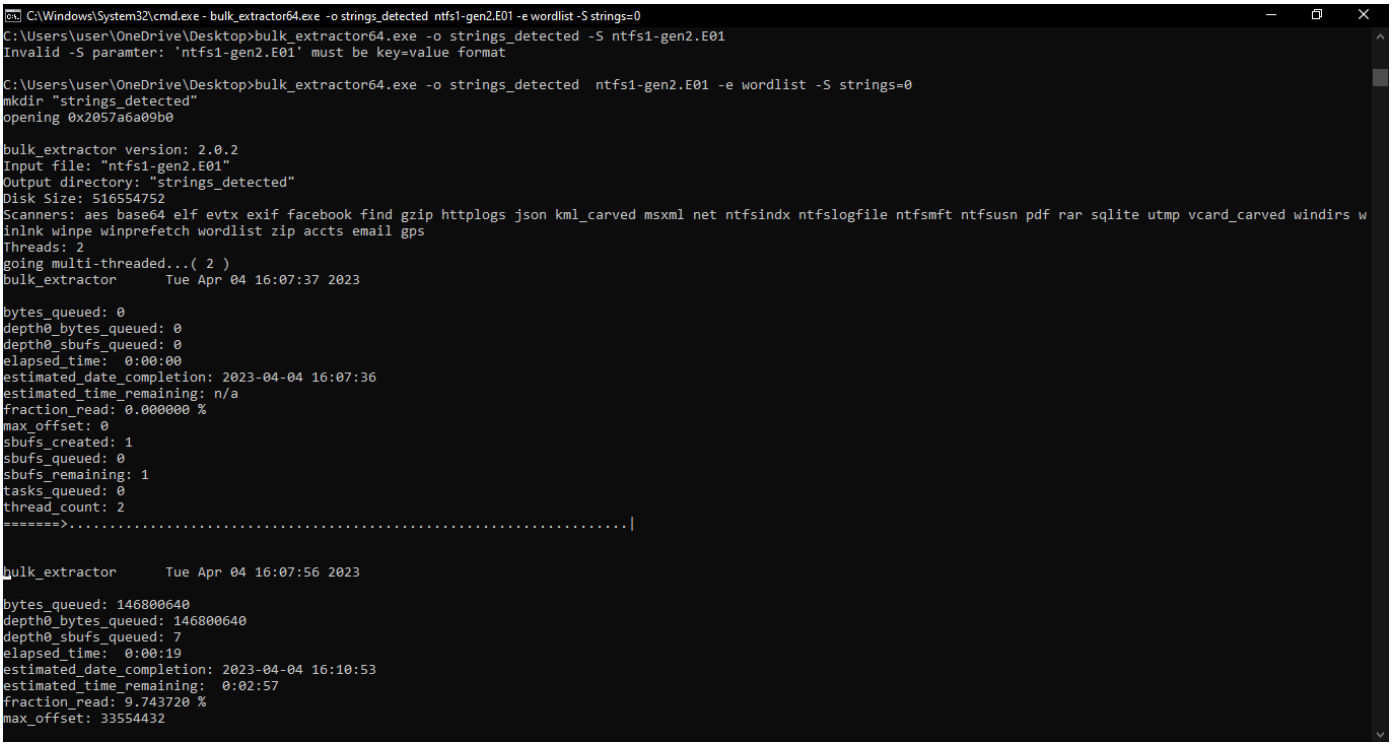


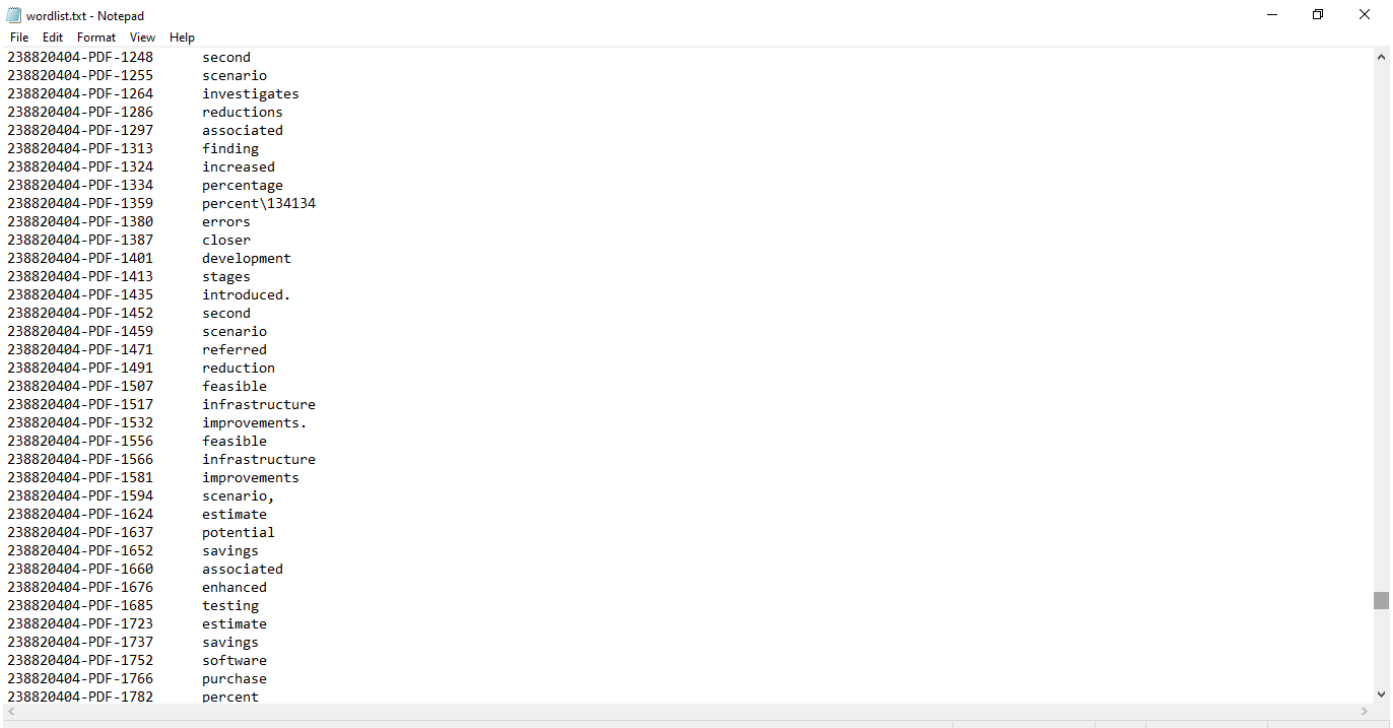


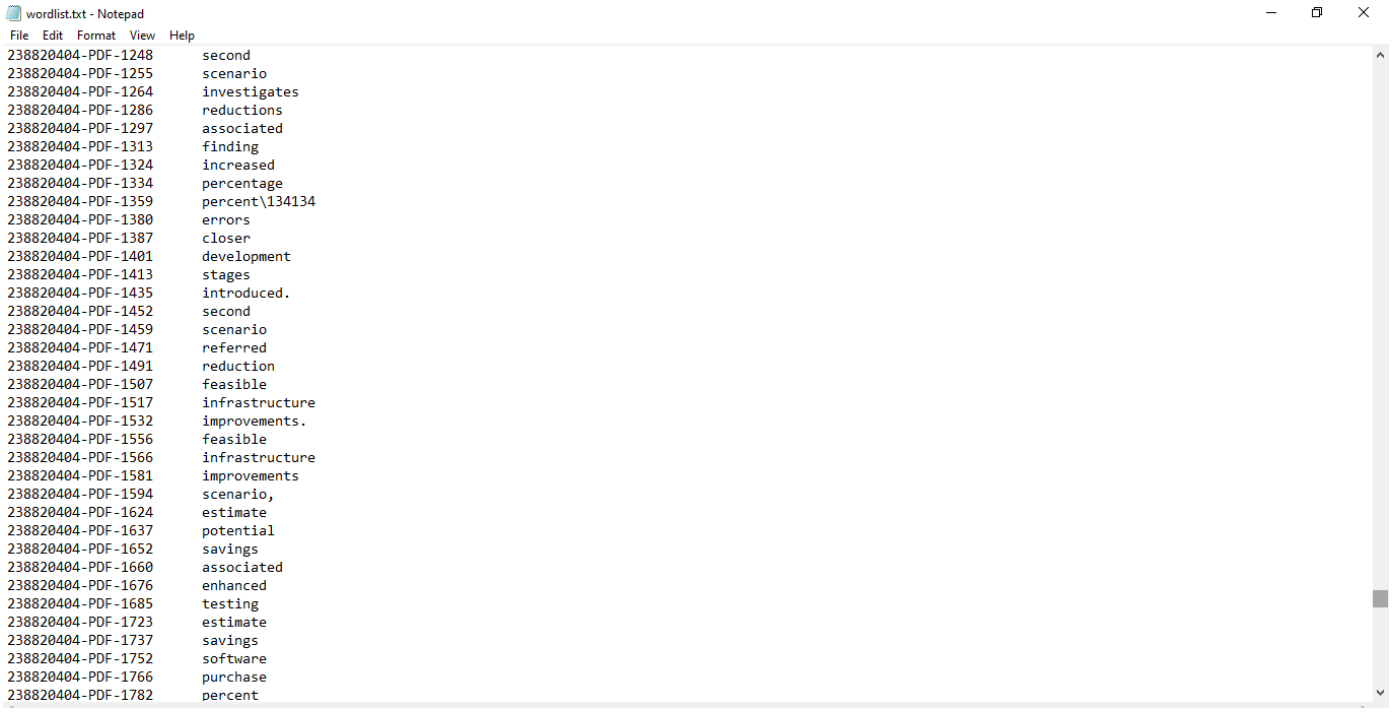
email.txt - Notepad

File Edit Format View Help

```
# BANNER FILE NOT PROVIDED (-b option)
# BULK_EXTRACTOR-Version: 2.0.2
# Feature-Recorder: email
# Filename: ntfs1-gen2.E01
# Feature-File-Version: 1.1
47346547        xL6@3ov.KE      ^g\214\007t\344g'DF\230\230\364F\264\204xL6@3ov.KE\343\341#\275\231\014\357&\213&u\226\203\350&\227
239767510-PDF-392       btinfo@bottomline.com    .bottomline.com btinfo@bottomline.com  Paybase Allows
239767510-PDF-1681      info@entegrity.com      w.entegrity.com info@entegrity.com Equifax E-Banki
239772263-PDF-1084      paul.wrenn@hicor.net    4 www.hicor.net paul.wrenn@hicor.net Mellon Global C
239772263-PDF-1296      gcm_direct_pgh@mellon.com       n.com/inst/gcm/ gcm_direct_pgh@mellon.com National City C
239772263-PDF-1673      mark_d_schulte@national-city.com        14 216-222-3633 mark_d_schulte@national-city.com \134(continued\134
239776426-PDF-437       gigiw@paytec.com        55 717-506-2200 gigiw@paytec.com Politzer & Hane
239776426-PDF-1747      ed_armstrong@stercomm.com       ingcommerce.com ed_armstrong@stercomm.com \134(continued\134
239780882-PDF-812       patricia.engelage@umb.com       045 www.umb.com patricia.engelage@umb.com
```


C:\Windows\System32\cmd.exe - bulk_extractor64.exe -o strings_detected ntfs1-gen2.E01 -e wordlist -S strings=0
C:\Users\user\OneDrive\Desktop>bulk_extractor64.exe -o strings_detected -S ntfs1-gen2.E01
