HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

NGUYỄN DIỆP ANH

NGHIÊN CỨU PHƯƠNG PHÁP PHÁT HIỆN TIẾN TRÌNH BẤT THƯỜNG TRÊN MÁY NGƯỜI DÙNG

Chuyên ngành: Hệ thống thông tin
Mã số: 8.48.01.04

TÓM TẮT LUẬN VĂN THẠC SỸ

(Theo định hướng ứng dụng)

Hà Nội - 2021

Luận văn được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: TS ĐỖ XUÂN CHỢ

Phản biện 1: TS. Hoàng Xuân Dậu

Phản biện 2: PGS TS. Nguyễn Hà Nam

Luận văn này được bảo vệ trước Hội đồng chấm luận văn thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: 11 giờ 00 phút ngày 28 tháng 8 năm 2021

Có thể tìm hiểu luận văn này tại:

Thư viện của Học viện Công nghệ Bưu chính Viễn thông

I. MỞ ĐẦU

1. Tính cấp thiết của đề tài

Các nguy cơ mất an toàn thông tin (ATTT) ngày càng gia tăng không chỉ về số lượng, quy mô tấn công mà còn về mức độ nguy hiểm của chúng gây ra cho các cơ quan tổ chức và doanh nghiệp. Tại Việt Nam theo thống kê từ Cục An Toàn Thông Tin thì năm 2019 trong tổng số 148 cổng, trang thông tin điện tử của cơ quan đơn vị có tới 2647 lỗ hổng, điểm yếu bao gồm: 19 lỗ hổng mức độ nghiêm trọng, 52 lỗ hổng ở mức cao, 270 lỗ hổng ở mức trung bình, 924 lỗ hổng ở mức thấp, 1382 lỗ hổng để lộ thông tin. Điều này thể hiện các biện pháp, kỹ thuật nhằm đảm bảo ATTT cho hệ thống thông tin của tổ chức, cơ quan đã mang lại hiệu quả nhưng chưa thực sự tốt.

Bên cạnh đó, cũng theo thống kê từ cục An Toàn Thông Tin trong năm 2019 thì có tới 95% sự cố, nguy cơ ATTT do lỗi của người dùng trong đó tấn công lừa đảo chiếm 96%, tấn công sử dụng mã độc chiếm 92,4%. Đặc biệt, trong báo cáo tổng kết năm 2019 theo ý kiến của đại diện Bộ Thông Tin và Truyền Thông, tấn công vào người dùng đã, vẫn, và sẽ là phương thức tấn công hiệu quả nhất.

Hiện nay, các hệ thống đảm bảo ATTT truyền thống như: phần mềm giám sát và phát hiện mã độc (Anti AV), hệ thống tường lửa, hệ thống phát hiện và ngăn chặn xâm nhập (IDS/IPS)...mặc dù đã phát sớm phát hiện và ngăn chặn được các cuộc tấn công, nhưng vẫn còn nhiều hạn chế. Đối với các phần mềm Anti AV mặc dù đã được tích hợp các công nghệ phát hiện mã độc tiên tiến dựa trên phân tích dấu hiệu và hành vi kết hợp với công cụ ảo hóa.

Tuy nhiên, trước sự thay đổi liên tục về cách thức phát tán và tấn công mã độc thì các hệ thống Anti AV đã không thể mang lại hiệu quả cao. Đặc biệt đối với hệ thống IDS/IPS đang gặp phải những khó khăn khi phải giám sát và phân tích lượng dữ liệu lớn để phát hiện dấu hiệu bất thường trong mạng. Điều này dẫn đến các hệ thống IDS/IPS thường chỉ phát hiện và ngăn chặn được tấn công khi kẻ tấn công đã ở giai đoạn đánh cắp dữ liệu.

Bên cạnh đó, đặc điểm chung của các hệ thống đảm bảo ATTT truyền thống là chưa cung cấp nhiều dịch vụ ứng phó và hỗ trợ xử lý các sự cố mất ATTT trong

các hệ thống thông tin. Từ những lý do trên, học viện lựa chọn đề tài "Nghiên cứu phương pháp phát hiện tiến trình bất thường trên máy người dùng".

2. Tổng quan về vấn đề nghiên cứu

Hiện nay các nghiên cứu lý thuyết liên quan đến phát hiện tiến trình bất thường trên các máy người dùng còn rất hạn chế. Tuy nhiên, liên quan đến vấn đề phát hiện mã độc trên máy Workstation ngoài các sản phẩm là các phần mềm Anti Virus thì xuất hiện một số sản phẩm hỗ trợ Phát hiện và Phản hồi Điểm cuối (Endpoint detection & Response- EDR). Sản phẩm EDR có chức năng phát hiện và theo dõi các sự cố bất thường trên Enduser để từ đó đưa ra các kịch bản ứng phó sự cố. Theo đó, có một số giải pháp và sản phẩm của EDR như sau: Sản phẩm EDR Apex One của Trend Micro có khả năng tự động phát hiện và ngăn chặn nhiều mối đe dọa điểm cuối nhất có thể, mà không cần bất cứ sự can thiệp thủ công nào từ người dùng. Apex One cũng sẽ tiến hành phát hiện và ngăn chặn việc khai thác lỗ hổng trong hệ điều hành trước khi các mối đe dọa xâm nhập vào điểm cuối với các bản vá ảo luôn được cập nhật liên tục bằng trí thông minh nhân tạo từ Trend Micro's Zero Day Initiative.

