



Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

Experiment No.8
To use bulk_extractor tool to detect Evidence related to email
Date of Performance:
Date of Submission:



Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

Aim: To use bulk_extractor tool to detect Evidence related to email

Objective: To make use of the bulk_extractor tool to recover email related evidence from a disk image

Theory:

Bulk_extractor is a C++ program that scans a disk image, a file, or a directory of files and extracts useful information without parsing the file system or file system structures. The results are stored in feature files that can be easily inspected, parsed, or processed with automated tools. bulk_extractor also creates histograms of features that it finds, as features that are more common tend to be more important.

Bulk Email Extractor is also an email client manager. The software provides a streamlined extraction of email. It supports every online email service on the internet. This helps users process all email addresses in an efficient and automated manner.

Bulk Email Extractor is created to handle tons of emails addresses. Collecting email addresses is a monumental job for anyone. Piled-up messages from clients and accounts can be a good source of email addresses. Some of them could be not existing or fake and most users don't have the luxury of checking individual email addresses. It will take several hours until all email addresses are collected and sorted out. This software is capable of extracting large amounts of email addresses. It runs within minutes compared to the manual opening of contact addresses. Any online email services or business directory websites are supported by this software and it even supports websites with email listings.

```
Average time each consumer spent waiting for data from producer: 0:00:00 (seconds)
*** More time spent waiting for workers. You need a faster CPU or more cores improved performance.
Total email features found: 1618

(kali@kali)-[~]
$ bulk_extractor -o test-case1 2020JimmyWilson.E01
```

Fig. 8.1 View of Bulk_Extractor



Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

Process:

Step 1, open the Bulk_extractor tool in kali Linux

Step 2. create or download the mirror image of the hard disk

Step 3. use the command as shown in figure 8.1

Step 4. open the output file to see email evidence

Output:

```
root@kali: /home/kali
bulk_extractor -o test-case1 ntfs1-gen2.E01
mkdir "test-case1"
opening ntfs1-gen2.E01

bulk_extractor version: 2.1.1
Input file: "ntfs1-gen2.E01"
Output directory: "test-case1"
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 zip accts email gps
Threads: 2
going multi-threaded...( 2 )
bulk_extractor Mon Mar 24 02:43:54 2025

available_memory: 4921139200
bytes_queued: 0
depth0_bytes_queued: 0
depth0_sbufs_queued: 0
elapsed_time: 0:00:00
estimated_date_completion: 2025-03-24 02:43:53
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 Mon Mar 24 02:43:55 2025

available_memory: 4874383360
bytes_queued: 104857600
depth0_bytes_queued: 104857600
depth0_sbufs_queued: 1
```



```
bulk_extractor Mon Mar 24 02:44:16 2025

available_memory: 4823572480
bytes_queued: 52953088
depth0_bytes_queued: 52953088
depth0_shufs_queued: 4
elapsed_time: 0:00:22
estimated_date_completion: 2025-03-24 02:44:16
estimated_time_remaining: 0:00:00
fraction_read: 100.000000 %
max_offset: 503316480
shufs_created: 866936
shufs_queued: 4
shufs_remaining: 1
tasks_queued: 2
thread-1: 503316480: accts (13238272 bytes)
thread-2: 503316480: net (13238272 bytes)
thread_count: 2
=====|
Intercept and store traffic on a network segment,
captures passwords,
conducts active eavesdropping against a number of common protocols:
Phase 2. Shutting down scanners
Computing final histograms and shutting down... (all and easy to use) filtering language that allows for custom
Phase 3. Generating stats and printing final usage information
All Threads Finished!
Elapsed time: 23.19 sec.
Total MB processed: 516
Overall performance: 22.28 MBytes/sec 11.14 (MBytes/sec/thread)
shufs created: 866936
shufs unaccounted: 0
Time producer spent waiting for scanners to process data: 0:00:16 (16.69 seconds)
Time consumer scanners spent waiting for data from producer: 0:00:00 (0.63 seconds)
Average time each consumer spent waiting for data from producer: 0:00:00 (0.00 seconds)
*** More time spent waiting for workers. You need a faster CPU or more cores for improved performance.
Total email features found: 9

(root@kali)-[/home/kali]
```

```
# BANNER FILE NOT PROVIDED (-b option)
# BULK_EXTRACTOR-Version: 2.1.1
# Feature-Recorder: email
# Filename: ntfsl-gen2.E01
# Feature-File-Version: 1.1
47346547 xL683ov.KE ^g\214\007t\344g'DF\230\230\364F\264\204xL683ov.KE\343\341#\275\231\014\357s\213su\226\203\350s\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@hicolor.net 4 www.hicolor.net paul.wrenn@hicolor.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
```

Conclusion:

The Bulk_extractor tool is highly significant in digital forensic investigations, especially for extracting email-related evidence. Its ability to scan raw disk images without relying on file system structures allows investigators to recover hidden or deleted email addresses and communication traces efficiently. By automatically parsing through large volumes of data and generating feature-rich output files, Bulk_extractor simplifies the process of identifying communication patterns, contacts, and potential leads. This makes it a valuable asset in cases involving email fraud, identity theft, cyberstalking, or any scenario where email communication plays a crucial role.