Invalid -S paramter: 'ntfs1-gen2.E01' must be key=value format
C:\Users\user\OneDrive\Desktop>bulk_extractor64.exe -o strings_detected ntfs1-gen2.E01 -e wordlist -S strings=0
mkdir "strings_detected"
opening 0x2057a6a09b0
bulk_extractor version: 2.0.2
Input file: "ntfs1-gen2.E01"
Output directory: "strings_detected"
Disk Size: 516554752
Scanners: aes base64 elf evtx exif facebook find gzip httplogs json kml_carved msxml net ntfsindx ntfslogfile ntfsmft ntfsusn pdf rar sqlite utmp vcard_carved windirs w
inlnk winpe winprefetch wordlist zip accts email gps
Threads: 2
going multi-threaded...(2)
bulk_extractor Tue Apr 04 16:07:37 2023
bytes_queued: 0
depth0_bytes_queued: 0
depth0_sbufs_queued: 0
elapsed_time: 0:00:00
estimated_date_completion: 2023-04-04 16:07:36
estimated_time_remaining: n/a
fraction_read: 0.000000 %
max_offset: 0
sbufs_created: 1
sbufs_queued: 0
sbufs_remaining: 1
tasks_queued: 0
thread_count: 2
=======>.....|
bulk_extractor Tue Apr 04 16:07:56 2023
bytes_queued: 146800640
depth0_bytes_queued: 146800640
depth0_sbufs_queued: 7
elapsed_time: 0:00:19
estimated_date_completion: 2023-04-04 16:10:53
estimated_time_remaining: 0:02:57
fraction_read: 9.743720 %
max_offset: 33554432


wordlist.txt - Notepad
File Edit Format View Help
238820404-PDF-1248 second
238820404-PDF-1255 scenario
238820404-PDF-1264 investigates
238820404-PDF-1286 reductions
238820404-PDF-1297 associated
238820404-PDF-1313 finding
238820404-PDF-1324 increased
238820404-PDF-1334 percentage
238820404-PDF-1359 percent\134134