Tương tự như Apex One của Trend Micro, sản phẩm Palo Alto Networks Traps của Palo Alto ngăn chặn các mối đe dọa trên ednpoint, phối hợp với bảo mật cloud và network để ngăn chặn các cuộc tấn công mạng. Traps ngăn chặn việc khởi chạy các tệp thực thi độc hại, tệp DLLs và tệp Office bằng nhiều phương pháp ngăn ngừa, giảm bề mặt tấn công và tăng tính chính xác của việc ngăn chặn phần mềm độc hại. Cách tiếp cận này ngăn chặn phần mềm độc hại đã biết và chưa biết từ việc lây nhiễm endpoint bằng cách kết hợp: WildFire threat intelligence, Local analysis via machine learning, WildFire inspection and analysis, Granular child process protection, Periodic scanning for dormant malware.

Kaspersky EDR có thể theo dõi liên tục và phân tích các hiện tượng bất thường, các quá trình khả nghi trên các máy trạm của nhân viên và phản ứng với các mối đe doạ ở chế độ thủ công, cũng như chế độ thụ động. Ngoài ra, Kaspersky EDR cho phép kiểm soát các sự cố tại các điểm cuối của mạng, phát hiện các mã độc và

các hành vi trái phép không thể nhận biết ở mức bảo vệ mạng, đồng thời phản ứng nhanh với chúng.

Ngoài ra có một số giải pháp khác như: VMware Carbon Black EDR, Falcon, Malwarebytes Endpoint Detection and Response.

3. Mục đích nghiên cứu

- Nghiên cứu phương pháp phát hiện tiến trình bất thường trên máy trạm.
- Thực nghiệm và đánh giá phương pháp phát hiện tiến trình trên máy trạm.

4. Đối tượng và phạm vi nghiên cứu

- Các giải pháp công nghệ EDR.
- Một số luật phát hiện tiến trình bất thường trên máy trạm.

5. Phương pháp nghiên cứu

Luận văn tập trung vào nghiên cứu và tìm hiểu về lý thuyết của các giải pháp công nghệ EDR. Bên cạnh đó, luận văn sẽ tìm hiểu cách thức để áp dụng tập luật để phát hiện tiến trình bất thường trên máy người dùng sử dụng hệ điều hành Window.

II. NỘI DUNG

Dự kiến luận văn sẽ được cấu trúc với các chương như sau:

Chương 1: Giới thiệu về hệ điều hành trong máy người dùng

1.1. Tổng quan về hệ điều hành

1.2. Phân loại hệ điều hành

1.3. Kết luận chương 1

Chương 2: Nghiên cứu phương pháp thu thập tiến trình trên hệ điều hành

2.1. Một số phương pháp và kỹ thuật thu thập tiến trình Windows

2.2. Một số phương pháp và kỹ thuật thu thập tiến trình Linux

2.3. Sử dụng giải pháp ELK để thu thập tiến trình bất thường

2.4. Kết luận chương 2

Chương 3: Cài đặt và thử nghiệm phát hiện tiến trình bất thường trên máy người dùng sử dụng hệ điều hành Window

3.1. Cài đặt cấu hình Elastic Search

3.2. Cài đặt cấu hình Kibana

3.3. Cài đặt và cấu hình Logstash

3.4. Cài đặt công cụ thu thập và vận chuyển tiến trình trên hệ điều hành Windows

3.5. Cài đặt công cụ giám sát tiến trình

3.6. Thực nghiệm và đánh giá

3.7. Kết luận chương 3

III. KẾT LUẬN LUẬN VĂN

CHƯƠNG 1: GIỚI THIỆU VỀ HỆ ĐIỀU HÀNH TRONG MÁY NGƯỜI DÙNG

1.1. Tổng quan về hệ điều hành

1.1.1. Khái niệm về hệ điều hành

Hệ điều hành là một hệ thống các chương trình hoạt động giữa người sử dụng (user) và phần cứng máy tính. Mục tiêu của hệ điều hành là cung cấp một môi trường để người sử dụng có thể thi hành các chương trình. Nó làm cho máy tính dễ sử dụng hơn, thuận lợi hơn và hiệu quả hơn. Hệ điều hành là một phần quan trọng của hầu hết các hệ thống máy tính. Một hệ thống máy tính thường được chia làm bốn phần chính: phần cứng, hệ điều hành, các chương trình ứng dụng và người sử dụng.

1.1.2. Các chức năng của hệ điều hành

1.2.1.1. Quản lý tiến trình

Một tiến trình là một chương trình đang được thực thi. Một tiến trình phải sử dụng tài nguyên như thời gian sử dụng CPU, bộ nhớ, tập tin, các thiết bị nhập xuất để hoàn tất công việc của nó. Các tài nguyên này được cung cấp khi tiến trình được tạo hay trong quá trình thực thi.

Vai trò của hệ điều hành trong việc quản lý tiến trình là:

- Tạo và hủy các tiến trình của người sử dụng và của hệ thống.
- Tạm dừng và thực hiện tiếp một tiến trình.
- Cung cấp các cơ chế đồng bộ tiến trình.
- Cung cấp các cơ chế giao tiếp giữa các tiến trình.
- Cung cấp cơ chế kiểm soát deadlock.

1.2.1.2. Quản lý bộ nhớ chính

Hệ điều hành có những vai trò như sau trong việc quản lý bộ nhớ chính:

- Lưu trữ thông tin về các vị trí trong bộ nhớ đã được sử dụng và tiến trình nào đang sử dụng.
- Quyết định tiến trình nào được nạp vào bộ nhớ chính, khi bộ nhớ đã có thể dùng được.
- Cấp phát và thu hồi bộ nhớ khi cần thiết.

