

**TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**

**ĐỒ ÁN TỐT NGHIỆP**  
**Phần mềm thu thập chứng cứ sự cố**  
**an toàn thông tin trên máy tính sử dụng**  
**hệ điều hành Linux**

**TRẦN QUANG HUY**

`huy.tq173180@sis.hust.edu.vn`

**Ngành Mạng và Truyền thông dữ liệu**  
**Chuyên ngành Kỹ thuật máy tính**

**Giảng viên hướng dẫn:** PGS. TS. Nguyễn Linh Giang \_\_\_\_\_  
Chữ ký của GVHD

**Bộ môn:** Truyền thông và Mạng máy tính  
**Viện:** Công nghệ thông tin - Truyền thông

**HÀ NỘI, 12/2021**

## **ĐỀ TÀI TỐT NGHIỆP**

Phần mềm thu thập chứng cứ sự cố an toàn thông tin trên máy tính sử dụng hệ điều hành Linux.

Giáo viên hướng dẫn  
Ký và ghi rõ họ tên

### **Lời cảm ơn**

Trước hết, tôi xin tỏ lòng kính trọng và biết ơn sâu sắc đối với thầy giáo - PGS.TS. Nguyễn Linh Giang, giảng viên viện Công nghệ Thông tin và Truyền thông, trường Đại học Bách khoa Hà Nội, người đã trực tiếp hướng dẫn tôi trong suốt thời gian thực hiện đề tài. Mặc dù thời gian có hạn và gặp nhiều rào cản, khó khăn bởi tình hình đại dịch COVID-19, thầy vẫn hết sức tận tình hướng dẫn và giúp đỡ tôi trong quá trình hoàn thiện đề tài.

Với tình cảm sâu sắc và chân thành nhất, tôi xin được bày tỏ lòng biết ơn quý thầy cô thuộc viện Công nghệ Thông tin và Truyền thông, trường Đại học Bách khoa Hà Nội, những người đã truyền đạt vốn kiến thức quý báu cho tôi trong suốt thời gian tham gia học tập tại trường. Đề tài của tôi có thể hoàn thiện như ngày hôm nay hoàn toàn là nhờ những lời hướng dẫn, dạy bảo của thầy cô.

Với điều kiện thời gian có hạn cũng như kinh nghiệm còn hạn chế, đề tài chắc chắn không thể tránh được nhiều thiếu sót. Tôi rất mong được tiếp thu những chỉ bảo, ý kiến đóng góp của quý thầy cô, để tôi có điều kiện bổ sung và phát triển đề tài của mình trong tương lai.

Tôi xin chân thành cảm ơn !

### **Tóm tắt nội dung đồ án**

Khi hệ thống thông tin của một doanh nghiệp, tổ chức bị tấn công hoặc gặp sự cố an toàn thông tin, tất cả các thiết bị, từ server, máy tính cá nhân hay cơ sở hạ tầng mạng đều có thể cung cấp những dấu vết kỹ thuật số quan trọng. Những dấu vết này đóng vai trò quan trọng trong việc tái tạo lại các sự kiện đã xảy ra, hoặc cung cấp bằng chứng về hành vi sai trái đã được thực hiện. Tuy nhiên, lượng thông tin này không hề nhỏ, và việc thu thập đầy đủ cũng như đảm bảo tính toàn vẹn của chúng là một thách thức lớn.

Hiểu được điều này, tôi đã nảy ra ý tưởng thực hiện đề tài, xây dựng một phần mềm đáp ứng những vấn đề khó khăn còn tồn tại trong việc thu thập chứng cứ từ các máy tính gặp sự cố an toàn thông tin. Sau khi đã đạt được những kết quả nhất định từ phần mềm thu thập chứng cứ trên Windows trong đề tài trước, lần này, tôi quyết định hướng tới thực hiện phát triển phần mềm trên nền tảng Linux cùng chức năng trên. Phần mềm sẽ thực hiện chức năng thu thập các chứng cứ liên quan đến hệ thống, kết nối mạng, vết sự kiện và tài khoản người dùng trên máy tính cần điều tra, đồng thời đảm bảo tính toàn vẹn của những chứng cứ thu thập được. Quá trình thực hiện đồ án tốt nghiệp này đã giúp tôi mở rộng thêm vốn hiểu biết về lĩnh vực thu thập chứng cứ trên máy tính gặp sự cố, và kết quả đồ án cũng hứa hẹn sẽ giải quyết được nhiều vấn đề còn tồn tại trong lĩnh vực này.

## MỤC LỤC

<b>CHƯƠNG 1. GIỚI THIỆU ĐỀ TÀI.....</b>	<b>1</b>
1.1 Đặt vấn đề .....	1
1.2 Mục tiêu và phạm vi đề tài.....	1
1.3 Định hướng giải pháp.....	2
1.4 Bố cục đồ án.....	3
<b>CHƯƠNG 2. KHẢO SÁT VÀ PHÂN TÍCH YÊU CẦU .....</b>	<b>4</b>
2.1 Khảo sát hiện trạng.....	4
2.2 Tổng quan chức năng .....	5
2.2.1 Biểu đồ use case tổng quan .....	5
2.2.2 Biểu đồ use case phân rã tạo thư mục chứa chứng cứ được thu thập	6
2.2.3 Biểu đồ use case phân rã thu thập chứng cứ hệ thống .....	7
2.2.4 Biểu đồ use case phân rã thu thập chứng cứ mạng .....	7
2.2.5 Biểu đồ use case phân rã thu thập vết sự kiện .....	7
2.2.6 Biểu đồ use case phân rã thu thập thông tin tài khoản người dùng	7
2.2.7 Biểu đồ use case phân rã cấu hình đảm bảo tính toàn vẹn chứng cứ thu thập	8
2.2.8 Quy trình nghiệp vụ .....	8
2.3 Đặc tả chức năng.....	9
2.3.1 Đặc tả use case tạo thư mục chứa chứng cứ được thu thập .....	9
2.3.2 Đặc tả use case thu thập chứng cứ hệ thống .....	10
2.3.3 Đặc tả use case thu thập chứng cứ mạng .....	10
2.3.4 Đặc tả use case thu thập vết sự kiện.....	11
2.3.5 Đặc tả use case thu thập thông tin tài khoản người dùng .....	11
2.3.6 Đặc tả use case đảm bảo tính toàn vẹn chứng cứ thu thập .....	12
2.4 Yêu cầu phi chức năng.....	12
<b>CHƯƠNG 3. CÔNG NGHỆ SỬ DỤNG .....</b>	<b>13</b>
3.1 Cơ sở lý thuyết .....	13
3.1.1 Quy chuẩn pháp y kỹ thuật số.....	13
3.1.2 Thu thập chứng cứ số.....	13
3.2 Công nghệ sử dụng.....	13
3.2.1 PyInstaller .....	13

3.2.2	Các thư viện, module Python sử dụng .....	13
<b>CHƯƠNG 4. PHÁT TRIỂN VÀ TRIỂN KHAI ỨNG DỤNG .....</b>		<b>15</b>
4.1	Xây dựng ứng dụng .....	15
4.1.1	Công cụ sử dụng .....	15
4.1.2	Yêu cầu chức năng .....	15
4.1.3	Lập trình xây dựng ứng dụng .....	16
4.1.1	Thiết kế giao diện .....	20
4.1.2	Thiết kế cấu trúc thư mục kết quả .....	21
4.1.3	Kết quả đạt được .....	22
4.1.4	Minh họa các chức năng chính .....	22
4.2	Kiểm thử .....	29
4.2.1	Kiểm thử chức năng thu thập tất cả chứng cứ .....	29
4.2.2	Kiểm thử chức năng đảm bảo tính toàn vẹn của chứng cứ .....	29
4.2.3	Kiểm thử phần mềm trên hệ điều hành Kali Linux .....	30
4.3	Triển khai .....	30
<b>CHƯƠNG 5. CÁC GIẢI PHÁP VÀ ĐÓNG GÓP NỔI BẬT .....</b>		<b>32</b>
5.1	Thu thập hầu hết chứng cứ trên máy tính cần điều tra .....	32
5.1.1	Chứng cứ hệ thống .....	32
5.1.2	Chứng cứ mạng .....	33
5.1.3	Vết sự kiện .....	33
5.1.4	Thông tin tài khoản người dùng .....	33
5.2	Phát triển những chức năng đặc trưng .....	34
5.2.1	Thu thập hàm băm của các tệp tin trong hệ thống .....	34
5.2.2	Thu thập lịch sử trình duyệt và xuất kết quả trực quan .....	34
5.2.3	Kiểm tra tính toàn vẹn của thông tin tài khoản người dùng có trong hệ thống .....	34
5.3	Xây dựng một phần mềm nhỏ gọn, tương thích với nhiều bản phân phối Linux .....	35
<b>CHƯƠNG 6. KẾT LUẬN .....</b>		<b>37</b>
6.1	Kết luận .....	37
6.2	Hướng phát triển của đề án trong tương lai .....	37
<b>TÀI LIỆU THAM KHẢO .....</b>		<b>38</b>

## DANH MỤC HÌNH VẼ

Hình 2-1 Biểu đồ use case tổng quan.....	6
Hình 2-2 Biểu đồ use case phân rã tạo thư mục chứa chứng cứ được thu thập.....	6
Hình 2-3 Biểu đồ use case phân rã thu thập chứng cứ hệ thống.....	7
Hình 2-4 Biểu đồ use case phân rã thu thập chứng cứ mạng.....	7
Hình 2-5 Biểu đồ use case phân rã thu thập vết sự kiện .....	7
Hình 2-6 Biểu đồ use case phân rã thu thập thông tin người dùng.....	7
Hình 2-7 Biểu đồ use case phân rã cấu hình đảm bảo tính toàn vẹn chứng cứ thu thập .....	8
Hình 2-8 Biểu đồ quy trình nghiệp vụ .....	8
Hình 4-1 Hàm thu thập lịch sử sử dụng dòng lệnh .....	17
Hình 4-2 Tạo thư mục lưu trữ kết quả .....	18
Hình 4-3 Tìm kiếm tệp tin cơ sở dữ liệu lịch sử trình duyệt trên Chrome .....	18
Hình 4-4 Đọc lịch sử trình duyệt Chrome từ cơ sở dữ liệu SQLite .....	19
Hình 4-5 Thực hiện nén thư mục lưu trữ kết quả.....	19
Hình 4-6 Xuất mã hóa SHA256 của tệp tin được nén .....	20
Hình 4-7 Minh họa giao diện phần mềm .....	21
Hình 4-8 Minh họa cấu trúc thư mục lưu trữ kết quả .....	22
Hình 4-9 Khởi động phần mềm.....	23
Hình 4-10 Giao diện lựa chọn đường dẫn chứa thư mục kết quả .....	23
Hình 4-11 Giao diện hiển thị chức năng .....	24
Hình 4-12 Chọn chức năng thu thập thông tin hệ thống .....	24
Hình 4-13 Thông báo thu thập thành công .....	25
Hình 4-14 Kết quả thu thập thông tin hệ thống.....	25
Hình 4-15 Màn hình tiếp tục thu thập .....	25
Hình 4-16 Chọn đồng thời nhiều chức năng.....	26
Hình 4-17 Thông báo thu thập thành công .....	26
Hình 4-18 Thư mục kết quả .....	27
Hình 4-19 Thư mục chứng cứ mạng .....	27
Hình 4-20 Thư mục vết sự kiện .....	28
Hình 4-21 Kết thúc phiên thu thập .....	28
Hình 4-22 Tệp tin nén chứng cứ thu thập .....	28
Hình 4-23 Thực thi câu lệnh pyinstaller .....	30
Hình 4-24 Tệp tin thực thi LinuxIncidentResponse.....	31
Hình 4-25 Khởi động phần mềm.....	31

## DANH MỤC BẢNG

Bảng 2-1 Một số phần mềm thu thập chứng cứ trên Linux .....	4
Bảng 2-2 Ưu nhược điểm của những phần mềm thu thập chứng cứ trên Linux ...	5
Bảng 2-3 Đặc tả use case tạo thư mục chứa chứng cứ được thu thập.....	9
Bảng 2-4 Đặc tả use case thu thập chứng cứ hệ thống.....	10
Bảng 2-5 Đặc tả use case thu thập chứng cứ mạng.....	10
Bảng 2-6 Đặc tả use case thu thập vết sự kiện.....	11
Bảng 2-7 Đặc tả use case thu thập thông tin tài khoản người dùng.....	11
Bảng 2-8 Đặc tả use case đảm bảo tính toàn vẹn chứng cứ thu thập.....	12
Bảng 4-1 Danh sách công cụ sử dụng.....	15
Bảng 4-2 Chi tiết các chức năng cần thực hiện.....	16
Bảng 4-3 Ý nghĩa màu sắc được sử dụng .....	21
Bảng 4-4 Mô tả thông tin sản phẩm.....	22
Bảng 4-5 Kiểm thử chức năng thu thập tất cả chứng cứ.....	29
Bảng 4-6 Kiểm thử chức năng đảm bảo tính toàn vẹn của chứng cứ .....	29
Bảng 4-7 Kiểm thử phần mềm trên hệ điều hành Kali Linux.....	30

# CHƯƠNG 1. GIỚI THIỆU ĐỀ TÀI

## 1.1 Đặt vấn đề

Trong thế kỉ XXI ngày này, việc đi vào một hiệu sách, đọc một tờ báo hay một cuốn tạp chí mà không nhìn hoặc nghe thấy những chủ đề về Internet là điều gần như không thể. Nó phổ biến đến mức mà gần như không một mục quảng cáo nào được hoàn thiện mà không xuất hiện trên một trang web nào đó. Sự phổ biến của Internet đã kéo theo sự phụ thuộc của con người vào các thiết bị điện tử như điện thoại thông minh, máy tính bảng, máy tính cá nhân. Lượng thông tin mà nhiều người lưu trữ trên thiết bị thông minh của mình đã vô tình biến chúng trở thành mục tiêu hàng đầu của những kẻ có ý đồ xấu.

Các cuộc tấn công liên quan đến công nghệ thông tin đang ngày càng gia tăng trong mấy năm trở lại đây, từ những thiết bị thông minh đơn lẻ của từng cá nhân đến những hệ thống thông tin lớn của các tổ chức. Khi hệ thống thông tin của một doanh nghiệp, tổ chức bị tấn công hoặc gặp sự cố an toàn thông tin, tất cả các thiết bị, từ server, máy tính cá nhân hay cơ sở hạ tầng mạng đều có thể cung cấp những dấu vết kỹ thuật số quan trọng. Những dấu vết này đóng vai trò quan trọng trong việc tái tạo lại các sự kiện đã xảy ra, hoặc cung cấp bằng chứng về hành vi sai trái được thực hiện. Tuy nhiên, quy trình thu thập và bảo quản những chứng cứ này còn nhiều hạn chế, đa phần do lượng thông tin cần thu thập quá lớn, lại không có quy chuẩn quốc tế chung để giúp chuyên viên phản ứng có thể áp dụng theo khi gặp sự cố. Quy trình thu thập chứng cứ sau sự cố của đa phần các tổ chức hiện giờ đều là tự xây dựng, phối hợp nhiều phần mềm cũng như phương pháp khác nhau để tổng hợp dữ liệu. Điều này có thể dẫn tới nhiều vấn đề, sai sót trong quá trình thực hiện, dẫn đến việc dữ liệu thu thập không đầy đủ, hoặc không có giá trị sử dụng.

Giải quyết được vấn đề trên sẽ không chỉ đem lại lợi ích cho các chuyên viên pháp y kỹ thuật số, các chuyên gia bảo mật đang tác nghiệp trong lĩnh vực này, mà còn góp phần củng cố hệ thống bảo mật, an toàn thông tin trong các doanh nghiệp, tổ chức, thông qua việc xác định được một quy chuẩn chung về xử lý sự cố giúp cải thiện khả năng khắc phục và ngăn chặn các sự cố tương tự xảy ra.

## 1.2 Mục tiêu và phạm vi đề tài

Hiện nay, số lượng các bộ công cụ, phần mềm cho phép thực hiện thu thập chứng cứ từ những máy tính gặp sự cố an toàn thông tin, lại đồng thời có thể đảm bảo tính toàn vẹn của chứng cứ còn rất hạn chế. Muốn thực hiện đầy đủ một phiên thu thập phải thông qua rất nhiều công cụ, phần mềm khác nhau. Sự khác nhau giữa các định dạng kết quả, cách thức hoạt động khiến cả người thu thập lẫn người phân tích chứng cứ đều gặp rất nhiều khó khăn trong quá trình điều tra và xử lý sự cố.