238820404-PDF-1380 errors
238820404-PDF-1387 closer
238820404-PDF-1401 development
238820404-PDF-1413 stages
238820404-PDF-1435 introduced.
238820404-PDF-1452 second
238820404-PDF-1459 scenario
238820404-PDF-1471 referred
238820404-PDF-1491 reduction
238820404-PDF-1507 feasible
238820404-PDF-1517 infrastructure
238820404-PDF-1532 improvements.
238820404-PDF-1556 feasible
238820404-PDF-1566 infrastructure
238820404-PDF-1581 improvements
238820404-PDF-1594 scenario,
238820404-PDF-1624 estimate
238820404-PDF-1637 potential
238820404-PDF-1652 savings
238820404-PDF-1660 associated
238820404-PDF-1676 enhanced
238820404-PDF-1685 testing
238820404-PDF-1723 estimate
238820404-PDF-1737 savings
238820404-PDF-1752 software
238820404-PDF-1766 purchase
238820404-PDF-1782 percent



Conclusion

Bulk Extractor is a powerful tool that can be used in digital forensics investigations. The tool can extract data from various sources, such as disk images, network traffic captures, and memory dumps. Bulk Extractor has various applications, such as recovering deleted data, password cracking, malware analysis, and incident response. In this report, we discussed the steps involved in using Bulk Extractor to potentially locate deleted emails and scan a disk for text strings to use as a password dictionary to crack encryption. These steps can be used as a starting point for digital forensics investigations involving Bulk Extractor.

Link of GitHub Repository

https://github.com/imlavish1012/Open-Source-Project