1.2.1.3. Quản lý bộ nhớ phụ

Vai trò của hệ điều hành trong việc quản lý bộ nhớ phụ:

- Quản lý vùng trống trên đĩa.
- Định vị lưu trữ.
- Lập lịch cho đĩa

1.2.1.4. Quản lý hệ thống vào/ra

Một hệ thống vào/ra bao gồm:

- Thành phần quản lý bộ nhớ vùng đệm, lưu trữ tạm thời và vùng chứa.
- Giao tiếp điều khiển thiết bị (device driver) tổng quát.
- Bộ điều khiển cho các thiết bị xác định.

1.2.1.5. Quản lý hệ thống tập tin

Vai trò của hệ điều hành trong quản lý tập tin:

- Tạo và xóa một tập tin.
- Tạo và xóa một thư mục.
- Hỗ trợ các thao tác trên tập tin và thư mục.
- Ánh xạ tập tin trên hệ thống lưu trữ phụ.
- Sao lưu dự phòng các tập tin trên các thiết bị lưu trữ.

1.1.3. Các tính chất cơ bản của hệ điều hành

Các hệ điều hành hiện đại ngày nay đều có các tính chất cơ bản như sau:

- Tin cậy
- An toàn
- Hiệu quả
- Tổng quát theo thời gian
- Thuận tiện

1.2. Một số hệ điều hành trên máy người dùng

1.2.1. Giới thiệu về hệ điều hành Windows

Microsoft Windows là tên của một họ hệ điều hành dựa trên giao diện người dùng đồ hoạ được phát triển và được phân phối bởi Microsoft. Nó bao gồm một vài các dòng hệ điều hành, mỗi trong số đó phục vụ một phần nhất định của ngành công nghiệp máy tính.

1.2.1.1. Quản lý tiến trình trong windows

Một bộ xử lý trong máy tính chạy Windows có hai chế độ khác nhau: chế độ người dùng và chế độ kernel. Bộ xử lý chuyển đổi giữa hai chế độ tùy thuộc vào loại mã nào đang chạy trên bộ xử lý. Các ứng dụng chạy ở chế độ người dùng và các thành phần hệ điều hành lõi chạy ở chế độ kernel. Trong khi nhiều trình điều khiển chạy ở chế độ kernel, một số trình điều khiển có thể chạy ở chế độ người dùng.

User mode: Khi một chương trình khởi tạo trên User Mode, Windows sẽ tạo một tiến trình (process) cho chương trình đó. Một tiến trình cung cấp cho chương trình đó một không gian địa chỉ ảo (Virtual address space). Bởi vì mỗi không gian địa chỉ ảo của mỗi tiến trình là riêng biệt, nên một chương trình khó có thể thể thay đổi dữ liệu thuộc về một chương trình khác. Do mỗi chương trình chạy một cách cô lập như vậy, nên nếu một chương trình gặp sự cố thì sự cố này sẽ được giới hạn trong chương trình đó. Những chương trình khác và hệ điều hành sẽ không bị ảnh hưởng. Không gian địa chỉ ảo của mỗi tiến trình ở User Mode bị giới hạn. Một chương trình chạy tại User Mode sẽ không xử lý được những địa chỉ ảo (Virtual address) được dự trữ để sử dụng bởi hệ điều hành. Sự giới hạn không gian địa chỉ nhằm bảo vệ những chương trình khỏi việc xử lý nhầm vào những vùng dữ liệu quan trọng, có thể gây hại tới hệ điều hành.

Kernel mode: Tất cả các mã chạy trong chế độ kernel chia sẻ một không gian địa chỉ ảo. Điều này có nghĩa là trình điều khiển chế độ kernel không bị cô lập với các trình điều khiển khác và chính hệ điều hành. Nếu trình điều khiển chế độ kernel vô tình ghi sai địa chỉ ảo, dữ liệu thuộc về hệ điều hành hoặc trình điều khiển khác có thể bị xâm phạm. Nếu trình điều khiển chế độ kernel gặp sự cố, toàn bộ hệ điều hành sẽ gặp sự cố.

- **Stack**: chứa dữ liệu tạm thời, thường chỉ chứa dữ liệu dùng trong thời gian rất ngắn hoặc dùng một lần.

- **Heap**: đây là vùng bộ nhớ được cấp phát động cho một tiến trình trong thời gian thực thi.

- **Data**: chứa các biến global (biến dùng chung ở cấp toàn bộ chương trình) và static (loại biến dùng chung ở cấp struct).

- **Text**: bao gồm hoạt động hiện tại được biểu thị bằng giá trị của Program Counter và nội dung của các thanh ghi trên vi xử lý.

Khi một tiến trình thực thi, nó có thể đi qua các trạng thái khác nhau. Các giai đoạn này có thể khác nhau đối với các hệ điều hành khác nhau và tên của các trạng thái này cũng không được chuẩn hóa cụ thể. Nhưng nhìn chung, một tiến trình có thể có một trong năm trạng thái sau đây tại một thời điểm:

- **Start**: trạng thái khi một tiến trình được tạo hoặc được khởi động lần đầu.

- **Ready**: tiến trình đang chờ để được chuyển cho vi xử lý. Thường thì khi một tiến trình khác đang chạy, CPU sẽ không cho phép một tiến trình khác xen vào mà sẽ phải đợi tới lượt để đảm bảo tiến trình hiện tại đã xử lý xong (lưu ý, tiến trình xử lý xong không có nghĩa là nó sẽ kết thúc). Trạng thái này thường xuất hiện ngay sau trạng thái Start.