Để giải quyết những vấn đề kể trên, mục tiêu của đề án này là xây dựng một phần mềm có khả năng xác định và thu thập bằng chứng từ các máy tính sử dụng nền tảng Linux trong các cơ quan, tổ chức. Tôi lựa chọn Linux là mục tiêu tiếp theo của mình, sau phần mềm thu thập chứng cứ trên Windows đã thực hiện ở đề



tài trước, bởi mặc dù Linux không phải hệ điều hành phổ biến nhất trên máy tính cá nhân, nhưng nó vẫn sở hữu vai trò riêng của mình và chiếm một vị trí quan trọng trong thời đại ngày nay. Khi mà Windows đang trên đà thống trị các máy tính cá nhân, vẫn có rất nhiều các tổ chức sử dụng Linux cho máy chủ của họ. Linux là sự lựa chọn của nhiều nhà cung cấp dịch vụ Internet và các công ty lớn như Google (họ thậm chí còn có bản phân phối Linux của riêng mình). Rất nhiều thiết bị xung quanh chúng ta đang chạy một phiên bản hệ điều hành nào đó của Linux. Nó là một công nghệ then chốt cho phép tạo ra các sản phẩm mới, và sẽ khó có thể bị thay thế trong một thời gian dài.

Phần mềm được phát triển trong đồ án cần thỏa mãn các yêu cầu sau:

- Đầu tiên, phần mềm có thể chạy được trên bất kỳ máy tính sử dụng hệ điều hành nền tảng Linux nào, kể cả khi máy tính đó sử dụng bản phân phối chỉ hỗ trợ dòng lệnh.
- Thứ hai, phần mềm này có thể thu thập và lưu trữ những chứng cứ quan trọng từ máy tính cần điều tra, bao gồm (i) chứng cứ dữ liệu hệ thống, (ii) chứng cứ mạng, (iii) vết sự kiện, và (iv) thông tin các tài khoản người dùng trong máy tính.
- Thứ ba, phần mềm cần có cơ chế đảm bảo tính toàn vẹn của bằng chứng, chống mất mát hoặc sửa đổi dữ liệu sau khi thu thập.
- Cuối cùng, phần mềm cần có giao diện thân thiện người dùng, dễ nâng cấp, sửa đổi các chức năng khi cần.

### 1.3 Định hướng giải pháp

Từ những nhiệm vụ cần giải quyết đã xác định ở mục 1.2, tôi dự định sẽ phát triển phần mềm thu thập chứng cứ sự cố an toàn thông tin trên máy tính sử dụng hệ điều hành Linux để đáp ứng những nhu cầu trên. Phần mềm sẽ được viết toàn bộ bằng ngôn ngữ lập trình Python để có thể tối thiểu hóa độ phức tạp của mã nguồn, ngoài ra cũng giúp phần mềm dễ nâng cấp và sửa đổi các chức năng khi cần. Ngoài ra, do đặc thù sử dụng của các hệ điều hành Linux, nhiều bản phân phối chỉ hỗ trợ dòng lệnh, phần mềm cũng hướng đến đối tượng người dùng là những người có kinh nghiệm sử dụng, quen thuộc với hệ điều hành Linux, nên phần mềm sẽ sử dụng giao diện dòng lệnh và tương tác với người dùng thông qua cửa sổ dòng lệnh quen thuộc của các hệ điều hành Linux.

Phần mềm sẽ được cài đặt trên một thiết bị thu thập, có thể là một USB hoặc ổ cứng di động có dung lượng lớn để có thể sẵn sàng thu thập nhiều dữ liệu chứng cứ nhất có thể. Phần mềm được xây dựng sẽ đảm nhiệm vai trò thu thập chứng cứ trực tiếp trên máy tính sử dụng hệ điều hành Linux còn khả năng khởi động được. Sau đó, người thu thập sẽ thực hiện tạo ảnh các ổ cứng có trong máy tính cần thu thập và lưu trữ trong thiết bị thu thập nhằm thực hiện các công đoạn pháp y khác sau này.

Cuối cùng, tôi muốn tổng kết lại những đóng góp chính của mình trong đồ án tốt nghiệp này: (i) xây dựng phần mềm thu thập tất cả chứng cứ cần thiết theo một quy trình xử lý sự cố được công nhận, (ii) đảm bảo tính toàn vẹn của những chứng cứ thu thập được, đủ chất lượng để có thể sử dụng trước tòa án pháp luật.

Phần mềm được xây dựng đã giải quyết được phần nào yêu cầu của bài toán đặt ra ban đầu, và với đặc điểm dễ nâng cấp, chỉnh sửa, nó sẽ có cơ hội phát triển và hoàn thiện trong tương lai.

#### **1.4 Bố cục đồ án**

Phần còn lại của báo cáo đồ án tốt nghiệp này được tổ chức như sau.

CHƯƠNG 2 sẽ thực hiện khảo sát chi tiết các phần mềm, bộ công cụ hiện có. Từ kết quả khảo sát, tôi thực hiện đánh giá, so sánh các sản phẩm trên, từ đó rút ra được những mục tiêu cần đáp ứng trên phần mềm của mình và xây dựng mô tả sơ lược về các tính năng phần mềm quan trọng cần phát triển.

Trong CHƯƠNG 3, tôi sẽ giới thiệu sơ lược về cơ sở lý thuyết của đồ án tốt nghiệp và các công nghệ được sử dụng để phát triển phần mềm.

Tiếp theo, CHƯƠNG 4 sẽ cung cấp thông tin về quá trình phát triển và triển khai ứng dụng. Trong chương này, tôi sẽ giới thiệu về công cụ sử dụng, yêu cầu chức năng, thiết kế chi tiết cùng quá trình lập trình xây dựng ứng dụng và minh họa của các chức năng chính của phần mềm. Quá trình kiểm thử và triển khai cũng sẽ được mô tả trong chương này.

CHƯƠNG 5 sẽ nói về các giải pháp và đóng góp nổi bật của tôi trong đồ án tốt nghiệp. Tôi sẽ mô tả những đột phá thiết yếu mà phần mềm đem đến cho người dùng.

Cuối cùng là CHƯƠNG 6. Phần này tôi sẽ nói về kết luận và hướng phát triển mà tôi muốn đồ án tốt nghiệp này có thể hướng tới.

## CHƯƠNG 2. KHẢO SÁT VÀ PHÂN TÍCH YÊU CẦU

### 2.1 Khảo sát hiện trạng

Theo Bruce Nikkel, mỗi cơ quan, tổ chức đều có những quy định cụ thể riêng đối với việc thu thập chứng cứ kỹ thuật số. Trong các tổ chức tư nhân, các quy chuẩn và thông lệ thường được tạo ra bởi các tổ chức và nhóm ngành khác nhau. Ví dụ, Hội đồng Cố vấn Đảm bảo Thông tin (Information Assurance Advisory Council - IAAC) đã cung cấp Hướng dẫn của Giám đốc và Cố vấn Công ty về Điều tra và Chứng cứ Kỹ thuật số [1].

Ngoài ra, còn có các nguồn khác, bao gồm các quy chuẩn, quy trình được ủy thác bởi luật pháp và các cơ quan quản lý: ví dụ, các yêu cầu đối với việc thu thập chứng cứ trong đạo luật Sarbanes-Oxley của Hoa Kỳ [1].

Một số yêu cầu về chứng cứ kỹ thuật số cũng có thể tùy thuộc vào ngành. Ví dụ, các tổ chức chăm sóc sức khỏe trong một khu vực có thể có quy định đặc biệt về đảm bảo an toàn dữ liệu và thu thập chứng cứ pháp y trong trường hợp xảy ra sự cố. Các nhà cung cấp dịch vụ viễn thông có thể có các quy định về lưu trữ vết sự kiện và quyền truy cập thực thi đối với dữ liệu thông tin liên lạc. Các cơ quan quản lý ngân hàng cũng có những quy định và yêu cầu riêng về chứng cứ kỹ thuật số liên quan đến gian lận tài chính (cụ thể là gian lận mạng). Một ví dụ điển hình có thể kể đến là MAS của Singapore, cung cấp các quy chuẩn chi tiết về lĩnh vực an ninh và phản ứng sự cố cho các hệ thống ngân hàng [1].

Chính vì không có bất kỳ quy chuẩn quốc tế nào xác định cách thực hiện quá trình pháp y trên hệ điều hành cho mọi trường hợp, nên các phần mềm và bộ công cụ hiện nay trên thế giới cũng rất khác nhau về phương pháp và dữ liệu thu thập được. Dưới đây là một số phần mềm thực hiện chức năng thu thập chứng cứ trên máy tính sử dụng hệ điều hành Linux hiện nay:

*Bảng 2-1 Một số phần mềm thu thập chứng cứ trên Linux*

Tên	URL	Chứng cứ thu thập
CyLR	<a href="https://github.com/orlikoski/CyLR">https://github.com/orlikoski/CyLR</a>	<ul style="list-style-type: none"><li>- Tập tin cấu hình hệ thống</li><li>- Tập tin cấu hình mạng</li><li>- Tập tin vết sự kiện</li><li>- Tập tin thông tin tài khoản người dùng</li></ul>
Live Response Collection (Cedarpelta Build)	<a href="https://www.brimorlabs.com/tools/">https://www.brimorlabs.com/tools/</a>	<ul style="list-style-type: none"><li>- Dữ liệu hệ thống từ bộ nhớ tạm và tập tin hệ thống</li><li>- Tập tin vết sự kiện</li><li>- Cấu hình và kết nối mạng</li><li>- Thông tin tài khoản người dùng</li><li>- Tạo ảnh ổ cứng</li></ul>

Đa phần các phần mềm và bộ công cụ pháp y kỹ thuật số trên Linux đều thiên về mục đích phân tích và điều tra chứng cứ. Những phần mềm hỗ trợ chức năng thu thập chứng cứ còn rất ít, và mỗi phần mềm lại có những ưu nhược điểm riêng.

Bảng 2-2 Ưu nhược điểm của những phần mềm thu thập chứng cứ trên Linux

Tên	Ưu điểm	Nhược điểm
CyLR	<ul style="list-style-type: none"> <li>- Tốc độ thu thập nhanh</li> <li>- Khả năng tùy chỉnh thông tin cần thu thập</li> <li>- Sử dụng giao diện dòng lệnh, dễ dàng sử dụng trên nhiều bản phân phối Linux</li> </ul>	<ul style="list-style-type: none"> <li>- Sao chép tệp nguyên bản, người thu thập cần phải có thêm công cụ để đọc các tệp tin định dạng đặc biệt sau khi thu thập</li> <li>- Chưa có chức năng đảm bảo tính toàn vẹn sau thu thập</li> </ul>
Live Response Collection (Cedarpelta Build)	<ul style="list-style-type: none"> <li>- Thông tin thu thập đa dạng, trực quan</li> <li>- Có chức năng đảm bảo tính toàn vẹn dữ liệu sau thu thập bằng mã hóa MD5 và SHA256</li> <li>- Sử dụng giao diện dòng lệnh, dễ dàng sử dụng trên nhiều bản phân phối Linux</li> </ul>	<ul style="list-style-type: none"> <li>- Chưa thu thập được lịch sử duyệt web trên máy tính</li> <li>- Không thể tùy chỉnh thông tin cần thu thập</li> <li>- Không thể tùy chỉnh vị trí tệp tin kết quả thu thập</li> </ul>

Đa số các phần mềm, công cụ thực hiện chức năng thu thập chứng cứ trên Linux hiện nay đều chưa thực sự đáp án được hết nhu cầu sử dụng của người dùng. Dựa trên những khảo sát đã thực hiện trên, phần mềm được xây dựng trong đồ án tốt nghiệp này cần đảm bảo các tính năng quan trọng như (i) thu thập được lượng lớn thông tin, đa dạng và trực quan; (ii) có thể tương tác, tùy chỉnh dễ dàng bởi người dùng; (iii) có chức năng đảm bảo tính toàn vẹn của chứng cứ sau thu thập; và (iv) có khả năng sử dụng trên nhiều bản phân phối Linux.

## 2.2 Tổng quan chức năng

Phần mềm bao gồm các chức năng chính sau: (i) thu thập chứng cứ hệ thống, (ii) thu thập chứng cứ mạng, (iii) thu thập vết sự kiện, (iv) thu thập thông tin tài khoản người dùng và (v) cấu hình đảm bảo tính toàn vẹn chứng cứ thu thập được.

### 2.2.1 Biểu đồ use case tổng quan

Tác nhân ở đây là người thu thập chứng cứ. Các use case chính bao gồm:

Use case tạo thư mục chứa chứng cứ được thu thập: bắt đầu một phiên thu thập chứng cứ.

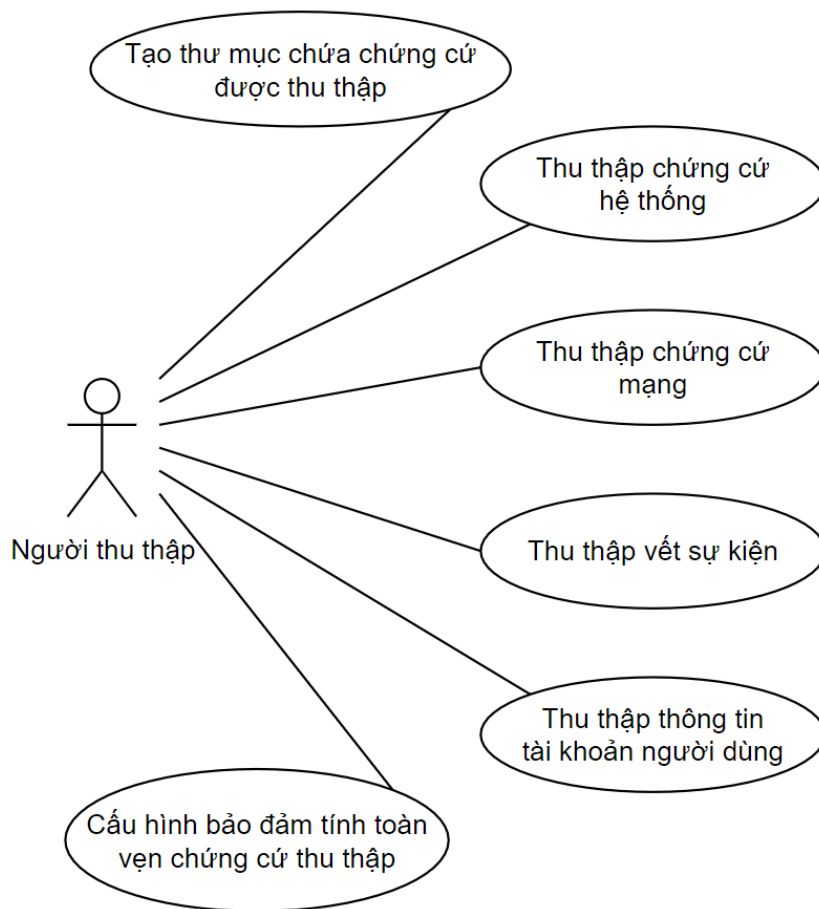
Use case thu thập chứng cứ hệ thống: thu thập những chứng cứ liên quan đến ứng dụng và hệ thống, bao gồm chứng cứ từ bộ nhớ tạm của máy tính và thông tin các tệp tin có trong máy.

Use case thu thập chứng cứ mạng: thu thập chứng cứ liên quan đến cấu hình và kết nối mạng.

Use case thu thập vết sự kiện: thu thập chứng cứ liên quan đến các phiên đăng nhập, lịch sử câu lệnh, lịch sử trình duyệt, log sự kiện, ...

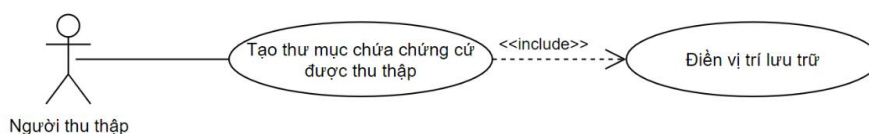
Use case thu thập thông tin tài khoản người dùng: thu thập chứng cứ liên quan đến tài khoản người dùng có trong máy tính.

Use case cấu hình đảm bảo tính toàn vẹn chứng cứ thu thập: thực hiện các tính năng đảm bảo tính toàn vẹn của chứng cứ sau khi quá trình thu thập kết thúc.



Hình 2-1 Biểu đồ use case tổng quan

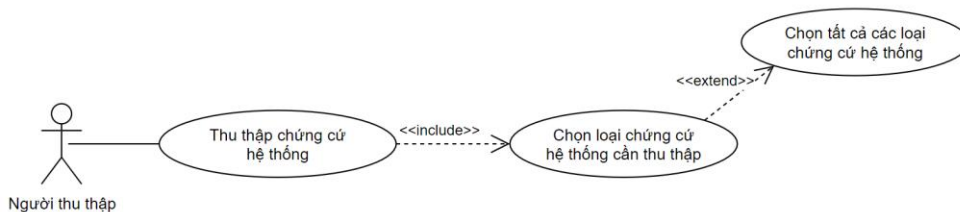
### 2.2.2 Biểu đồ use case phân rã tạo thư mục chứa chứng cứ được thu thập



Hình 2-2 Biểu đồ use case phân rã tạo thư mục chứa chứng cứ được thu thập

Sau khi khởi động phần mềm, người thu thập thực hiện điền vị trí lưu trữ - nơi tạo thư mục chứa chứng cứ được thu thập. Sau đó, phần mềm sẽ thực hiện tạo thư mục lưu trữ chứng cứ trong quá trình thu thập tại vị trí được chỉ định, và trả về màn hình giao diện chính của phần mềm trên cửa sổ dòng lệnh.

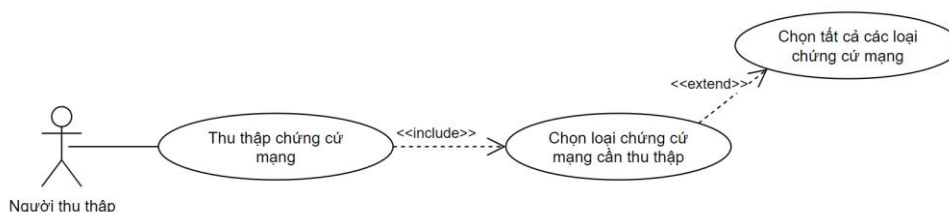
### 2.2.3 Biểu đồ use case phân rã thu thập chứng cứ hệ thống



Hình 2-3 Biểu đồ use case phân rã thu thập chứng cứ hệ thống

Người thu thập tiến hành chọn những loại chứng cứ liên quan đến hệ thống, có thể chọn riêng các loại chứng cứ mong muốn hoặc chọn tất cả. Sau đó chuyển sang các loại chứng cứ khác hoặc tiến hành thu thập.

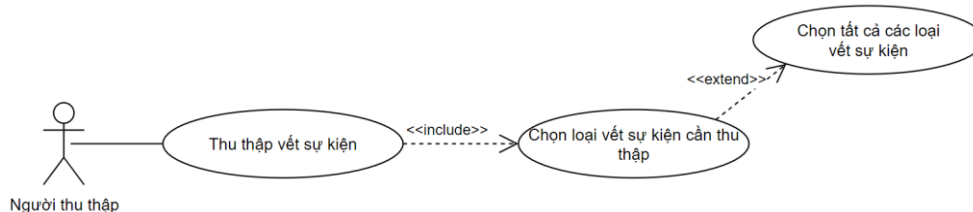
### 2.2.4 Biểu đồ use case phân rã thu thập chứng cứ mạng



Hình 2-4 Biểu đồ use case phân rã thu thập chứng cứ mạng

Người thu thập tiến hành chọn những loại chứng cứ liên quan đến hệ thống mạng, có thể chọn riêng các loại chứng cứ mong muốn hoặc chọn tất cả. Sau đó chuyển sang các loại chứng cứ khác hoặc tiến hành thu thập.

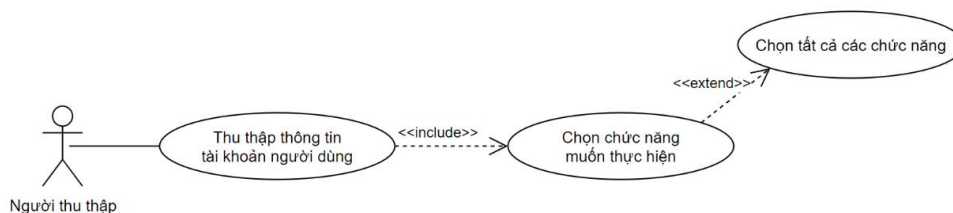
### 2.2.5 Biểu đồ use case phân rã thu thập vết sự kiện



Hình 2-5 Biểu đồ use case phân rã thu thập vết sự kiện

Người thu thập tiến hành chọn những loại vết sự kiện, có thể chọn riêng các loại vết sự kiện mong muốn hoặc chọn tất cả. Sau đó chuyển sang các loại chứng cứ khác hoặc tiến hành thu thập.

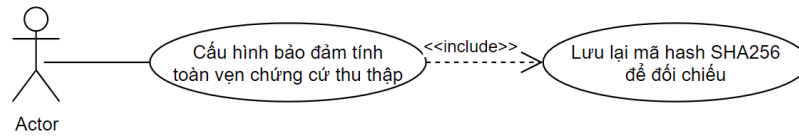
### 2.2.6 Biểu đồ use case phân rã thu thập thông tin tài khoản người dùng



Hình 2-6 Biểu đồ use case phân rã thu thập thông tin người dùng

Người thu thập tiến hành chọn các chức năng muốn thực hiện, có thể chọn riêng chỉ thu thập thông tin tài khoản người dùng có trong máy tính, hoặc thêm chức năng phân tích xác định tài khoản có dấu hiệu bất thường trên máy tính. Sau đó chuyển sang các loại chứng cứ khác, hoặc tiến hành thu thập.

### 2.2.7 Biểu đồ use case phân rã cấu hình đảm bảo tính toàn vẹn chứng cứ thu thập

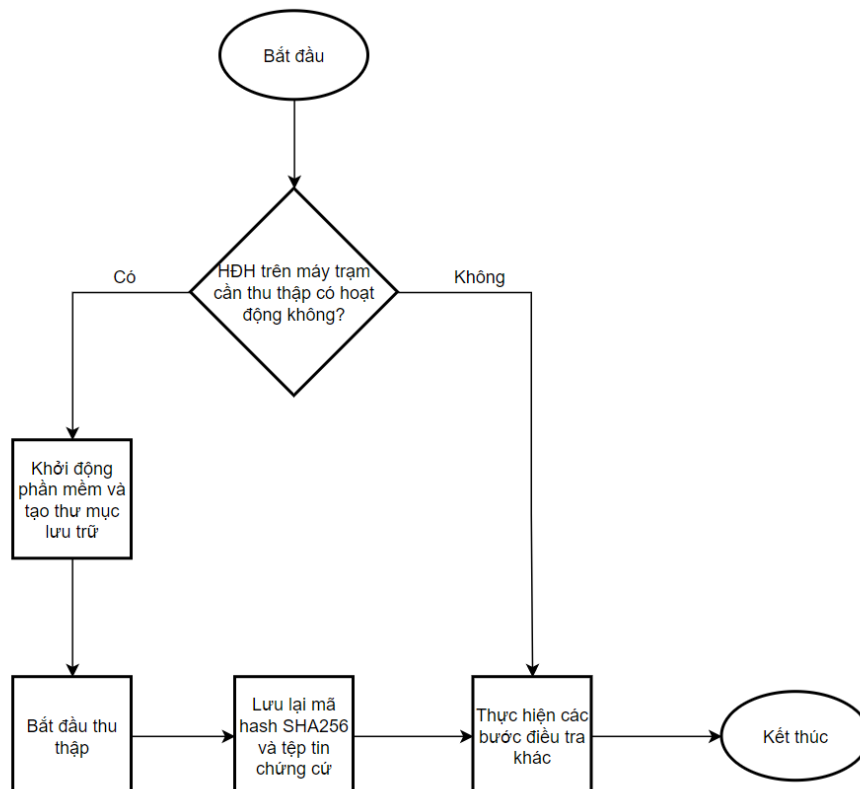


Hình 2-7 Biểu đồ use case phân rã cấu hình đảm bảo tính toàn vẹn chứng cứ thu thập

Người thu thập chọn chức năng cấu hình đảm bảo tính toàn vẹn chứng cứ thu thập. Tại đây, phần mềm sẽ tự động thực hiện nén thư mục chứng cứ vừa thu thập thành tệp tin nén có phần mở rộng “.zip”, sau đó xuất ra mã hóa SHA256 của tệp tin vừa nén để người thu thập lưu lại và đối chiếu kiểm tra tính toàn vẹn của chứng cứ khi cần.

### 2.2.8 Quy trình nghiệp vụ

Mục tiêu của phần mềm là có thể thu thập toàn bộ chứng cứ sự cố an toàn thông tin từ máy tính sử dụng hệ điều hành Linux.



Hình 2-8 Biểu đồ quy trình nghiệp vụ

Chú thích: HĐH - Hệ điều hành

Để tiến hành thu thập chứng cứ, hệ điều hành của máy tính đang xét cần khởi động được.

Bước 1: Khởi động hệ điều hành của máy tính cần thu thập và cắm thiết bị thu thập vào máy tính. Mở cửa sổ dòng lệnh, điều hướng đến thư mục chứa phần mềm và chạy tệp tin LinuxIncidentResponse dưới quyền quản trị viên trên cửa sổ dòng lệnh bằng câu lệnh “sudo ./LinuxIncidentResponse”. Nhập vào đường dẫn đến thư mục dự định chứa chứng cứ thu thập.

Bước 2: Sau khi nhập vào đường dẫn đến thư mục dự định chứa chứng cứ thu thập, phần mềm sẽ tạo một thư mục có tên được đặt dựa theo tên thiết bị và thời gian thực hiện quá trình thu thập. Sau đó, trên màn hình cửa sổ dòng lệnh sẽ hiển thị giao diện chính của phần mềm với các chức năng thu thập cho người thu thập tiến hành lựa chọn. Sau khi lựa chọn các chức năng và loại chứng cứ mong muốn, người thu thập xác nhận lựa chọn của mình trên cửa sổ dòng lệnh để phần mềm bắt đầu quá trình thu thập.

Bước 3: Người thu thập đợi đến quá trình thu thập kết thúc, màn hình cửa sổ dòng lệnh trả về thông báo “Exit...”. Lúc này tiến trình thu thập và đảm bảo tính toàn vẹn đã hoàn thành. Người thu thập kiểm tra lại tệp tin chứng cứ một lần nữa và lưu lại mã hóa SHA256 được hiển thị trên màn hình cửa sổ dòng lệnh để tiến hành đối chiếu sau này. Nếu mọi thứ đều hoạt động bình thường, người thu thập có thể tắt cửa sổ dòng lệnh, rút thiết bị và chuyển sang máy tính khác để tiếp tục thu thập.

## 2.3 Đặc tả chức năng

### 2.3.1 Đặc tả use case tạo thư mục chứa chứng cứ được thu thập

*Bảng 2-3 Đặc tả use case tạo thư mục chứa chứng cứ được thu thập*

<b>Tên use case</b>	Tạo thư mục chứa chứng cứ được thu thập
<b>Mô tả</b>	Người thu thập tạo một thư mục chứa chứng cứ trong quá trình thu thập và bắt đầu phiên thu thập
<b>Tác nhân</b>	Người thu thập
<b>Tiền điều kiện</b>	- Người thu thập có quyền quản trị viên - Thiết bị thu thập được kết nối với máy tính cần thu thập
<b>Hậu điều kiện</b>	- Một thư mục chứa chứng cứ thu thập được tạo - Phần mềm sẵn sàng bắt đầu quá trình thu thập
<b>Luồng sự kiện chính</b>	1 Người thu thập khởi động phần mềm 2 Người thu thập nhập vào đường dẫn tới thư mục chứa chứng cứ thu thập mong muốn



### 2.3.2 Đặc tả use case thu thập chứng cứ hệ thống

Bảng 2-4 Đặc tả use case thu thập chứng cứ hệ thống

<b>Tên use case</b>	Thu thập chứng cứ hệ thống
<b>Mô tả</b>	Người thu thập muốn thu thập các loại chứng cứ liên quan đến ứng dụng, hệ thống; bao gồm chứng cứ từ bộ nhớ tạm của máy tính và thông tin các tệp tin có trong máy
<b>Tác nhân</b>	Người thu thập
<b>Tiền điều kiện</b>	- Người thu thập có quyền quản trị viên - Thiết bị thu thập được kết nối với máy tính cần thu thập - Người thu thập đã tạo thư mục chứa chứng cứ và bắt đầu phiên thu thập
<b>Hậu điều kiện</b>	- Chứng cứ liên quan đến ứng dụng và hệ thống
<b>Luồng sự kiện chính</b>	1 Người thu thập chọn các loại chứng cứ hệ thống muốn thu thập và điền vào cửa sổ dòng lệnh 2 Người thu thập nhấn Enter
<b>Luồng sự kiện phát sinh</b>	Phần mềm sẽ trả về thông tin những lỗi gặp phải trong quá trình thu thập

### 2.3.3 Đặc tả use case thu thập chứng cứ mạng

Bảng 2-5 Đặc tả use case thu thập chứng cứ mạng

<b>Tên use case</b>	Thu thập chứng cứ mạng
<b>Mô tả</b>	Người thu thập muốn thu thập các loại chứng cứ liên quan đến mạng
<b>Tác nhân</b>	Người thu thập
<b>Tiền điều kiện</b>	- Người thu thập có quyền quản trị viên - Thiết bị thu thập được kết nối với máy tính cần thu thập - Người thu thập đã tạo thư mục chứa chứng cứ và bắt đầu phiên thu thập
<b>Hậu điều kiện</b>	- Chứng cứ liên quan đến mạng
<b>Luồng sự kiện</b>	1 Người thu thập chọn các loại chứng cứ mạng muốn thu thập và điền vào cửa sổ dòng lệnh 2 Người thu thập nhấn Enter
<b>Luồng sự kiện phát sinh</b>	Phần mềm sẽ trả về thông tin những lỗi gặp phải trong quá trình thu thập

### 2.3.4 Đặc tả use case thu thập vết sự kiện

Bảng 2-6 Đặc tả use case thu thập vết sự kiện

<b>Tên use case</b>	Thu thập vết sự kiện
<b>Mô tả</b>	Người thu thập muốn thu thập các loại vết sự kiện
<b>Tác nhân</b>	Người thu thập
<b>Tiền điều kiện</b>	<ul style="list-style-type: none"><li>- Người thu thập có quyền quản trị viên</li><li>- Thiết bị thu thập được kết nối với máy tính cần thu thập</li><li>- Người thu thập đã tạo thư mục chứa chứng cứ và bắt đầu phiên thu thập</li></ul>
<b>Hậu điều kiện</b>	- Vết sự kiện
<b>Luồng sự kiện</b>	1 Người thu thập chọn các loại vết sự kiện muốn thu thập và điền vào cửa sổ dòng lệnh 2 Người thu thập nhấn Enter
<b>Luồng sự kiện phát sinh</b>	Phần mềm sẽ trả về thông tin những lỗi gặp phải trong quá trình thu thập

### 2.3.5 Đặc tả use case thu thập thông tin tài khoản người dùng

Bảng 2-7 Đặc tả use case thu thập thông tin tài khoản người dùng

<b>Tên use case</b>	Thu thập thông tin tài khoản người dùng
<b>Mô tả</b>	Người thu thập muốn thu thập thông tin tài khoản người dùng có trong máy tính
<b>Tác nhân</b>	Người thu thập
<b>Tiền điều kiện</b>	<ul style="list-style-type: none"><li>- Người thu thập có quyền quản trị viên</li><li>- Thiết bị thu thập được kết nối với máy tính cần thu thập</li><li>- Người thu thập đã tạo thư mục chứa chứng cứ và bắt đầu phiên thu thập</li></ul>
<b>Hậu điều kiện</b>	- Thông tin tài khoản người dùng
<b>Luồng sự kiện</b>	1 Người thu thập chọn các chức năng liên quan đến tài khoản người dùng muốn sử dụng và điền vào cửa sổ dòng lệnh 2 Người thu thập nhấn Enter
<b>Luồng sự kiện phát sinh</b>	Phần mềm sẽ trả về thông tin những lỗi gặp phải trong quá trình thu thập

### 2.3.6 Đặc tả use case đảm bảo tính toàn vẹn chứng cứ thu thập

Bảng 2-8 Đặc tả use case đảm bảo tính toàn vẹn chứng cứ thu thập

<b>Tên use case</b>	Đảm bảo tính toàn vẹn chứng cứ thu thập
<b>Mô tả</b>	Người thu thập muốn đảm bảo tính toàn vẹn chứng cứ thu thập được
<b>Tác nhân</b>	Người thu thập
<b>Tiền điều kiện</b>	<ul style="list-style-type: none"><li>- Người thu thập có quyền quản trị viên</li><li>- Thiết bị thu thập được kết nối với máy tính cần thu thập</li><li>- Người thu thập đã tạo thư mục chứa chứng cứ và thu thập xong chứng cứ cần thiết</li></ul>
<b>Hậu điều kiện</b>	<ul style="list-style-type: none"><li>- Chứng cứ được nén lại, đảm bảo tính toàn vẹn bằng mã hóa SHA256 để đối chiếu</li></ul>
<b>Luồng sự kiện</b>	<ol style="list-style-type: none"><li>1 Người thu thập chọn chức năng đảm bảo tính toàn vẹn của chứng cứ thu thập được và điền vào cửa sổ dòng lệnh</li><li>2 Người thu thập nhấn Enter</li></ol>
<b>Luồng sự kiện phát sinh</b>	Phần mềm sẽ trả về thông tin những lỗi gặp phải trong quá trình thu thập

### 2.4 Yêu cầu phi chức năng

Bên cạnh các yêu cầu chính cần thiết để đáp ứng các vấn đề được đặt ra ban đầu, có những yêu cầu phi chức năng khác cũng cần được giải quyết. Những yêu cầu đó bao gồm (i) phần mềm cần dễ sử dụng, dễ bảo trì, có thể nâng cấp và tùy chỉnh để phù hợp với nhiều mục đích; (ii) phần mềm phải gọn nhẹ, không yêu cầu cài đặt nhằm hạn chế thay đổi gây ra trong hệ thống; (iii) để giảm thiểu tối đa dấu vết để lại trong bộ nhớ và trên hệ thống, chỉ sử dụng phần mềm được lưu trữ trên thiết bị thu thập cắm vào máy tính gặp sự cố, không sao chép phần mềm trực tiếp lên máy tính.