- **Running**: trạng thái thực thi của tiến trình, khi bộ lập lịch Scheduler đã chuyển nó cho vi xử lý, vi xử lý sẽ tiến hành thực hiện các tính toán hoặc chỉ dẫn trên nó.

- **Wait**: trạng thái này xảy ra nếu tiến trình cần chờ một tài nguyên để có thể sử dụng, ví dụ như chờ người dùng nhập liệu hoặc chờ đợi tệp tin đang được sử dụng bởi tiến trình khác (vấn đề này có thể được thảo luận liên quan đến xử lý I/O).

- **Terminated**: xảy ra khi một tiến trình đã xử lý xong hoặc bị ngắt bởi yêu cầu từ hệ thống hoặc người dùng, trạng thái này có thể được gọi là Exit.

1.2.1.2. Cấu trúc tập tin thực thi trong Windows

Cấu trúc cơ bản của PE File: Sau đây là các phần phổ biến và quan trọng nhất trong một tập tin PE:

- Section.text chứa các hướng dẫn cho CPU thực thi. Tất cả các Section khác lưu trữ dữ liệu và thông tin hỗ trợ. Nói chung, đây là một Section duy nhất mà có thể thực thi và nó là section duy nhất chứa mã code của chương trình.

- Section.rdata thường chứa các thông tin về các hàm import và export, những thông tin này thường được cung cấp bởi hai công cụ Dependency Walker và

Peview. Section này cũng có thể lưu trữ những dữ liệu thuộc loại read-only khác được sử dụng bởi chương trình. Đôi khi một tập tin sẽ chứa một Section.idata và .edata, những Section này cũng lưu trữ các thông tin về các hàm import và export.

- Section.data chứa dữ liệu toàn bộ của chương trình, nó có thể truy cập từ bất cứ nơi nào trong chương trình. Dữ liệu cục bộ thì không được lưu trữ trong Section này hoặc ở bất cứ nơi nào khác trong tập tin PE.

- Section.rsrc bao gồm các tài nguyên được sử dụng bởi tiến trình thực thi, những tài nguyên này như là icon, images, menus, và các strings. Các Strings có thể được lưu trữ trong Section.rsrc hoặc trong chương trình chính, tuy nhiên chúng thường được lưu trữ trong Section.rsrc để hỗ trợ đa ngôn ngữ.

1.2.2. Giới thiệu về hệ điều hành Linux

Linux là tên gọi của một hệ điều hành máy tính và cũng là tên hạt nhân của hệ điều hành. Phiên bản Linux đầu tiên do Linus Torvalds viết vào năm 1991.

1.2.2.1. Quản lý tiến trình trong Linux

Những hệ thống máy tính ban đầu cho phép chỉ một chương trình được thực thi tại một thời điểm. Chương trình này có toàn quyền điều khiển hệ thống và có truy xuất tới tất cả tài nguyên của hệ thống. Những hệ thống máy tính hiện nay cho phép nhiều chương trình được nạp vào bộ nhớ và được thực thi đồng hành.

Mỗi tiến trình trong Linux được biểu diễn bằng một cấu trúc dữ liệu task_struct. Linux sử dụng task_vector để quản lý các con trỏ đến các task_struct, mặc định là có 512 phần tử.

Khi một tiến trình được tạo ra, một task_struct mới được cấp phát trong bộ nhớ và được thêm vào vector_task. Linux hỗ trợ hai loại tiến trình là loại bình thường và loại thời gian thực.

1.2.2.2. Cấu trúc tập tin thực thi trên Linux

Cấu trúc tập tin thực thi trên linux có tên gọi là Executable and Linkable Format (ELF format) là định dạng tập tin tiêu chuẩn phổ biến cho các tập tin thực thi, mã đối tượng, thư viện dùng chung và kết xuất lõi.

ELF được thiết kế có tính linh hoạt, khả năng mở rộng, hỗ trợ đa nền tảng cho các định dạng và kích thước địa chỉ khác nhau. Thiết kế ELF không bị giới hạn trong một bộ xử lý, tập lệnh hoặc kiến trúc phần cứng cụ thể.

Do đó, ELF được sử dụng bởi nhiều hệ điều hành khác như: Solaris, Ubuntu, OpenBSD, QNX, BeOS, Haiku và Fuchsia.

1.3. Kết luận chương 1

Kết thúc chương 1, luận văn đã trình bày tổng quan về hệ điều hành, trong đó trình bày chi tiết về hai hệ điều hành phổ biến nhất hiện nay là Windows và Linux. Mỗi hệ điều hành, luận vă đã nghiên cứu về cấu trúc tập tin thực thi và cách thức quản lý tiến trình. Trong chương tiếp theo, luận văn sẽ tập trung nghiên cứu về các phương thức thu thập dữ liệu liên quan tới tiến trình trên hệ thống.

CHƯƠNG 2: NGHIÊN CỨU PHƯƠNG PHÁP THU THẬP TIẾN TRÌNH TRÊN HỆ ĐIỀU HÀNH

2.1. Một số phương pháp và kỹ thuật thu thập tiến trình Windows

2.1.1. Process Monitor

Process Monitor là công cụ giám sát và theo dõi dành cho hệ điều hành Windows được phát triển từ hai công cụ là Filemon và Regmon, được sử dụng để giám sát các tệp và hoạt động của chúng. Process Monitor là công cụ giám sát hệ thống vô cùng mạnh mẽ, với khả năng giám sát các tập tin, cửa sổ đăng ký của một ứng dụng. Process Monitor có khả năng thu thập rất nhiều dữ liệu trên máy tính Windows, đặc biệt tập trung vào các hoạt động (Input/Output) sau:

- Registry: liên quan tới các việc tạo, xóa, truy vấn registry.
- File System: liên quan tới hoạt động tạo, ghi, xóa tập tin.
- Network: hiển thị lưu lượng TCP/UDP.
- Process: thông tin liên quan tới tiến trình như khởi tạo, xóa, suspend…
- Cấu hình: những thông tin liên quan tới bộ nhớ, thời gian…

2.1.2. System Monitor

Sysmon là trình điều khiển thiết bị và dịch vụ hệ thống Windows, sau khi được cài đặt sẽ tồn tại trong toàn hệ thống để theo dõi và ghi nhật ký hoạt động của hệ thống vào nhật ký sự kiện của Windows. Sysmon cung cấp thông tin chi tiết về các process được tạo, kết nối mạng và thay đổi thời gian tạo tệp. Bằng cách thu thập các sự kiện mà nó tạo ra bằng Windows Agent Collection hoặc các tác nhân SIEM và sau đó phân tích chúng để xác định hoạt động độc hại hoặc bất thường, từ đó tìm hiểu các kẻ xâm nhập hoạt động trên mạng.

Tổng quan về khả năng của Sysmon:

- Ghi nhật ký quá trình tạo tiến trình với đầy đủ dòng lệnh cho cả tiến trình hiện tại và tiến trình cha.

- Ghi lại giá trị băm của tập tin khởi tạo tiến trình với giá trị mặc định là SHA1, ngoài ra còn có các giá trị khác như MD5, SHA256 hoặc IMPHASH.

- Lưu trữ nhiều giá trị băm.

- Lưu trữ giá trị GUID trong quá trình tạo sự kiện để cho phép tổng hợp, phân tích sự tương quan giữa các sự kiện ngay cả khi hệ điều hành Windows sử dụng lại PID.

- Giá trị Session GUID trong mỗi sự kiện để cho phép tổng hợp, phân tích tương quan các sự kiện trên cùng một phiên đăng nhập.

- Nhật ký quá trình tải, quá trình điều khiển hệ thống, thư viện với chữ ký và giá trị băm của từng tập tin.

- Lưu nhật ký mở, đọc ổ đĩa.

- Lưu nhật ký kết nối mạng, bao gồm thông tin tiến trình kết nối, địa chỉ IP, cổng, địa chỉ kết nối.

- Phát hiện các thay đổi về thời gian tạo tập tin để thu thông tin liên quan tới sự kiện thay đổi tập tin. Thay đổi thời gian tạo tập tin là một trong những kỹ thuật hay được mã độc sử dụng.

- Tự động tải lại cấu hình nếu có thay đổi trong registry.

- Cơ chế lọc giúp loại trừ các sự kiện một cách linh hoạt

- Theo dõi, giám sát hệ thống ngay từ khi quá trình khởi động để nắm bắt được toàn bộ hoạt động của các tiến trình trên hệ thống Windows.

Hiện tại, Microsoft đã phát hành Sysmon phiên bản 11, người dùng có thể tải về từ trang chủ của hệ thống microsoft hoặc thông qua địa chỉ: https://docs.microsoft.com/vi-vn/sysinternals/. Trang web Sysinternals được tạo ra từ năm 1996 bởi Mark Russinovich để lưu trữ các tiện ích hệ thống và thông tin kỹ thuật tiến tiến của ông. Ngay cả những chuyên gia an ninh mạng hàng đầu thế giới, hay nhà phát triển phần mềm Windows, đều truy cập địa chi web này để tìm các tiện ích, công cụ liên quan tới việc giám sát Windows.

Các tiến trình thu thập được bằng Sysmon trên Event Viewer:

- **Event ID 1:** Process creation cung cấp thông tin mở rộng về một process mới được tạo.

- **Event ID 2:** A process changed a file creation time được đăng ký khi thời gian tạo tệp được sửa đổi bởi một process sự kiện này giúp theo dõi thời gian tạo thực sự của một tập tin. Kẻ tấn công có thể thay đổi thời gian tạo tập tin để làm cho

nó trông giống như nó đã được cài đặt với hệ điều hành. Lưu ý rằng nhiều quy trình thay đổi hợp pháp thời gian tạo tệp, nó không nhất thiết chỉ ra hoạt động độc hại.

- Event ID 3: Network connection là sự kiện kết nối mạng ghi lại các kết nối TCP / UDP trên máy. Nó bị tắt theo mặc định. Mỗi kết nối được liên kết với một quy trình thông qua các trường ProcessID và ProcessGUID. Sự kiện này cũng chứa tên máy chủ nguồn và đích địa chỉ IP, số port và trạng thái IPv6.

- Event ID 4: Sysmon service state changed là sự kiện thay đổi trạng thái dịch vụ Sysmon (đã bắt đầu hoặc đã dừng).

- Event ID 5: Process terminated là quá trình báo cáo sự kiện khi một quá trình kết thúc. Nó cung cấp UtcTime, ProcessGuid và ProcessID của quy trình.

- Event ID 6: Driver loaded cung cấp thông tin về driver đang được tải trên hệ thống.

- Event ID 7: Image loaded là nhật ký sự kiện tải hình ảnh khi một module được tải trong một process cụ thể. Sự kiện này nên được cấu hình cẩn thận vì giám sát tất cả các sự kiện tải hình ảnh sẽ tạo ra một số lượng lớn các sự kiện.