## CHƯƠNG 3. CÔNG NGHỆ SỬ DỤNG

Sau khi phân tích yêu cầu của phần mềm, tại chương này tôi sẽ đi sâu hơn vào lý thuyết và công nghệ được sử dụng.

### 3.1 Cơ sở lý thuyết

#### 3.1.1 Quy chuẩn pháp y kỹ thuật số

Theo Bruce Nikkel, so với các quá trình pháp y thông thường, có rất ít tiêu chuẩn chung dành cho việc phân tích hệ điều hành. Quy trình phân tích pháp y hệ điều hành thường được thực hiện theo các yêu cầu, chính sách từ các phòng thí nghiệm kỹ thuật số và theo giới hạn khả năng của phần mềm phân tích pháp y. Không có bất kỳ quy chuẩn quốc tế nào xác định cách thực hiện quá trình pháp y trên hệ điều hành. Hệ điều hành nói chung quá đa dạng, phức tạp, và thay đổi quá nhanh để có thể xác định một quy chuẩn chung [1].

Mặc dù vậy, vẫn có một số nguyên tắc hướng dẫn chung cần tuân thủ khi thực hiện quá trình pháp y. Các nguyên tắc này bao gồm (i) duy trì tính toàn vẹn của chứng cứ, (ii) đảm bảo vết sự kiện đúng trình tự, (iii) tuân theo những quy tắc tiêu chuẩn, và (iv) ghi chép lại đầy đủ mọi thứ [2].

#### 3.1.2 Thu thập chứng cứ số

Dựa theo quy trình thu thập trực tiếp chứng cứ số trên máy tính sử dụng hệ điều hành Linux còn khả năng hoạt động của Philip Polstra, các dữ liệu cần được kiểm tra để làm chứng cứ thu thập bao gồm (i) dữ liệu từ bộ nhớ tạm, (ii) thông tin tệp tin trên máy tính, (iii) lịch sử các câu lệnh sử dụng, (iv) thu thập các tệp tin vết sự kiện, và (v) tạo ảnh các ổ đĩa có trên máy tính [2].

### 3.2 Công nghệ sử dụng

#### 3.2.1 PyInstaller

PyInstaller đóng gói một ứng dụng Python và tất cả các công cụ cần thiết của nó vào một package duy nhất. Người dùng có thể chạy ứng dụng được đóng gói trong package mà không cần cài đặt trình thông dịch Python hoặc bất kỳ module nào.

PyInstaller được sử dụng trong đề tài nhằm mục đích tạo ra môi trường phù hợp để chạy phần mềm trên tất cả các máy tính, kể cả các máy tính chưa cài đặt Python. Nhờ vậy, phần mềm có thể tối thiểu hóa dấu vết để lại trên bộ nhớ và hệ thống của máy tính cần thu thập, do có thể sử dụng trực tiếp phần mềm từ thiết bị thu thập mà không cần cài đặt thêm công cụ gì.

#### 3.2.2 Các thư viện, module Python sử dụng

##### 3.2.2.1. *os*

*os* là module Python cung cấp khả năng sử dụng các chức năng của hệ điều hành. Module được sử dụng trong đề tài với chức năng thực hiện các lệnh hệ thống nhằm mục đích thu thập chứng cứ và tương tác với các đường dẫn thư mục trong hệ điều hành.

### 3.2.2.2. *subprocess*

*subprocess* là module Python cung cấp phương tiện cho phép tạo ra những process mới và truy xuất kết quả của chúng. Module này được tạo ra để thay thế cho một vài module và chức năng cũ, ví dụ như `os.system` và `os.spawn`.

Module được sử dụng trong đề tài với chức năng thực hiện các lệnh hệ thống nhằm mục đích thu thập chứng cứ từ hệ điều hành.

### 3.2.2.3. *sqlite3*

*sqlite3* là module Python được viết bởi Gerhard Häring, cung cấp khả năng tương tác với dữ liệu SQLite tuân theo đặc tả DB-API 2.0 được mô tả bởi PEP 249. Module yêu cầu SQLite phiên bản 3.7.15 hoặc mới hơn.

Đa phần các trình duyệt hiện tại lưu trữ lịch sử duyệt web tại cơ sở dữ liệu SQLite, nên module được sử dụng với chức năng tìm nạp lịch sử duyệt web của các trình duyệt trên máy tính từ cơ sở dữ liệu lưu trữ.

### 3.2.2.4. *csv*

Module Python *csv* triển khai các lớp để đọc và ghi dữ liệu dạng bảng ở định dạng CSV. Lập trình viên có thể tùy chỉnh định dạng CSV để tệp tin có thể đọc được bởi Excel và các ứng dụng khác, hoặc định nghĩa lại định dạng CSV cho mục đích riêng của họ.

Module được sử dụng trong đề tài với chức năng ghi dữ liệu lịch sử duyệt web từ máy tính ra tệp tin chứng cứ có định dạng CSV.

### 3.2.2.5. *re*

*re* là module Python cung cấp khả năng so sánh các regular expression tương tự như trong ngôn ngữ lập trình Perl.

Module được sử dụng trong đề tài với chức năng so sánh để xác định các tài khoản người dùng có dấu hiệu khả nghi trên máy tính trong quá trình thu thập chứng cứ.

### 3.2.2.6. *shutil*

Module Python *shutil* cung cấp nhiều thao tác cấp cao trên tệp tin hoặc tập hợp các tệp tin. Cụ thể, các chức năng được hỗ trợ bao gồm sao chép và xóa tệp tin. Bên cạnh đó, các tiện ích cấp cao để tạo và đọc các tệp tin nén cũng được module này hỗ trợ.

Module này được sử dụng trong đề tài với chức năng nén thư mục lưu trữ chứng cứ sau khi thu thập.

### 3.2.2.7. *hashlib*

Module Python *hashlib* cung cấp nhiều thuật toán băm và phân tích thông điệp khác nhau, bao gồm các thuật toán băm an toàn SHA1, SHA224, SHA256, SHA384, và SHA512 (được định nghĩa trong FIPS 180-2) cũng như thuật toán MD5 của RSA (được định nghĩa trong internet RFC 1321).

Module này được sử dụng trong đề tài với chức năng xuất mã hóa SHA256 của tệp tin nén chứa chứng cứ thu thập.

## CHƯƠNG 4. PHÁT TRIỂN VÀ TRIỂN KHAI ỨNG DỤNG

### 4.1 Xây dựng ứng dụng

#### 4.1.1 Công cụ sử dụng

Để xây dựng và phát triển phần mềm thu thập chứng cứ trên máy tính sử dụng hệ điều hành Linux, ta cần một số công cụ để hỗ trợ quá trình lập trình và kiểm thử. Dưới đây là danh mục các công cụ mà tôi sử dụng trong đồ án này.

*Bảng 4-1 Danh sách công cụ sử dụng*

Mục đích	Công cụ	Địa chỉ URL
IDE lập trình	Visual Studio Code 1.62.3	<a href="https://code.visualstudio.com/">https://code.visualstudio.com/</a>
Tạo môi trường ảo hóa để kiểm thử phần mềm	Vmware Workstation Pro 16.2.1	<a href="https://www.vmware.com/products/workstation-pro.html">https://www.vmware.com/products/workstation-pro.html</a>
Ngôn ngữ lập trình	Python 3.8	<a href="https://www.python.org/">https://www.python.org/</a>

#### 4.1.2 Yêu cầu chức năng

Yêu cầu chức năng cần thực hiện của phần mềm được thể hiện trong bảng dưới. Mô tả chi tiết các chứng cứ thu thập sẽ được trình bày ở phần 5.1.

Bảng 4-2 Chi tiết các chức năng cần thực hiện

Chức năng	Chi tiết
Tạo thư mục chứa chứng cứ được thu thập	<ul style="list-style-type: none"> <li>- Kiểm tra đường dẫn được người dùng nhập vào, nếu đường dẫn không tồn tại, tạo đường dẫn theo yêu cầu của người dùng</li> <li>- Tạo thư mục chứa chứng cứ được thu thập tại đường dẫn theo yêu cầu</li> </ul>
Thu thập chứng cứ hệ thống	<ul style="list-style-type: none"> <li>- Thu thập thông tin cơ bản của hệ thống</li> <li>- Thu thập thông tin về các tệp tin trên hệ thống đang hoặc đã được mở bởi các tiến trình</li> <li>- Thu thập thông tin các tiến trình đang chạy</li> <li>- Thu thập thông tin dung lượng ổ cứng đang sử dụng và còn sẵn trong hệ thống</li> <li>- Thu thập thông tin các phân vùng được mount trên hệ thống</li> <li>- Thu thập thông tin các module đang được sử dụng bởi Linux kernel</li> <li>- Thu thập thông tin chi tiết của các tệp tin trong hệ thống</li> <li>- Thu thập hàm băm của các tệp tin trong hệ thống</li> </ul>
Thu thập chứng cứ mạng	<ul style="list-style-type: none"> <li>- Thu thập thông tin cấu hình giao diện hệ thống mạng</li> <li>- Thu thập thông tin các cổng mạng được mở</li> <li>- Thu thập thông tin bảng định tuyến</li> </ul>
Thu thập vết sự kiện	<ul style="list-style-type: none"> <li>- Thu thập thông tin người dùng mới đăng nhập</li> <li>- Thu thập thông tin những lần đăng nhập thất bại</li> <li>- Thu thập lịch sử sử dụng dòng lệnh</li> <li>- Thu thập các tệp tin vết sự kiện</li> <li>- Thu thập lịch sử trình duyệt</li> </ul>
Thu thập thông tin tài khoản người dùng	<ul style="list-style-type: none"> <li>- Thu thập thông tin tài khoản người dùng có trong hệ thống</li> <li>- Kiểm tra tính toàn vẹn của thông tin tài khoản người dùng có trong hệ thống</li> </ul>
Cấu hình đảm bảo tính toàn vẹn chứng cứ thu thập	<ul style="list-style-type: none"> <li>- Nén thư mục chứa chứng cứ được thu thập thành tệp tin có phần mở rộng “.zip”</li> <li>- Tạo mã hóa SHA256 để đối chiếu tính toàn vẹn khi cần</li> </ul>

#### 4.1.3 Lập trình xây dựng ứng dụng

Do độ dài của báo cáo có hạn, nên trong phần này tôi sẽ chỉ trình bày chi tiết một số hàm đặc trưng của phần mềm.

##### 4.1.3.1. Hàm thu thập lịch sử sử dụng dòng lệnh

Hàm này thuộc module thu thập vết sự kiện, có tác dụng kiểm tra và thu thập những câu lệnh được sử dụng của tất cả các tài khoản người dùng tồn tại trong máy tính.

Trên Linux, lịch sử sử dụng dòng lệnh được lưu lại trong tệp tin `.bash_history`, lưu trữ tại thư mục riêng của mỗi tài khoản người dùng trong thư mục `/home`. Để thu thập lịch sử sử dụng dòng lệnh có trên máy, phần mềm sử dụng module subprocess thực hiện câu lệnh `grep` có sẵn trên Linux, ghi lại lịch sử dòng lệnh của tất cả tài khoản người dùng được lưu trữ trong `.bash_history` ra một tệp tin kết quả có phần mở rộng `".txt"`.

Tham số đầu vào của hàm bao gồm đường dẫn thư mục kết quả và eID của chứng cứ. Kết quả đầu ra là tệp tin `14.BashHistory_eID.txt` nằm trong thư mục `C.Event_eID` tại thư mục kết quả thu thập.

Nếu gặp lỗi trong quá trình thực hiện, hàm sẽ trả về thông báo lỗi gặp phải trên màn hình dòng lệnh.

```
def getBashHistory(outp, eID):
    fName = "14.BashHistory_" + eID + ".txt"
    try:
        api.exeCmd(
            'sudo grep -e "$pattern" /home/*/.bash_history'
            + ">"
            + os.path.join(outp, fName)
        )
        print(bcolors.OKGREEN + fName + " export successfully!" + bcolors.ENDC)
    except Exception as e:
        print(bcolors.FAIL + fName + " export failed!")
        print("NameError: " + str(e) + bcolors.ENDC)
```

Hình 4-1 Hàm thu thập lịch sử sử dụng dòng lệnh

#### 4.1.3.2. Hàm thu thập lịch sử trình duyệt

Hàm này thuộc module thu thập vết sự kiện, có tác dụng kiểm tra và thu thập lịch sử duyệt web trên trình duyệt của tất cả các tài khoản người dùng tồn tại trong máy tính.

Lịch sử duyệt web của mỗi trình duyệt đa phần đều có điểm chung là được lưu trữ trong cơ sở dữ liệu SQLite, nằm trong thư mục tệp tin cấu hình riêng của mỗi trình duyệt. Phần mềm sẽ thực hiện tìm kiếm và thu thập các tệp tin cơ sở dữ liệu chứa lịch sử duyệt web của từng tài khoản người dùng với mỗi trình duyệt, và sử dụng module sqlite3 từ Python để đọc dữ liệu trong những tệp tin này. Hiện tại, phần mềm có thể tìm và đọc lịch sử duyệt web từ 2 trình duyệt phổ biến nhất trên Linux là Google Chrome và Mozilla Firefox.

Tham số đầu vào của hàm bao gồm đường dẫn thư mục kết quả và eID của chứng cứ. Kết quả đầu ra là thư mục `16.BrowserHistory_eID` chứa lịch sử duyệt web nằm trong thư mục `C.Event_eID` tại thư mục kết quả thu thập. Các tệp tin lịch sử trình duyệt có phần mở rộng `".csv"` để thuận tiện cho công việc điều tra sau này.

Nếu gặp lỗi trong quá trình thực hiện, hàm sẽ trả về thông báo lỗi gặp phải trên màn hình dòng lệnh.



```

def getBrowserHistory(outp, eID):
    fName = "16.BrowserHistory_" + eID
    folder = os.path.join(outp, fName)
    if not os.path.exists(folder):
        os.makedirs(folder)
    try:
        users = api.exeCmd("cd /home; ls")
        for user in users.split():
            userFolder = os.path.join(folder, user)

```

Hình 4-2 Tạo thư mục lưu trữ kết quả

```

if os.path.exists(
    "/home/"
    + user
    + "/.config/google-chrome/"
    + "Guest Profile"
    + "/History"
):
    if not os.path.exists(userFolder):
        os.makedirs(userFolder)
    if not os.path.exists(userChromeFolder):
        os.makedirs(userChromeFolder)
    getChromeHistory(userChromeFolder, user, "Guest Profile")

if os.path.exists(
    "/home/" + user + "/.config/google-chrome/" + "Default" + "/History"
):
    if not os.path.exists(userFolder):
        os.makedirs(userFolder)
    if not os.path.exists(userChromeFolder):
        os.makedirs(userChromeFolder)
    getChromeHistory(userChromeFolder, user, "Default")

i = 1
while True:
    cProfile = "Profile " + str(i)
    if os.path.exists(
        "/home/" + user + "/.config/google-chrome/" + cProfile + "/History"
    ):
        if not os.path.exists(userFolder):
            os.makedirs(userFolder)
        if not os.path.exists(userChromeFolder):
            os.makedirs(userChromeFolder)
        getChromeHistory(userChromeFolder, user, cProfile)
        i = i + 1
    else:
        break

```

Hình 4-3 Tìm kiếm tệp tin cơ sở dữ liệu lịch sử trình duyệt trên Chrome

```
def getChromeHistory(path, user, profile):
    fname = profile + ".csv"
    try:
        con = sqlite3.connect(
            "/home/" + user + "/.config/google-chrome/" + profile + "/History"
        )
        c = con.cursor()
        c.execute(
            "SELECT datetime(last_visit_time / 1000000 - 11644473600, 'unixepoch', 'localtime'), url, title, visit_count FROM urls ORDER BY last_visit_time DESC"
        )
        results = c.fetchall()

        f = open(os.path.join(path, fname), "w")
        cols = ("Datetime", "URL", "Title", "Visit Count")
        write = csv.writer(f)
        write.writerow(cols)
        write.writerows(results)

        c.close()
    except Exception as e:
        print(bcolors.FAIL + fname + " export failed!")
        print("NameError: " + str(e) + bcolors.ENDC)
```

Hình 4-4 Đọc lịch sử trình duyệt Chrome từ cơ sở dữ liệu SQLite

#### 4.1.3.3. Hàm đảm bảo tính toàn vẹn chứng cứ thu thập

Hàm này thực hiện chức năng nén và xuất mã hóa SHA256 để đối chiếu kiểm tra tính toàn vẹn của chứng cứ khi cần.

Phần mềm sử dụng module Python shutil để nén thư mục lưu trữ kết quả, sau đó xuất mã hóa SHA256 của tệp tin vừa nén bằng module hashlib. Mã hóa SHA256 sẽ được hiển thị ra màn hình dòng lệnh với màu sắc đặc trưng để người dùng dễ nhận biết và chú ý lưu trữ.

Nếu gặp lỗi trong quá trình thực hiện, hàm sẽ trả về thông báo lỗi gặp phải trên màn hình dòng lệnh.

```
try:
    outp = os.path.join(rawOutp, folderName)
    shutil.make_archive(outp, "zip", rawOutp, folderName)

    zipOutp = outp + ".zip"
    api.exeCmd("sudo chmod 777 -R " + zipOutp)

    print(bcolors.OKGREEN + "Compress files successfully!" + bcolors.ENDC)
except Exception as e:
    print(bcolors.FAIL + "There was an error while compressing " + folderName + "!")
    print("NameError: " + str(e) + bcolors.ENDC)
```

Hình 4-5 Thực hiện nén thư mục lưu trữ kết quả

```

try:
    print(bcolors.OKCYAN + "Exporting encrypted SHA256 signature..." + bcolors.ENDC)
    sha256_hash = hashlib.sha256()
    with open(zipOutp, "rb") as f:
        for byte_block in iter(lambda: f.read(4096), b''):
            sha256_hash.update(byte_block)

    print(
        bcolors.OKGREEN
        + "Export encrypted SHA256 signature successfully!"
        + bcolors.ENDC
    )

    print(
        bcolors.OKGREEN
        + "SHA256 Signature: "
        + bcolors.ENDC
        + bcolors.HEADER
        + sha256_hash.hexdigest()
        + bcolors.ENDC
    )
except Exception as e:
    print(
        bcolors.FAIL
        + "There was an errors while exporting "
        + folderName
        + ".zip encrypted SHA256 signature!"
    )
print("NameError: " + str(e) + bcolors.ENDC)

```

Hình 4-6 Xuất mã hóa SHA256 của tệp tin được nén

#### 4.1.1 Thiết kế giao diện

Do đặc thù sử dụng của các hệ điều hành Linux, nhiều bản phân phối chỉ hỗ trợ dòng lệnh, phần mềm cũng hướng đến đối tượng người dùng là những người có kinh nghiệm sử dụng, quen thuộc với hệ điều hành Linux, nên phần mềm sẽ sử dụng giao diện dòng lệnh và tương tác với người dùng thông qua cửa sổ dòng lệnh quen thuộc của các hệ điều hành Linux.

Sử dụng giao diện dòng lệnh sẽ có những ưu điểm khác, ví dụ như: (i) giảm dung lượng phần mềm, (ii) tối thiểu hóa tác động đến máy tính cần thu thập, (iii) giúp việc chỉnh sửa và nâng cấp phần mềm trong tương lai dễ dàng hơn. Tuy nhiên, nó cũng đi kèm một số nhược điểm. Giao diện dòng lệnh kém trực quan hơn giao diện đồ họa thông thường, và cũng yêu cầu một số kiến thức cơ bản sử dụng dòng lệnh trên Linux.

Để khắc phục những nhược điểm trên, phần mềm sử dụng những thông báo, yêu cầu chi tiết và thân thiện với người dùng trong quá trình vận hành để người dùng có thể dễ dàng làm quen và thao tác. Ngoài ra, tôi cũng sử dụng màu sắc để phân biệt các thông tin được hiển thị trên màn hình dòng lệnh, giúp người dùng có thể dễ dàng phân biệt giữa thông báo và yêu cầu trong quá trình phần mềm vận hành.

Bảng 4-3 Ý nghĩa màu sắc được sử dụng

Màu sắc	Nội dung
Tím	Biểu ngữ của phần mềm
Vàng	Yêu cầu người dùng cần thực hiện
Xanh Neon	Menu chức năng Thông báo đang thu thập
Xanh lá	Thông báo thu thập thành công
Đỏ	Thông báo lỗi

Minh họa giao diện phần mềm được thiết kế:

```

root@kali: /media/flash/LinuxForensics
File Actions Edit View Help
Select one or more of the options below to launch the program by choosing the number
before it and entering into the area below:
*Options are separated by spaces.
00 - Select all

A. System Data
01 - Collect system information
02 - Collect list of open files
03 - Collect process status
04 - Collect disk filesystems information
05 - Collect list of mounted filesystems
06 - Collect loaded kernel modules information
07 - Collect file metadata
08 - Collect file hashes

B. Network Data
09 - Collect network interfaces
10 - Collect network statistics
11 - Collect network routing tables

C. Event Data
12 - Collect list of last logged-in users
13 - Collect list of failed logins
14 - Collect bash history
15 - Collect log files
16 - Collect browser history

D. User Data
17 - Collect local users information
18 - Check the integrity of local user credentials

E. Compress files, export encrypted SHA256 signature for integrity check and exit

Enter your options:
>>> █
  
```

Hình 4-7 Minh họa giao diện phần mềm

#### 4.1.2 Thiết kế cấu trúc thư mục kết quả

Việc thu thập chứng cứ trên nhiều máy tính mỗi khi có sự cố an toàn thông tin là một điều thường xuyên xảy ra. Cần phải có giải pháp để phân biệt dữ liệu chứng cứ giữa các máy tính khác nhau.

Cung cấp chức năng cho phép người dùng lựa chọn tên định dạng cho chứng cứ thu thập là một giải pháp. Tuy nhiên, chứng cứ thu thập được có thể sẽ phải qua nhiều bước điều tra khác nhau, tên gợi nhớ đối với người thu thập có thể lại không có ý nghĩa gì đối với chuyên viên điều tra. Ngoài ra, đối với các chứng cứ mang tính chất pháp lý, có thể được sử dụng trước tòa án pháp luật, cần phải có ít sự tác động, chỉnh sửa từ người dùng nhất có thể. Vì những lý do trên, chúng ta cần xây dựng một hệ thống định danh cho chứng cứ thu thập được từ phần mềm. Phần mềm sử dụng một chuỗi định danh đặc biệt để phân biệt giữa các máy tính được điều tra, và giữa các lần thu thập trên cùng một máy tính. Chuỗi định danh này được dựa trên cấu trúc:

$$eID = hostname\_epochTime$$

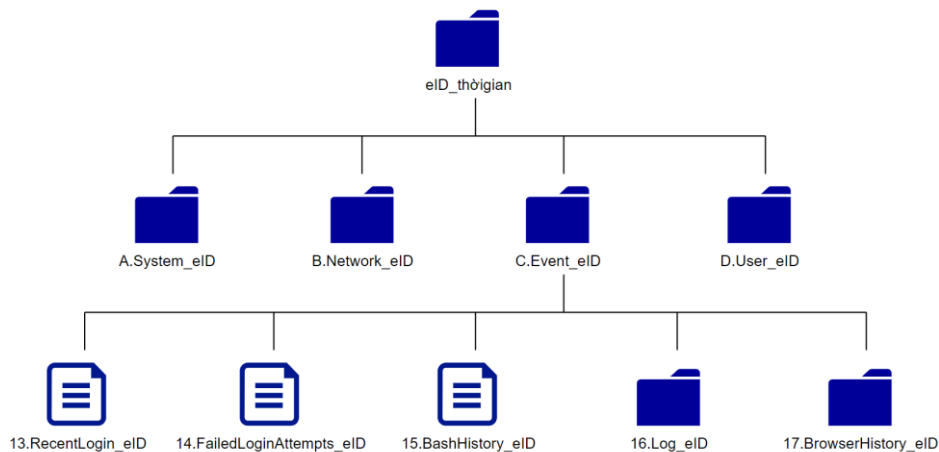
với *hostname* là tên máy tính, *epochTime* là số giây đếm tăng dần từ 01/01/1970 (UTC) 00:00:00 đến thời điểm bắt đầu thu thập. Chuỗi định danh này sẽ được sử dụng như phần đính kèm theo tên mặc định của tất cả các tệp tin và thư mục chứng cứ được tạo trong quá trình thu thập.

Ví dụ: *10.NetworkInterfaces\_kali\_1639107539*

- *10.NetworkInterfaces* là tên mặc định của tệp tin chứng cứ
- *kali* là tên máy tính
- *1639107539* là số giây đếm tăng dần từ 01/01/1970 (UTC) 00:00:00 đến thời điểm bắt đầu thu thập

Riêng thư mục cha bao gồm tất cả các thư mục, tệp tin kết quả khác sẽ có định dạng *eID* đính kèm thời gian bắt đầu thu thập.

Ví dụ: *kali\_1639107539\_10-12-21\_10:38:59*



Hình 4-8 Minh họa cấu trúc thư mục lưu trữ kết quả

### 4.1.3 Kết quả đạt được

Dưới đây là bảng mô tả thông tin về sản phẩm sau quá trình phát triển:

Bảng 4-4 Mô tả thông tin sản phẩm

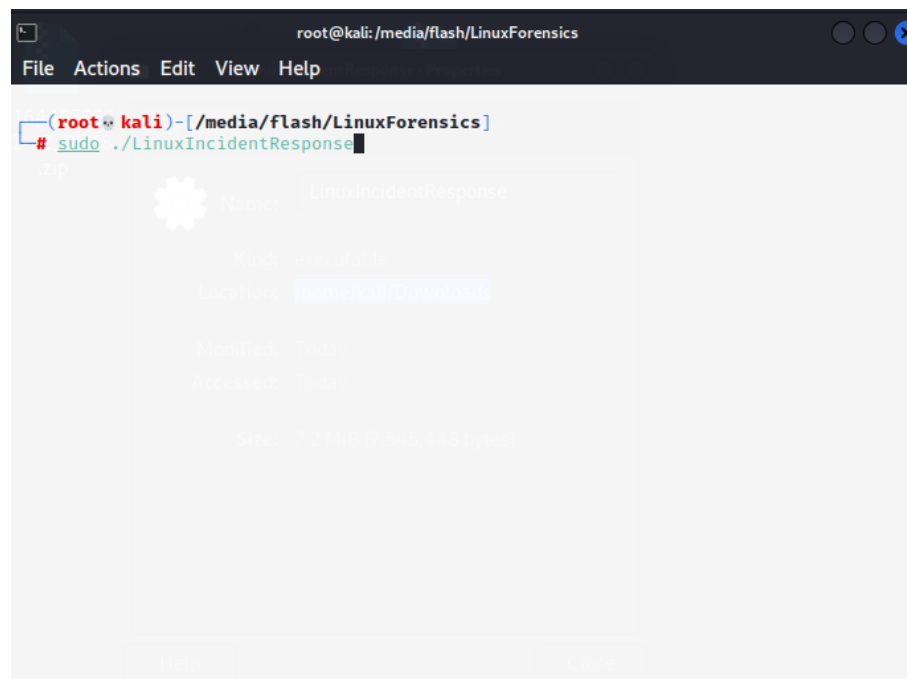
<b>Tên phần mềm</b>	LinuxIncidentResponse
<b>Số dòng code</b>	1123 dòng
<b>Dung lượng toàn bộ mã nguồn</b>	63.44 KB
<b>Dung lượng sản phẩm đóng gói</b>	7.369 KB

Sản phẩm đóng gói gồm một tệp thực thi duy nhất LinuxIncidentResponse thực hiện chức năng thu thập chứng cứ trên máy tính sử dụng hệ điều hành Linux gặp sự cố an toàn thông tin.

### 4.1.4 Minh họa các chức năng chính

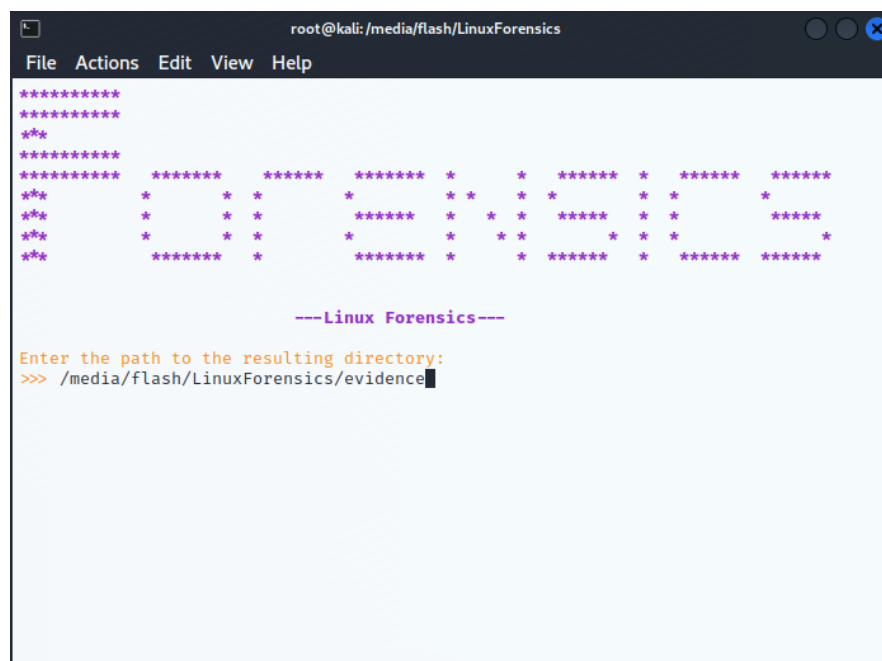
Sau khi xây dựng và triển khai thành công, phần này sẽ minh họa chi tiết chức năng chính của phần mềm.