- Event ID 8: CreateRemoteThread sự kiện CreatRemoteThread phát hiện khi một tiến trình tạo một luồng trong tiến trình khác. Kỹ thuật này được sử dụng bởi phần mềm độc hại để tiêm mã và ẩn trong các quy trình khác. Sự kiện chỉ ra quá trình nguồn và đích. Nó cung cấp thông tin về mã sẽ được chạy trong luồng mới: StartAddress, StartModule và StartFunction. Lưu ý rằng các trường StartModule và StartFunction được suy ra, chúng có thể trống nếu địa chỉ bắt đầu nằm ngoài các module được tải hoặc các hàm xuất đã biết.

- Event ID 9: RawAccessRead sự kiện RawAccessRead phát hiện khi một quá trình tiến hành các hoạt động đọc từ ổ đĩa bằng cách sử dụng ký hiệu "\\.**",** Kỹ thuật này thường được sử dụng bởi phần mềm độc hại để lọc dữ liệu của các tệp bị khóa để đọc, cũng như để tránh các công cụ kiểm tra truy cập tệp. Sự kiện chỉ ra quá trình nguồn và thiết bị đích.

- Event ID 10: ProcessAccess là quá trình truy cập báo cáo sự kiện khi một quy trình mở ra một quy trình khác, một hoạt động thường được theo sau bởi các truy vấn thông tin hoặc đọc và ghi không gian địa chỉ của quy trình đích. Điều này

cho phép phát hiện các công cụ hack đọc nội dung bộ nhớ của các quy trình như Local Security Agency (Lsass.exe) đánh cắp thông tin đăng nhập để sử dụng trong các cuộc tấn công Pass-the-Hash. Việc kích hoạt nó có thể tạo ra số lượng ghi nhật ký đáng kể nếu có các tiện ích chẩn đoán hoạt động liên tục mở các quy trình để truy vấn trạng thái của chúng, do đó thường chỉ nên thực hiện với các bộ lọc loại bỏ các truy cập dự kiến.

- Event ID 11: FileCreate là hoạt động tạo tập tin được ghi lại khi một tập tin được tạo hoặc ghi đè. Sự kiện này theo dõi các vị trí tự khởi động như thư mục Startup, cũng như các thư mục tải xuống tạm thời và là những nơi phổ biến khiến phần mềm độc hại bị loại bỏ trong quá trình lây nhiễm ban đầu.

- Event ID 12: RegistryEvent (Object create and delete) khóa đăng ký và giá trị khởi tạo và xóa operations map theo loại sự kiện này, có thể hữu ích cho việc theo dõi các thay đổi đối với các vị trí tự khởi động của Sổ đăng ký hoặc phần mềm thay đổi độc hại.

- Event ID 13: RegistryEvent (Value Set) xác định sửa đổi giá trị Registry. Sự kiện này ghi lại giá trị được ghi cho các giá trị Registry của loại DWORD và QWORD.

- Event ID 14: RegistryEvent (Key and Value Rename) Registry key và đổi tên operations map thành loại sự kiện này, ghi lại tên mới của khóa hoặc giá trị đã được đổi tên.

- Event ID 15: FileCreateStreamHash sự kiện này ghi lại khi tạo một dòng tệp và nó tạo ra các sự kiện ghi lại hàm băm nội dung của tệp mà luồng được gán cũng như nội dung của luồng được đặt tên.

- Event ID 17: PipeEvent (Pipe Created) sự kiện này tạo ra khi đường ống được tạo ra. Phần mềm độc hại thường sử dụng các đường ống được đặt tên để liên lạc giữa các quá trình.

- Event ID 18: PipeEvent (Pipe Connected) sự kiện này ghi lại khi kết nối đường ống được thực hiện giữa máy khách và máy chủ.

- Event ID 19: WmiEvent (WmiEventFilter activity detected) khi bộ lọc sự kiện WMI được đăng ký, đây là phương thức được phần mềm độc hại sử dụng để thực thi, sự kiện này sẽ ghi lại không gian tên WMI, tên bộ lọc và biểu thức lọc.

- Event ID 20: WmiEvent (WmiEventConsumer activity detected) sự kiện này ghi lại đăng ký của người tiêu dùng WMI, ghi lại tên người dùng, nhật ký và đích đến.

- Event ID 21: WmiEvent (WmiEventConsumerToFilter activity detected) khi người tiêu dùng liên kết với bộ lọc, sự kiện này sẽ ghi lại tên người tiêu dùng và đường dẫn bộ lọc.

- Event ID 22: DNSEvent (DNS query) sự kiện này tạo ra khi một quá trình thực hiện truy vấn DNS, cho dù kết quả thành công hay thất bại, được lưu trữ hay không. Sự kiện này đã được thêm từ Windows 8.1 vì vậy nó không có sẵn trên Windows 7 trở về trước.

- Event ID 255: Error sự kiện này được tạo khi xảy ra lỗi trong Sysmon. Chúng có thể xảy ra nếu hệ thống chịu tải nặng và không thể thực hiện một số nhiệm vụ nhất định hoặc có lỗi trong dịch vụ Sysmon.

2.2. Một số phương pháp và kỹ thuật thu thập tiến trình Linux

Cấu trúc cơ sở của Linux khác biệt hoàn toàn so với Windows. Trong khi Windows duy trì một hệ thống cấu trúc khá chuẩn, với các bản cập nhật và các phiên bản phân tầng, thì Linux lại phức tạp hơn rất nhiều. Được viết ban đầu bởi sinh viên Phần Lan là Linus Torvalds, hạt nhân Linux ngày nay trở thành cơ sở cho tất cả các hệ điều hành Linux.