Truy cập thư mục chứa phần mềm và khởi động cửa sổ dòng lệnh. Khởi động phần mềm dưới quyền quản trị viên:



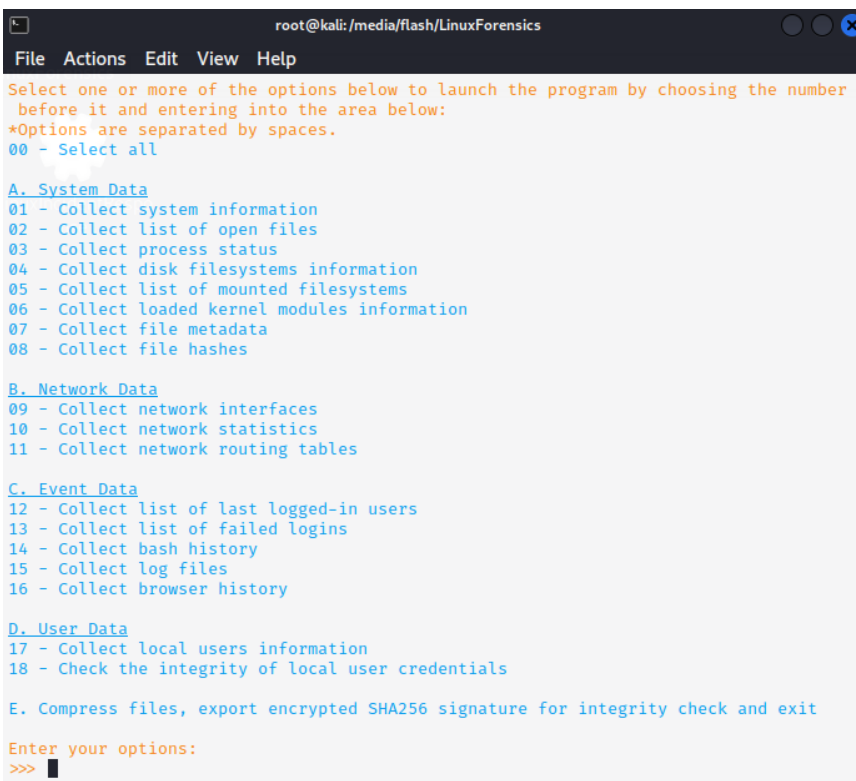
Hình 4-9 Khởi động phần mềm

Lựa chọn đường dẫn chứa thư mục kết quả:



Hình 4-10 Giao diện lựa chọn đường dẫn chứa thư mục kết quả

Màn hình giao diện hiển thị chức năng của phần mềm:



```
root@kali: /media/flash/LinuxForensics
File Actions Edit View Help
Select one or more of the options below to launch the program by choosing the number
before it and entering into the area below:
*Options are separated by spaces.
00 - Select all

A. System Data
01 - Collect system information
02 - Collect list of open files
03 - Collect process status
04 - Collect disk filesystems information
05 - Collect list of mounted filesystems
06 - Collect loaded kernel modules information
07 - Collect file metadata
08 - Collect file hashes

B. Network Data
09 - Collect network interfaces
10 - Collect network statistics
11 - Collect network routing tables

C. Event Data
12 - Collect list of last logged-in users
13 - Collect list of failed logins
14 - Collect bash history
15 - Collect log files
16 - Collect browser history

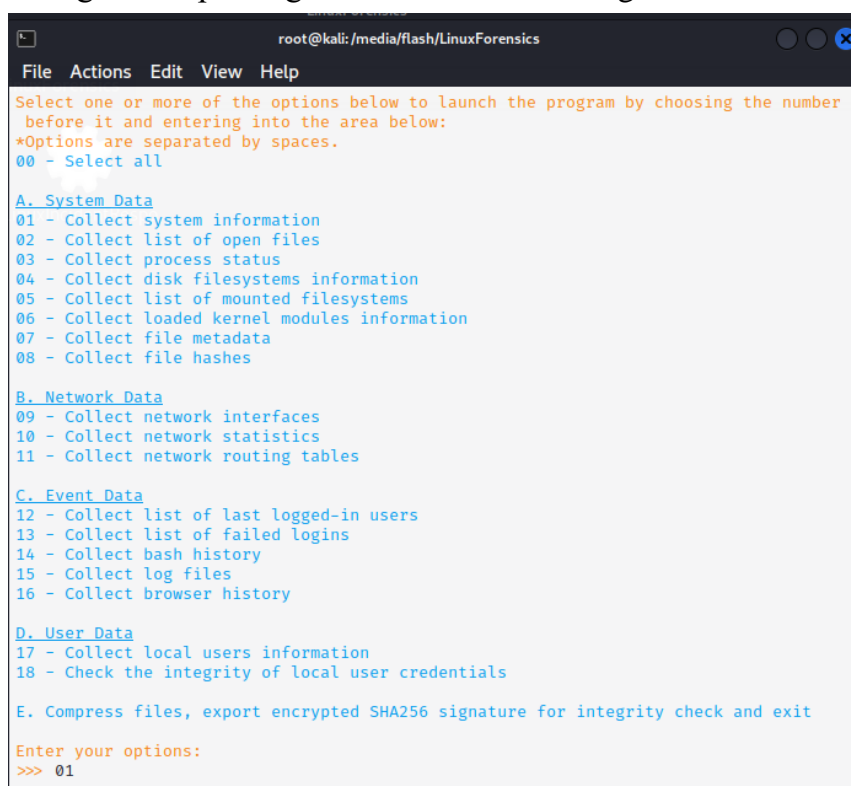
D. User Data
17 - Collect local users information
18 - Check the integrity of local user credentials

E. Compress files, export encrypted SHA256 signature for integrity check and exit

Enter your options:
>>> █
```

Hình 4-11 Giao diện hiển thị chức năng

Chọn chức năng thu thập thông tin cơ bản của hệ thống:



```
root@kali: /media/flash/LinuxForensics
File Actions Edit View Help
Select one or more of the options below to launch the program by choosing the number
before it and entering into the area below:
*Options are separated by spaces.
00 - Select all

A. System Data
01 - Collect system information
02 - Collect list of open files
03 - Collect process status
04 - Collect disk filesystems information
05 - Collect list of mounted filesystems
06 - Collect loaded kernel modules information
07 - Collect file metadata
08 - Collect file hashes

B. Network Data
09 - Collect network interfaces
10 - Collect network statistics
11 - Collect network routing tables

C. Event Data
12 - Collect list of last logged-in users
13 - Collect list of failed logins
14 - Collect bash history
15 - Collect log files
16 - Collect browser history

D. User Data
17 - Collect local users information
18 - Check the integrity of local user credentials

E. Compress files, export encrypted SHA256 signature for integrity check and exit

Enter your options:
>>> 01
```

Hình 4-12 Chọn chức năng thu thập thông tin hệ thống

Màn hình thông báo quá trình thu thập đã thành công:

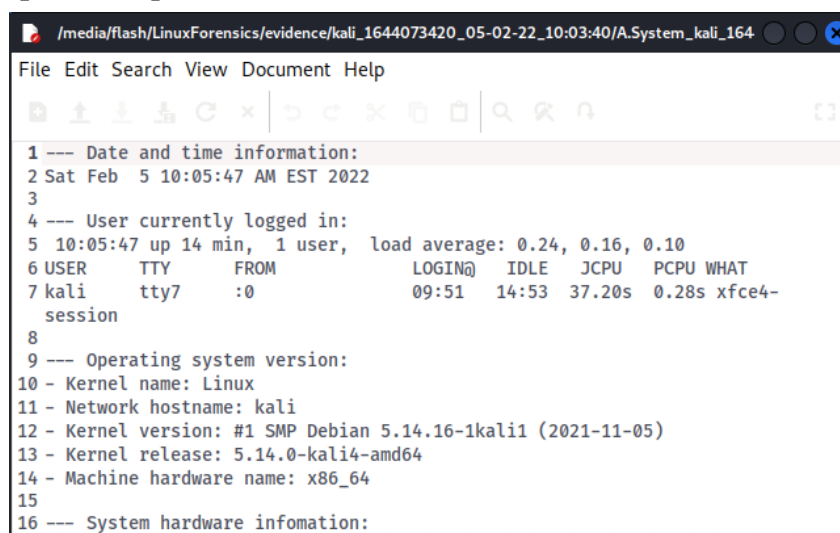
```
Enter your options:
>>> 01

Collecting system information (1)...
1.SystemInfomation_kali_1644073420.txt export successfully!

Do you want to do further forensics? If you choose NO, the system will compress the
forensic results and terminate the program. (y/n)
>>> █
```

Hình 4-13 Thông báo thu thập thành công

Kiểm tra tệp tin kết quả:

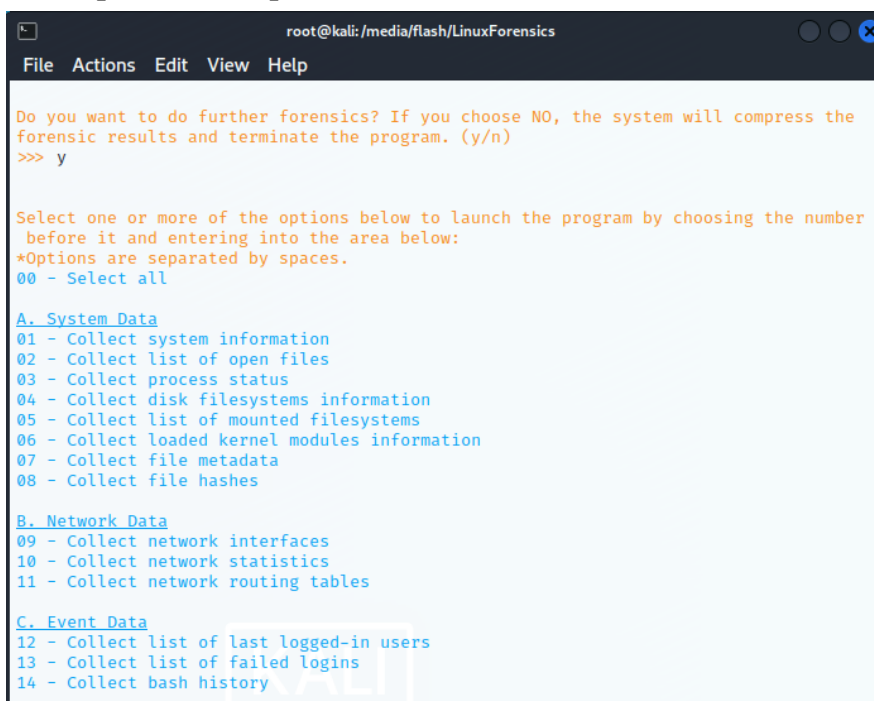


The screenshot shows a file manager window with the title bar indicating the file path: /media/flash/LinuxForensics/evidence/kali\_1644073420\_05-02-22\_10:03:40/A.System\_kali\_164. The window contains the following text:

```
1 --- Date and time information:
2 Sat Feb 5 10:05:47 AM EST 2022
3
4 --- User currently logged in:
5 10:05:47 up 14 min, 1 user, load average: 0.24, 0.16, 0.10
6 USER      TTY      FROM          LOGIN@   IDLE   JCPU   PCPU   WHAT
7 kali      tty7      :0            09:51    14:53  37.20s 0.28s xfce4-
8 session
9 --- Operating system version:
10 - Kernel name: Linux
11 - Network hostname: kali
12 - Kernel version: #1 SMP Debian 5.14.16-1kali1 (2021-11-05)
13 - Kernel release: 5.14.0-kali4-amd64
14 - Machine hardware name: x86_64
15
16 --- System hardware information:
```

Hình 4-14 Kết quả thu thập thông tin hệ thống

Nhấn “y” để tiếp tục thu thập. Màn hình hiển thị lại danh sách chức năng:



The screenshot shows a terminal window with the title bar indicating the user is root@kali and the directory is /media/flash/LinuxForensics. The window contains the following text:

```
Do you want to do further forensics? If you choose NO, the system will compress the
forensic results and terminate the program. (y/n)
>>> y

Select one or more of the options below to launch the program by choosing the number
before it and entering into the area below:
*Options are separated by spaces.
00 - Select all

A. System Data
01 - Collect system information
02 - Collect list of open files
03 - Collect process status
04 - Collect disk filesystems information
05 - Collect list of mounted filesystems
06 - Collect loaded kernel modules information
07 - Collect file metadata
08 - Collect file hashes

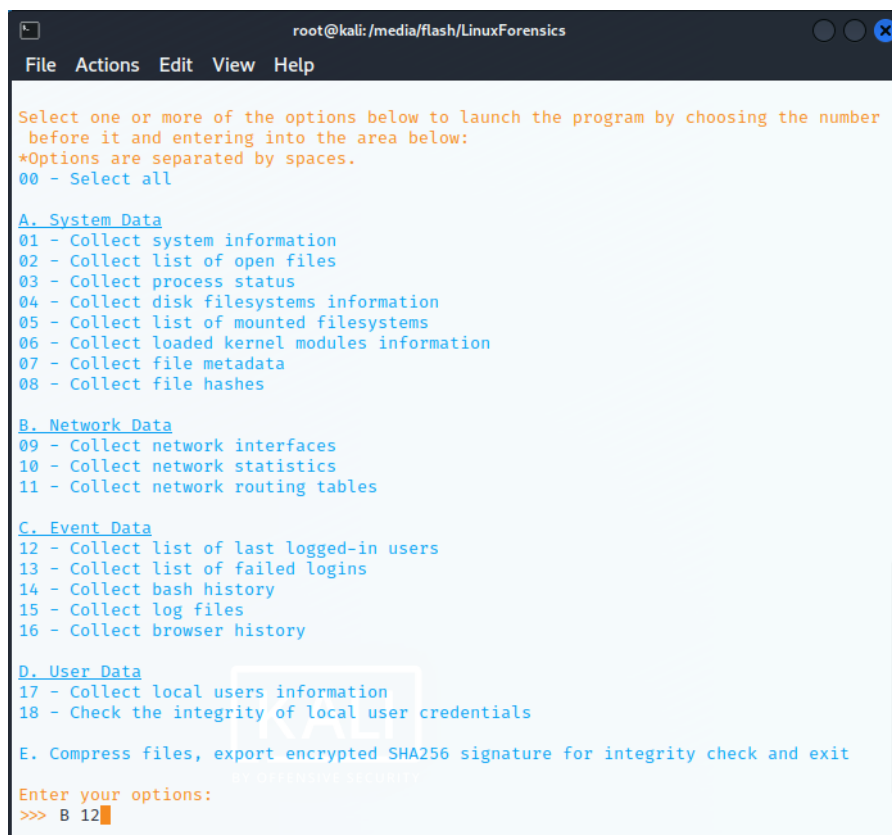
B. Network Data
09 - Collect network interfaces
10 - Collect network statistics
11 - Collect network routing tables

C. Event Data
12 - Collect list of last logged-in users
13 - Collect list of failed logins
14 - Collect bash history
```

Hình 4-15 Màn hình tiếp tục thu thập



Chọn đồng thời nhiều chức năng, thu thập tất cả chứng cứ mạng và danh sách tài khoản người dùng mới đăng nhập:



```
root@kali: /media/flash/LinuxForensics
File Actions Edit View Help

Select one or more of the options below to launch the program by choosing the number
before it and entering into the area below:
*Options are separated by spaces.
00 - Select all

A. System Data
01 - Collect system information
02 - Collect list of open files
03 - Collect process status
04 - Collect disk filesystems information
05 - Collect list of mounted filesystems
06 - Collect loaded kernel modules information
07 - Collect file metadata
08 - Collect file hashes

B. Network Data
09 - Collect network interfaces
10 - Collect network statistics
11 - Collect network routing tables

C. Event Data
12 - Collect list of last logged-in users
13 - Collect list of failed logins
14 - Collect bash history
15 - Collect log files
16 - Collect browser history

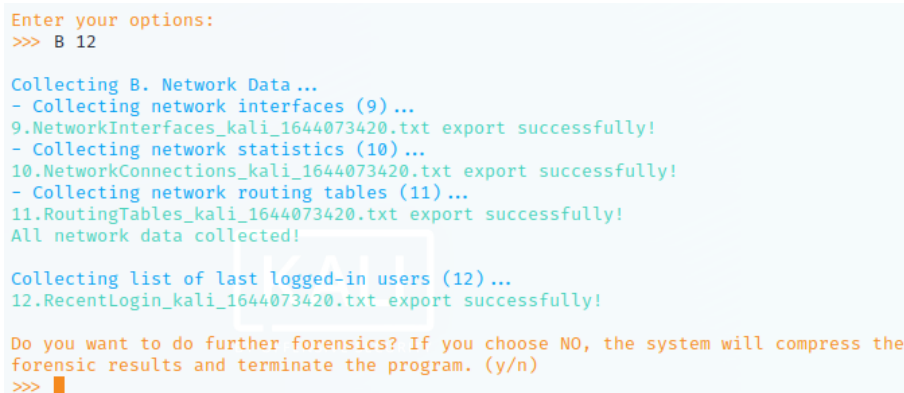
D. User Data
17 - Collect local users information
18 - Check the integrity of local user credentials

E. Compress files, export encrypted SHA256 signature for integrity check and exit

Enter your options:
>>> B 12
```

Hình 4-16 Chọn đồng thời nhiều chức năng

Màn hình thông báo quá trình thu thập thành công:



```
Enter your options:
>>> B 12

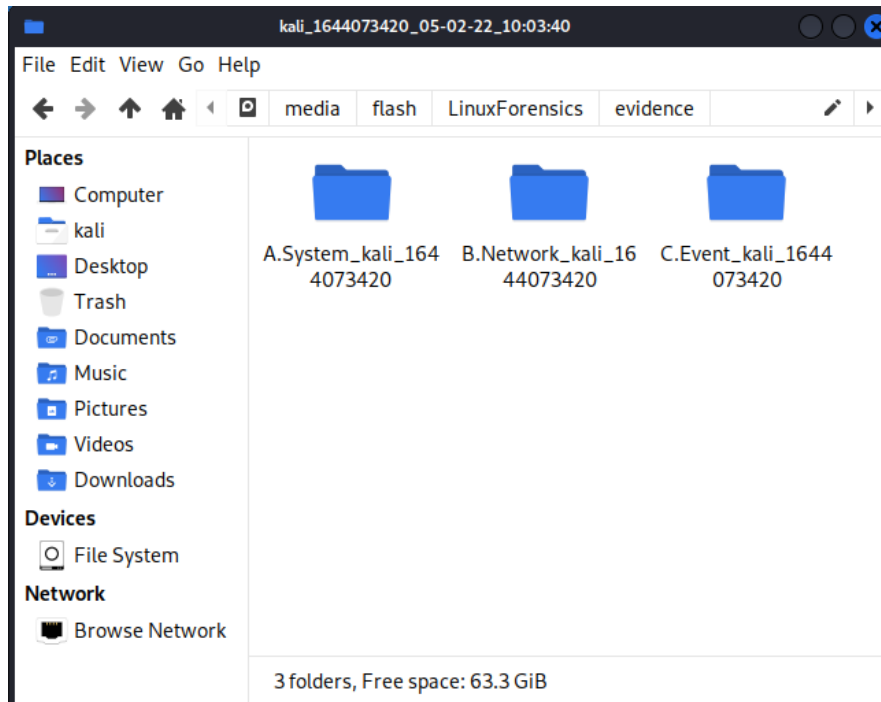
Collecting B. Network Data...
- Collecting network interfaces (9)...
9.NetworkInterfaces_kali_1644073420.txt export successfully!
- Collecting network statistics (10)...
10.NetworkConnections_kali_1644073420.txt export successfully!
- Collecting network routing tables (11)...
11.RoutingTables_kali_1644073420.txt export successfully!
All network data collected!

Collecting list of last logged-in users (12)...
12.RecentLogin_kali_1644073420.txt export successfully!

Do you want to do further forensics? If you choose NO, the system will compress the
forensic results and terminate the program. (y/n)
>>>
```

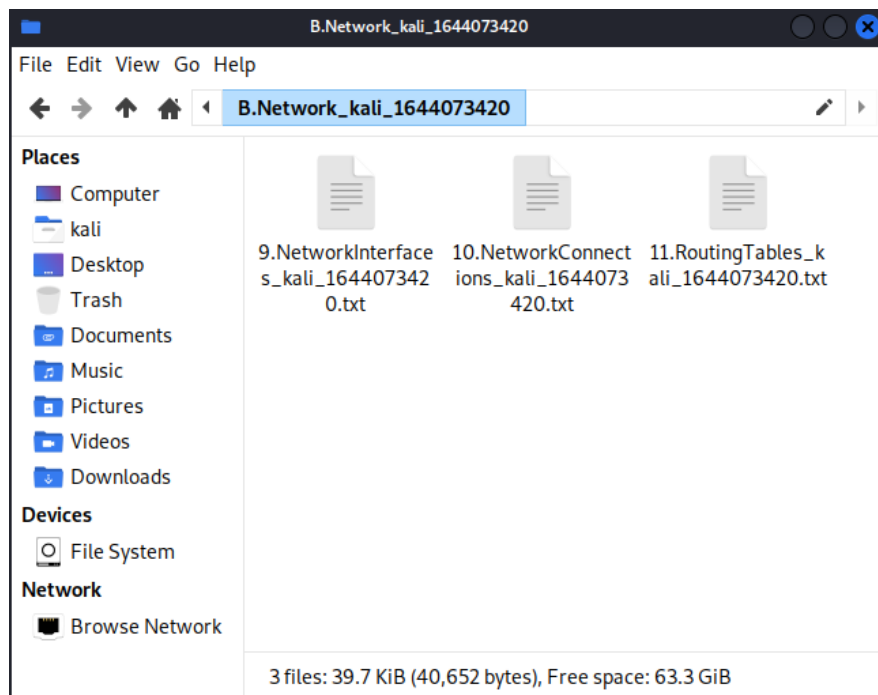
Hình 4-17 Thông báo thu thập thành công

Kiểm tra thư mục kết quả lúc này:



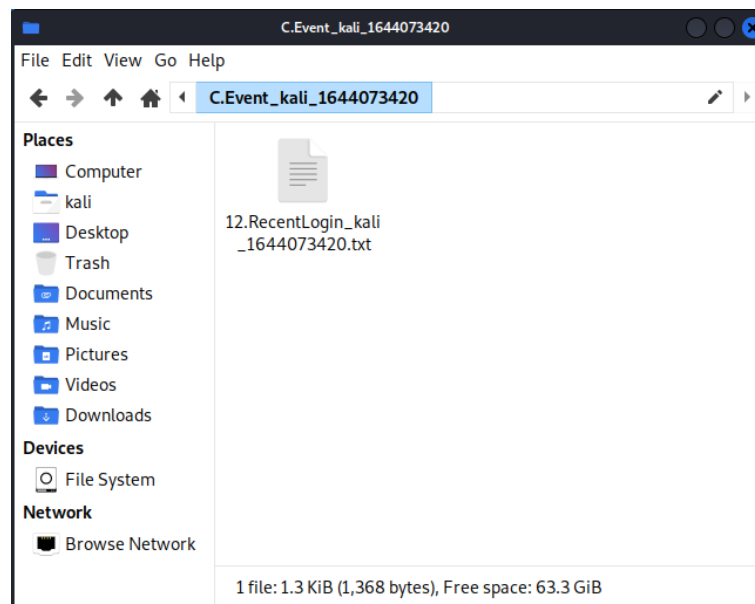
Hình 4-18 Thư mục kết quả

Kiểm tra thư mục chứa chứng cứ mạng:



Hình 4-19 Thư mục chứng cứ mạng

Kiểm tra thư mục vết sự kiện:



Hình 4-20 Thư mục vết sự kiện

Chọn “n” để kết thúc phiên thu thập. Lưu lại mã hóa SHA256 được thông báo trên màn hình:

```
Do you want to do further forensics? If you choose NO, the system will compress the
forensic results and terminate the program. (y/n)
>>> n

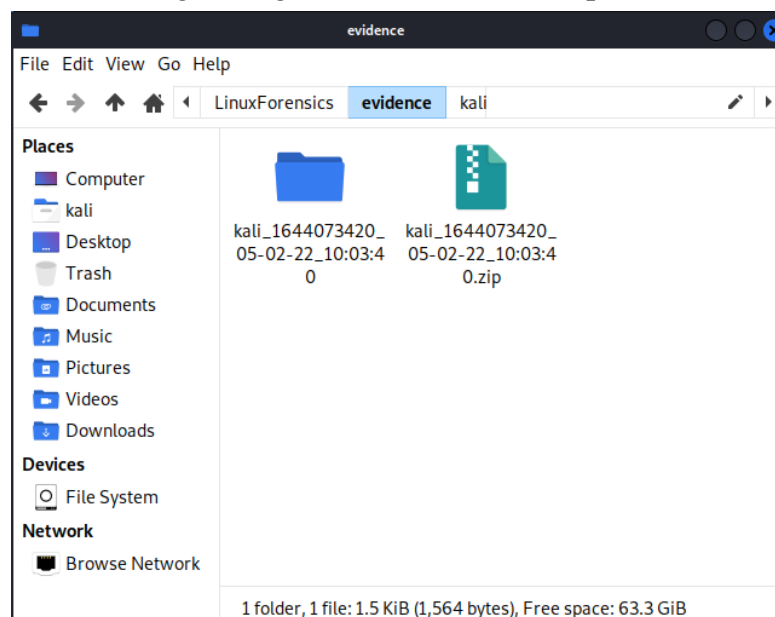
Compressing files...
Compress files successfully!
Exporting encrypted SHA256 signature...
Export encrypted SHA256 signature successfully!
SHA256 Signature: b3cbdcc5786817365a4013d7fad4f76807e5cb6fb75ee5fd68f20533d1f6105c

Exit ...

(root@kali)-[/media/flash/LinuxForensics]
```

Hình 4-21 Kết thúc phiên thu thập

Kiểm tra tệp tin nén trong đường dẫn tới thư mục kết quả:



Hình 4-22 Tệp tin nén chứng cứ thu thập

## 4.2 Kiểm thử

### 4.2.1 Kiểm thử chức năng thu thập tất cả chứng cứ

*Bảng 4-5 Kiểm thử chức năng thu thập tất cả chứng cứ*

Mô tả	Kiểm thử chức năng thu thập tất cả chứng cứ
Kỹ thuật kiểm thử	Kiểm thử hộp đen
Môi trường kiểm thử	Hệ điều hành Ubuntu 20.04
Bước thực hiện	<ul style="list-style-type: none"><li>- Khởi động phần mềm bằng quyền quản trị viên</li><li>- Nhập đường dẫn tới thư mục lưu trữ</li><li>- Chọn chức năng thu thập tất cả chứng cứ</li></ul>
Kết quả	<ul style="list-style-type: none"><li>- Xảy ra lỗi khi thu thập lịch sử trình duyệt Google Chrome</li><li>- Các chứng cứ khác thu thập đầy đủ</li></ul>

Kết quả: 9/10

Phân tích nguyên nhân và đưa ra hướng giải quyết: Lỗi thu thập lịch sử trình duyệt Google Chrome do trình duyệt đang hoạt động trên máy tính thực hiện thu thập. Cần tắt trình duyệt trước khi thực hiện thu thập lịch sử.

Sau khi giải quyết: chức năng thu thập lịch sử trình duyệt Google Chrome thực hiện bình thường, không có lỗi xảy ra.

### 4.2.2 Kiểm thử chức năng đảm bảo tính toàn vẹn của chứng cứ

*Bảng 4-6 Kiểm thử chức năng đảm bảo tính toàn vẹn của chứng cứ*

Mô tả	Kiểm thử chức năng đảm bảo tính toàn vẹn của chứng cứ
Kỹ thuật kiểm thử	Kiểm thử hộp đen
Môi trường kiểm thử	Hệ điều hành Ubuntu 20.04
Bước thực hiện	<ul style="list-style-type: none"><li>- Khởi động phần mềm bằng quyền quản trị viên</li><li>- Nhập đường dẫn tới thư mục lưu trữ</li><li>- Chọn chức năng thu thập tất cả chứng cứ và bảo vệ tính toàn vẹn chứng cứ sau khi thu thập</li></ul>
Kết quả	Các chứng cứ sau khi thu thập được nén thành tệp tin có phần mở rộng “.zip” và xuất mã hóa SHA256 của tệp tin trên màn hình

Kết quả: 10/10

Chức năng thực hiện thành công, không có lỗi gì xảy ra.

### 4.2.3 Kiểm thử phần mềm trên hệ điều hành Kali Linux

Bảng 4-7 Kiểm thử phần mềm trên hệ điều hành Kali Linux

Mô tả	Kiểm thử phần mềm trên hệ điều hành Kali Linux
Kỹ thuật kiểm thử	Kiểm thử hộp đen
Môi trường kiểm thử	Hệ điều hành Kali Linux 2021.4
Bước thực hiện	<ul style="list-style-type: none"><li>- Khởi động phần mềm bằng quyền quản trị viên</li><li>- Nhập đường dẫn tới thư mục lưu trữ</li><li>- Chọn chức năng thu thập tất cả chứng cứ và bảo vệ tính toàn vẹn chứng cứ sau khi thu thập</li></ul>
Kết quả	Thu thập thành công tất cả các loại chứng cứ yêu cầu, các chứng cứ sau khi thu thập được nén thành tệp tin có phần mở rộng “.zip” và xuất mã hóa SHA256 của tệp tin trên màn hình

Kết quả: 10/10

Tất cả các chức năng thực hiện thành công, không có lỗi gì xảy ra.

### 4.3 Triển khai

Ở phần này, chúng ta sẽ đi vào quá trình tạo tệp tin thực thi phần mềm dựa trên mã nguồn Python đã xây dựng.

Môi trường cài đặt sử dụng trong ví dụ này là hệ điều hành Ubuntu 20.04 LTS. Phiên bản hệ điều hành này đã được cài đặt sẵn Python 3.8.10. Đây cũng là phiên bản Python được sử dụng để phát triển phần mềm.

**Bước 1:** Mở cửa sổ dòng lệnh. Cài đặt trình quản lý thư viện PIP của Python bằng các câu lệnh sau:

```
$ sudo apt update
```

```
$ sudo apt install python3-pip
```

**Bước 2:** Cài đặt PyInstaller bằng câu lệnh:

```
$ pip install pyinstaller
```

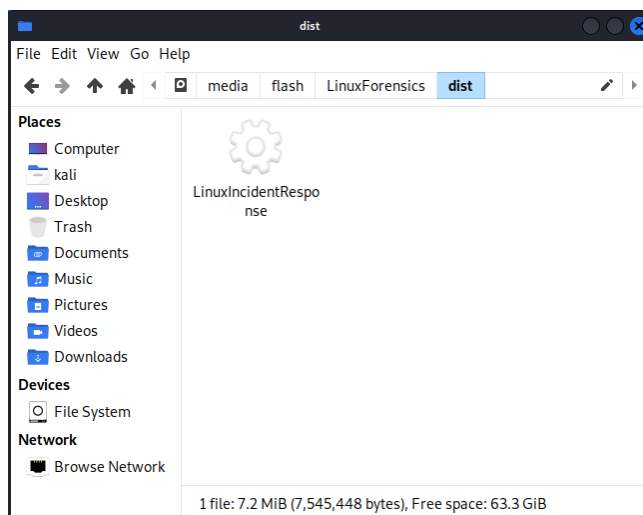
**Bước 3:** Truy cập thư mục chính của phần mềm, nơi chứa tệp tin execute.py. Mở cửa sổ dòng lệnh tại đây và chạy câu lệnh sau:

```
$ pyinstaller -D -F -n LinuxIncidentResponse -c "execute.py"
```

```
kooriboh@MingdiKubiboh:~/Documents/LinuxForensics$ pyinstaller -D -F -n LinuxIncidentResponse -c "execute.py"
44 INFO: PyInstaller: 4.8
44 INFO: Python: 3.8.10
55 INFO: Platform: Linux-5.13.0-28-generic-x86_64-with-glibc2.29
55 INFO: wrote /home/kooriboh/Documents/LinuxForensics/LinuxIncidentResponse.spec
57 INFO: UPX is not available.
59 INFO: Extending PYTHONPATH with paths
['/home/kooriboh/Documents/LinuxForensics']
185 INFO: checking Analysis
185 INFO: Building Analysis because Analysis-00.toc is non-existent
186 INFO: Initializing module dependency graph...
187 INFO: Caching module graph hooks...
192 INFO: Analyzing base_library.zip ...
```

Hình 4-23 Thực thi câu lệnh pyinstaller

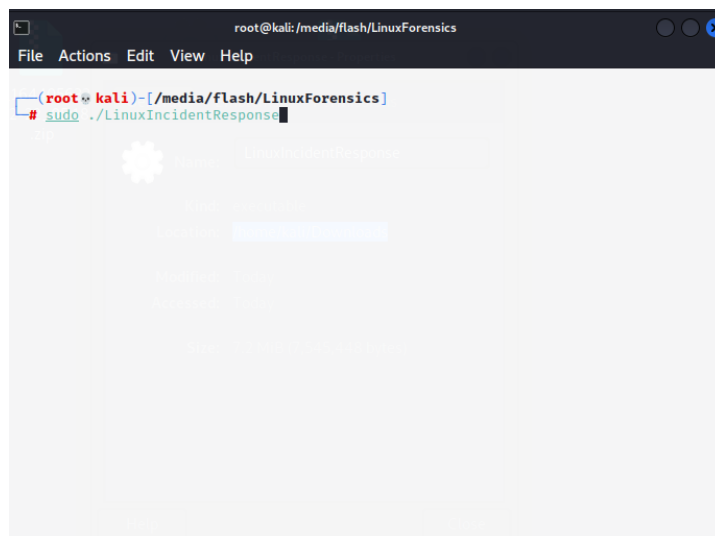
Sau khi câu lệnh được thực thi, một thư mục mới có tên “dist” sẽ xuất hiện tại thư mục chính của phần mềm. Truy cập thư mục này, ta thấy tệp tin LinuxIncidentResponse xuất hiện.



Hình 4-24 Tệp tin thực thi LinuxIncidentResponse

Tệp tin thực thi có thể được chạy trên các máy tính sử dụng hệ điều hành Linux, yêu cầu cần có GLIBC phiên bản từ 2.29 trở lên (do yêu cầu của thư viện từ Python 3.8 và PyInstaller).

Để chạy tệp tin thực thi, người dùng cần mở cửa sổ dòng lệnh tại thư mục chứa tệp tin và sử dụng câu lệnh: `sudo ./LinuxIncidentResponse`.



Hình 4-25 Khởi động phần mềm

## CHƯƠNG 5. CÁC GIẢI PHÁP VÀ ĐÓNG GÓP NỔI BẬT

Sau khi đã mô tả chi tiết quá trình phát triển và triển khai ứng dụng, chương này tôi sẽ trình bày những giải pháp và đóng góp nổi bật của mình trong quá trình thực hiện đồ án tốt nghiệp. Đây cũng là những tính năng nổi bật, tạo nên sự khác biệt giữa phần mềm được xây dựng với các công cụ thu thập chứng cứ hiện tại.