2.2.1. Process Status

Process Status (PS) là một tiện ích của Unix/Linux dùng để theo dõi, giám sát thông tin của các tiến trình đang chạy trong hệ thống. PS được sử dụng để liệt kê các tiến trình hiện đang chạy bao gồm cả PID (Process ID) của tiến trình giúp hệ điều hành phân biệt các tiến trình chạy trên hệ thống cùng với một số thông tin khác phụ thuộc vào các tùy chọn khác nhau.

PS đọc thông tin liên quan tới tiến trình từ tập tin ảo tại đường dẫn /proc file-system.

2.2.2. Top

Top là công cụ Unix khác được sử dụng rộng rãi giống như PS và có khả năng cập nhật thông tin tiến trình theo thời gian thực. Top là một trong những công cụ được sử dụng nhiều nhất để theo dõi hiệu suất thời gian thực của các tiến trình trong hệ thống Unix kể từ khi được giới thiệu từ năm 1984 trong BSD Unix4.1.

2.2.3. XosView

Xosview OS Monitor (XosView) là công cụ giám sát hiệu suất hệ thống thời gian thực cung cấp giao diện đồ họa cho các hệ thống Unix. Đây là công cụ đơn giản so với các công cụ được mô tả từ trước cho đến nay vì nó không giám sát các tiến trình đơn lẻ, thay vào đó chỉ giám sát toàn bộ hệ thống.

2.2.4. TreePs

TreePs là công cụ giám sát tiến trình thời gian thực cung cấp đồ họa cho các hệ thống Unix chạy X Window System. TreePs có khả năng tự động cập nhật thông tin theo thời gian thực về các hoạt động của tiến trình, lấy mẫu dữ liệu cho các tiến trình hoạt đồng nhiều và ít hoạt động. TreePs hiển thị tất cả các trường từ cấu trúc của tiến trình bao gồm: thời gian CPU chạy, bộ nhớ hiện tại được sử dụng và một số thông tin cơ bản về tiến trình như ProcessID, tiến trình cha.

2.3. Kết luận chương 2

Kết thúc chương 2, luận văn đã trình bày được về các phương pháp, công cụ thu thập dữ liệu nhật ký, thông tin liên quan tới tiến trình trên hệ thống Windows và Linux. Trong chương tiếp theo luận văn tập trung nghiên cứu, xây dựng ứng dụng phát hiện tiến trình độc trên máy người dùng.

CHƯƠNG 3. CÀI ĐẶT VÀ THỬ NGHIỆM PHÁT HIỆN TIẾN TRÌNH BẤT THƯỜNG TRÊN MÁY NGƯỜI DÙNG SỬ DỤNG HỆ ĐIỀU HÀNH WINDOW

3.1. Cài đặt công cụ thu thập và vận chuyển tiến trình trên hệ điều hành Window

Để cài đặt Sysmon, luân văn thực hiện tải tập tin cài đặt tại địa chỉ: *https://docs.microsoft.com/en-us/sysinternals/downloads/Sysmon* và sử dụng tập tin cài đặt Sysmon được các chuyên gia an ninh mạng của trung tâm ứng cứu sự cố Nhật Bản (JPCert) tại địa chỉ: https://github.com/SwiftOnSecurity/Sysmon-config.

```
<RuleGroup name="" groupRelation="or">
        <ProcessCreate onmatch="exclude">
                <!--SECTION: Microsoft Windows-->
                <CommandLine condition="begin with"> "C:\Windows\system32\wermgr.exe" "-queuereporting_svc" </CommandLine> <!--Wind
                <CommandLine condition="begin with">C:\Windows\system32\DllHost.exe /Processid</CommandLine> <!--Windows-->
                <CommandLine condition="begin with">C:\Windows\system32\wbem\wmiprvse.exe -Embedding</CommandLine> <!--Windows: WMI
                <CommandLine condition="begin with">C:\Windows\system32\wbem\wmiprvse.exe -secured -Embedding</CommandLine> <!--Win
                <CommandLine condition="is">C:\Windows\system32\wermgr.exe -upload</CommandLine> <!--Windows:Windows error reportin
                <CommandLine condition="is">C:\Windows\system32\SearchIndexer.exe /Embedding</CommandLine> <!--Windows: Search Inde
                <CommandLine condition="is">C:\windows\system32\wermgr.exe -queuereporting</CommandLine> <!--Windows:Windows error
                <CommandLine condition="is">\??\C:\Windows\system32\autochk.exe *</CommandLine> <!--Microsoft:Bootup: Auto Check Ut
                <CommandLine condition="is">\SystemRoot\System32\smss.exe</CommandLine> <!--Microsoft:Bootup: Windows Session Manag
                <CommandLine condition="is">C:\Windows\System32\RuntimeBroker.exe -Embedding</CommandLine> <!--Windows:Apps permiss
                <Image condition="is">C:\Program Files (x86)\Common Files\microsoft shared\ink\TabTip32.exe</Image> <!--Windows: To
                <Image condition="is">C:\Windows\System32\TokenBrokerCookies.exe</Image> <!--Windows: SSO sign-in assistant for Mic
                <Image condition="is">C:\Windows\System32\plasrv.exe</Image> <!--Windows: Performance Logs and Alerts DCOM Server--
                <Image condition="is">C:\Windows\System32\wifitask.exe</Image> <!--Windows: Wireless Background Task-->
                <Image condition="is">C:\Windows\system32\CompatTelRunner.exe</Image> <!--Windows: Customer Experience Improvement-
                <Image condition="is">C:\Windows\system32\PrintIsolationHost.exe</Image> <!--Windows: Printing-->
                <Image condition="is">C:\Windows\system32\SppExtComObj.Exe</Image> <!--Windows: KMS activation-->
                <Image condition="is">C:\Windows\system32\audiodg.exe</Image> <!--Windows: Launched constantly-->
                <Image condition="is">C:\Windows\system32\conhost.exe</Image> <!--Windows: Command line interface host process-->
                <Image condition="is">C:\Windows\system32\mobsync.exe</Image> <!--Windows: Network file syncing-->
                <Image condition="is">C:\Windows\system32\musNotification.exe</Image> <!--Windows: Update pop-ups-->
                <Image condition="is">C:\Windows\system32\musNotificationUx.exe</Image> <!--Windows: Update pop-ups-->
                <Image condition="is">C:\Windows\system32\powercfg.exe</Image> <!--Microsoft:Power configuration management-->
                <Image condition="is">C:\Windows\system32\sndVol.exe</Image> <!--Windows: Volume control-->
```