### 5.1 Thu thập hầu hết chứng cứ trên máy tính cần điều tra

Phần mềm đã thực hiện chức năng thu thập và đảm bảo tính toàn vẹn cho một lượng lớn chứng cứ có mặt trên một máy tính gặp sự cố an toàn thông tin. Dưới đây là chi tiết các loại chứng cứ mà phần mềm thu thập.

#### 5.1.1 Chứng cứ hệ thống

- Thu thập thông tin cơ bản của hệ thống: phần mềm thu thập thông tin về ngày giờ thực hiện, tài khoản người dùng đang đăng nhập, và thông tin cơ bản của hệ điều hành trên máy tính thực hiện thu thập.
- Thu thập thông tin về các tệp tin trên hệ thống đang hoặc đã được mở bởi các tiến trình: phần mềm thu thập chi tiết thông tin các tệp được mở bởi tiến trình.
- Thu thập thông tin các tiến trình đang chạy: phần mềm thu thập thông tin chi tiết các tiến trình đang chạy trong tất cả các tài khoản người dùng đang sử dụng trên máy tính. Đây là một dữ liệu quan trọng trong việc điều tra các tiến trình khả nghi, ví dụ như các tiến trình được chạy bằng quyền quản trị viên mà không được sự cho phép của người dùng, hay những tài khoản hệ thống không được phép đăng nhập đang thực hiện chạy tiến trình.
- Thu thập thông tin dung lượng ổ cứng đang sử dụng và còn sẵn trong hệ thống: phần mềm thu thập thông tin đầy đủ về việc sử dụng không gian đĩa cứng còn sẵn và được sử dụng của hệ thống tập tin trên hệ thống Linux, là thông tin quan trọng trong việc điều tra sự cố. Ví dụ như kiểm tra xem phân vùng nào đầy không rõ lý do, phân vùng nào được kết nối trái phép.
- Thu thập thông tin các phân vùng được mount trên hệ thống : phần mềm thu thập thông tin chi tiết về các quyền và tùy chọn được sử dụng để gắn kết mỗi phân vùng.
- Thu thập thông tin các module đang được sử dụng bởi Linux kernel: phần mềm cung cấp chi tiết các module kernel được cài đặt, từ đó có thể kiểm tra sự tồn tại của các trojan trong hệ thống nếu có, hoặc những driver được cài đặt mà không có sự cho phép của người dùng.
- Thu thập thông tin chi tiết của các tệp tin trong hệ thống: phần mềm thu thập các thông tin chi tiết về các tệp tin trên máy tính, bao gồm các vết thời gian, quyền thực thi, chủ sở hữu tệp tin và kích thước của tệp.
- Thu thập hàm băm của các tệp tin trong hệ thống: phần mềm thu thập mã hóa SHA1 của tất cả các tệp tin trong hệ thống nhằm phân tích nhanh tệp tin trong bước điều tra sau này.

### 5.1.2 Chứng cứ mạng

- Thu thập thông tin cấu hình giao diện hệ thống mạng: phần mềm thu thập các thông tin liên quan đến cấu hình hệ thống mạng trên máy tính, ví dụ như địa chỉ được gán cho giao diện mạng, mặt nạ mạng, ...
- Thu thập thông tin các cổng mạng được mở: phần mềm thu thập thông tin chi tiết về các cổng mạng đang mở, từ đó sử dụng để phân tích điều tra các kết nối khả nghi.
- Thu thập thông tin bảng định tuyến: phần mềm thu thập lại bảng định tuyến trên máy tính cần điều tra. Đây là một thông tin quan trọng sử dụng để phân tích kết nối mạng, ví dụ kiểm tra lưu lượng truy cập của người dùng có đang được định tuyến lại và giám sát bởi kẻ tấn công, hay có bất kỳ gateway nào đã được thay đổi.

### 5.1.3 Vết sự kiện

- Thu thập thông tin người dùng mới đăng nhập: phần mềm thu thập thông tin chi tiết về thời điểm khởi động hệ thống cũng như thời gian, tài khoản sử dụng để đăng nhập, đăng xuất.
- Thu thập thông tin về những lần đăng nhập thất bại: phần mềm thu thập danh sách những lần đăng nhập thất bại trên máy tính.
- Thu thập lịch sử sử dụng dòng lệnh: phần mềm thu thập danh sách những câu lệnh đã được sử dụng ở mỗi tài khoản người dùng. Đây là chức năng được sử dụng thường xuyên nhất trên Linux, và thông tin thu thập được từ chức năng này sẽ cung cấp vai trò quan trọng trong quá trình điều tra sự cố.
- Thu thập các tệp tin vết sự kiện: phần mềm thu thập các thư mục và tệp tin trong thư mục vết sự kiện /var/log của hệ điều hành Linux trên máy tính.
- Thu thập lịch sử trình duyệt: phần mềm thu thập lịch sử duyệt web chi tiết trên các trình duyệt phổ biến của hệ điều hành Linux như Google Chrome và Mozilla Firefox, sau đó xuất ra tệp tin .csv để thuận tiện cho việc điều tra.

### 5.1.4 Thông tin tài khoản người dùng

- Thu thập thông tin tài khoản người dùng có trong hệ thống: phần mềm thu thập tệp tin /etc/passwd và /etc/shadow - những tệp tin chứa thông tin tài khoản người dùng của hệ điều hành Linux.
- Kiểm tra tính toàn vẹn của thông tin tài khoản người dùng có trong hệ thống: phần mềm phân tích dữ liệu từ tệp tin /etc/passwd và /etc/shadow, từ đó xác định được những tài khoản người dùng khả nghi, có dấu hiệu đăng nhập trái phép vào máy tính.

**Đánh giá:** đây là phần khó khăn nhất trong quá trình phát triển phần mềm, do yêu cầu cần phải hiểu rõ cấu trúc tệp hệ thống của Linux và có khả năng phối hợp sử dụng những câu lệnh được tích hợp sẵn trên hệ điều hành để tăng hiệu suất thu thập. Tất cả những thông tin được thu thập trên đây đều là những dữ liệu quan trọng, sẽ ảnh hưởng lớn đến quá trình điều tra sau này.



## 5.2 Phát triển những chức năng đặc trưng

Phần mềm không chỉ thực hiện chức năng thu thập thông thường mà còn có những tính năng đặc biệt, hỗ trợ quá trình điều tra sự cố sau này.

### 5.2.1 Thu thập hàm băm của các tệp tin trong hệ thống

Đây là chức năng mở rộng của phần mềm, không có trên bất kỳ phần mềm thu thập chứng cứ nào khác trên hệ điều hành Linux. Chức năng thực hiện thu thập mã hóa SHA1 của tất cả các tệp tin trong hệ thống nhằm phân tích nhanh tệp tin trong các bước điều tra sau này.

Có rất nhiều các cơ sở dữ liệu hàm băm trên Internet giúp ta có thể kiểm tra các tệp tin trong hệ thống có thuộc danh sách những tệp tin đã được chứng nhận an toàn hoặc nguy hiểm hay không. Thu thập và kiểm tra những hàm băm này có thể giúp ta loại trừ một lượng lớn những tệp tin an toàn khỏi diện điều tra, và mặc dù đây không phải là cách tốt nhất để tìm kiếm những tệp tin độc hại trong hệ thống, nhưng nó cũng nhanh hơn rất nhiều so với việc phân tích từng tệp tin bằng phương pháp dịch ngược hoặc sử dụng phần mềm diệt virus thông thường. Vì vậy, kiểm tra hàm băm có thể sử dụng như bước khởi đầu để tìm manh mối của mã độc trong quá trình điều tra.

**Đánh giá:** chức năng thu thập hàm băm của phần mềm sẽ hỗ trợ rất lớn trong quá trình điều tra sự cố sau này, tuy nhiên, việc xuất hàm băm của tất cả các tệp tin trong hệ thống có thể sẽ tiêu tốn rất nhiều thời gian thực hiện. Vì vậy, cần cân nhắc việc không nên sử dụng chức năng này khi thu thập chứng cứ trên các hệ thống lớn.

### 5.2.2 Thu thập lịch sử trình duyệt và xuất kết quả trực quan

Hầu hết các phần mềm thực hiện chức năng thu thập chứng cứ hiện này đều thu thập lịch sử duyệt web bằng cách sao chép toàn bộ cơ sở dữ liệu của trình duyệt, chuyên viên điều tra sẽ cần sử dụng thêm công cụ và các bước phân tích khác để có thể đọc những dữ liệu này.

Phần mềm sẽ thực hiện đọc trực tiếp thông tin lịch sử duyệt web từ tệp tin cơ sở dữ liệu của mỗi trình duyệt, và xuất dữ liệu này thành tệp tin có phần mở rộng “.csv” - định dạng có thể dễ dàng đọc bởi Excel hay các công cụ bảng tính thông dụng khác, thuận tiện cho quá trình điều tra sau này.

**Đánh giá:** đây là chức năng hữu ích, hỗ trợ quá trình điều tra sau thu thập.

### 5.2.3 Kiểm tra tính toàn vẹn của thông tin tài khoản người dùng có trong hệ thống

Thông tin về tài khoản người dùng trên Linux được lưu trữ trong tệp tin passwd và shadow trong thư mục /etc của hệ thống. Ban đầu, tất cả thông tin về tài khoản người dùng đều được lưu trữ trong /etc/passwd. Tuy nhiên, nó được coi là vấn đề bảo mật khi lưu trữ thông tin mật khẩu trong một tệp tin có thể đọc được bởi mọi người dùng. Do vậy, mật khẩu được băm trong /etc/passwd đã được thay thế bằng một chữ “x” duy nhất, biểu thị rằng mật khẩu băm tương ứng phải được tra cứu trong tệp tin /etc/shadow. Tác dụng phụ của việc này là mật khẩu băm trong

/etc/passwd vẫn được hỗ trợ, và tất cả cài đặt trong /etc/passwd có thể ghi đè thông tin đăng nhập trong /etc/shadow [3].

Phần mềm tham khảo một số hàm Python sử dụng trong cuốn *Mastering Python Forensics* [3] của Michael Spreitzenbarth và Johann Uhrmann để xây dựng chức năng này, kiểm tra tính toàn vẹn của thông tin tài khoản người dùng có trong hệ thống thông qua các yếu tố:

- Kiểm tra sự tồn tại của nhiều tài khoản sử dụng chung ID người dùng: bằng cách tạo một tài khoản mới có chung ID với tài khoản sẵn có, kẻ tấn công có thể tạo một bí danh với mật khẩu riêng và sử dụng bí danh này để truy cập vào tài khoản chung ID. Phần mềm sẽ thực hiện kiểm tra sự tồn tại của những tài khoản này.
- Kiểm tra sự tồn tại của tài khoản chỉ xuất hiện trong /etc/passwd hoặc /etc/shadow: trong quá trình hoạt động bình thường, thông tin tài khoản của hai tệp tin này sẽ luôn khớp nhau. Nếu có tài khoản chỉ xuất hiện tại một trong hai tệp tin, điều này báo hiệu việc quản lý tài khoản người dùng của hệ điều hành đã bị can thiệp. Phần mềm sẽ thực hiện kiểm tra sự tồn tại của những tài khoản này.
- Kiểm tra sự tồn tại của tài khoản không có mật khẩu: thông thường, các tài khoản trong hệ thống đều có mật khẩu. Ngoài ra, tất cả các mật khẩu băm phải nằm trong tệp /etc/shadow và tất cả các thông tin tài khoản trong /etc/passwd đều phải tham chiếu đến thông tin tương ứng trong /etc/shadow. Phần mềm sẽ thực hiện kiểm tra sự tồn tại của những tài khoản không thỏa mãn những điều kiện trên.
- Kiểm tra các hàm băm bất thường: Linux sử dụng chung thuật toán băm với mọi mật khẩu trong tệp /etc/shadow. Nếu có thuật toán sai lệch thì rất có thể quy trình mã hóa của hệ thống đã bị xâm phạm. Phần mềm sẽ thực hiện kiểm tra điều này.

**Đánh giá:** đây là một chức năng hữu ích, không tốn nhiều thời gian để thực thi, có thể sử dụng thường xuyên trong các phiên thu thập để hỗ trợ quá trình điều tra.

### **5.3 Xây dựng một phần mềm nhỏ gọn, tương thích với nhiều bản phân phối Linux**

Một trong những yêu cầu quan trọng nhất của việc thu thập chứng cứ kỹ thuật số là phải tối thiểu hóa những xáo trộn, thay đổi gây ra đối với máy tính được thu thập. Để thực hiện điều này, phần mềm được phát triển phải đáp ứng mục tiêu không yêu cầu cài đặt để sử dụng. Quá trình cài đặt có thể tạo ra những thay đổi đáng kể đến hệ thống, tạo ra nhiều tệp tin, thư mục mới và ảnh hưởng lớn đến chất lượng chứng cứ được thu thập.

Với sự trợ giúp của PyInstaller, phần mềm có thể hoạt động với chỉ một tệp tin thực thi được lưu trữ trên thiết bị thu thập. Tất cả các tệp tin, thư mục chứng cứ sẽ được tạo và lưu trữ trên thiết bị thu thập, đảm bảo hạn chế tối thiểu dấu vết để lại trên bộ nhớ và hệ thống của máy tính cần thu thập.

Bên cạnh đó, nhờ hoạt động hoàn toàn trên giao diện dòng lệnh, phần mềm cũng tương thích với số lượng lớn các bản phân phối Linux hiện nay có mặt trên thị trường, đảm bảo tính tiện lợi trong quá trình sử dụng.

**Đánh giá:** phần mềm có khả năng tương thích tốt, đảm bảo các yêu cầu đặt ra đối với một sản phẩm sử dụng trong lĩnh vực an toàn thông tin. Tuy nhiên, do sử dụng PyInstaller để triển khai phần mềm, sẽ có một số bản phân phối Linux không tương thích với phần mềm do sở hữu GLIBC phiên bản cũ, không hỗ trợ các thư viện Python mới. Cần tìm giải pháp khắc phục vấn đề này.

## CHƯƠNG 6. KẾT LUẬN

### 6.1 Kết luận

Thực nghiệm tiến hành trong quá trình kiểm thử đã khẳng định toàn bộ yêu cầu đặt ra của đề án tốt nghiệp đã được đáp ứng đầy đủ. Phần mềm đã hoàn thành mục tiêu ban đầu, có đầy đủ khả năng để đưa vào sử dụng thực tế và hỗ trợ quá trình thu thập chứng cứ từ máy tính sử dụng hệ điều hành Linux gặp sự cố an toàn thông tin.

Trong quá trình thực hiện đề án dưới sự hướng dẫn của PGS.TS. Nguyễn Linh Giang, tôi đã thu được nhiều bài học mới, trong đó có thể kể đến (i) kiến thức về hệ điều hành Linux, (ii) kỹ năng thu thập chứng cứ kỹ thuật số và (iii) khả năng lập trình, xây dựng một phần mềm hoàn chỉnh. Mặc dù chưa thật sự hoàn thành xuất sắc những mục tiêu đề ra do sự hạn chế về kiến thức và kinh nghiệm, nhưng tôi vẫn cảm thấy hài lòng với những trải nghiệm đã đạt được trong suốt quá trình thực hiện đề án. Đó là món quà quý báu giúp tôi vững vàng hơn trong con đường học hỏi và trau dồi kiến thức phía trước.

### 6.2 Hướng phát triển của đề án trong tương lai

Do phạm vi của đề án có hạn, nên vẫn còn nhiều chức năng, thành phần tôi mong muốn phát triển nhưng chưa đủ khả năng để thực hiện. Trong đó có thể kể đến (i) trích xuất bộ nhớ tạm từ máy tính, (ii) tạo ảnh ổ cứng có trong máy tính và (iii) cải tiến phần mềm để có thể tương thích với các bản phân phối Linux đời cũ. Trong tương lai tôi sẽ cố gắng để hoàn thiện phần mềm và phát triển những tính năng mới, hướng tới việc thay thế hoàn toàn các công cụ, phần mềm thu thập chứng cứ hiện tại, rút ngắn quy trình phản ứng khi gặp sự cố an toàn thông tin.

## **TÀI LIỆU THAM KHẢO**

- [1] B. Nikkel, Practical Linux Forensics, No Starch Press, 2021.
- [2] P. Polstra, Linux Forensics, Pentester Academy, 2015.
- [3] M. Spreitzenbarth and J. Uhrmann, Mastering Python Forensics, Packt Publishing Ltd, 2015.