Hình 3.1. Tập tin cấu hình Sysmon

Để tiến hành cài đặt Sysmon với tập tin cấu hình cài đặt riêng biệt, luân văn thực hiện câu lệnh sau:

```
Sysmon –I –acceptcula –I Sysmonconfig-export.xml
```

Thực hiện gửi dữ liệu Sysmon thu thập được trên máy tính client về máy chủ phân tích, luân văn sử dụng công cụ "Winlogbeat" để thực hiện. Trong đó cấu hình cài đặt WinlogBeat.yml như sau:

```
winlogbeat.event_logs:
  - name: "Microsoft-Windows-Sysmon/Operational"
# It is necessary to delete the following default processors configuration:
#    processors:
#      - script:
```

```
#         lang: javascript
#         id: Sysmon
#         file: ${path.home}/module/Sysmon/config/winlogbeat-Sysmon.js
...
# Configure ElasticSearch Server IP address
output.elasticsearch:
  hosts: ["Elasticserach server IP address:9200"]
```

Hình 3.5. Cấu hình WinlogBeat

Sau khi cấu hình thành công Winlogbeat, thực hiện chạy lệnh như sau để gửi dữ liệu nhật ký event của Sysmon thu thập được trên máy tính Windows 7 về máy chủ phân tích dữ liệu.

Winlogbeat.exe -e -c winlogbeat.yml

3.2. Cài đặt công cụ giám sát tiến trình

Để xây dựng cơ sở dữ liệu về tiến trình độc, luận văn sử dụng các dấu hiệu thu thập từ Virustotal. Các thành phần gồm: mã băm của mã độc, IP độc, domain độc… Để xây dựng cơ sở dữ liệu dấu hiệu nhằm phát hiện tiến trình độc, trong báo cáo này tác giả sử dụng công cụ MongoDB.

MongoDB có thể cài đặt trên nhiều hệ điều hành khác nhau nhưng trong đề tài này Ubuntu sẽ được chọn làm hệ điều hành để xây dựng môi trường phát triển hệ thống, do tính phổ biến, miễn phí và dễ tùy chỉnh.

3.2.1. Cài đặt Mongo DB

Để cài đặt MongoDB trên Ubuntu cần thực hiện các bước sau:

- Bước 1: Thêm MongoDB public GPG key
- Bước 2: Thêm MongoDB repository vào thư mục sources.list.d
- Bước 3: Cập nhật repositories
- Bước 4: Cài đặt MongoDB
- Bước 5: Chạy MongoDB

3.2.2. Cấu hình Mongo DB

File cấu hình của MongoDB được lưu ở trong thư mục /etc/mongod.conf, và được viết dưới định dạng YAML.

3.2.3. Truy vấn MongoDB sử dụng Python

Trước khi sử dụng cần cài đặt python và python-pip. Sau đó chạy lệnh sau để cài đặt pymongo để kết nối với MongoDB: sudo pip install pymongo.

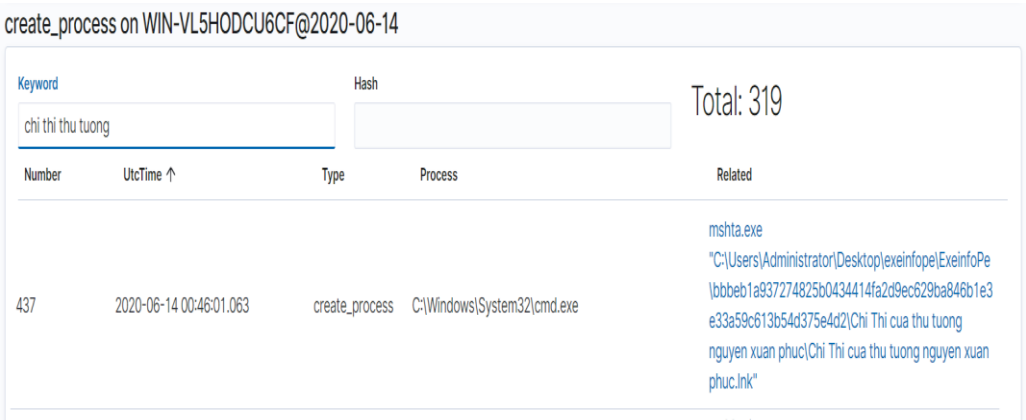
3.3. Thực nghiệm và đánh giá

Cụ thể, những ngày gần đây, tin tặc đã phát tán mã độc qua thư điện tử có đính kèm tập tin word có tiêu đề "Chi Thi cua Thu tuong nguyen xuan phuc.lnk" giả dạng thông báo của Thủ tướng Chính phủ về dịch COVID-19.

Bảng 3.2. Thông tin chi tiết về mã độc sử dụng

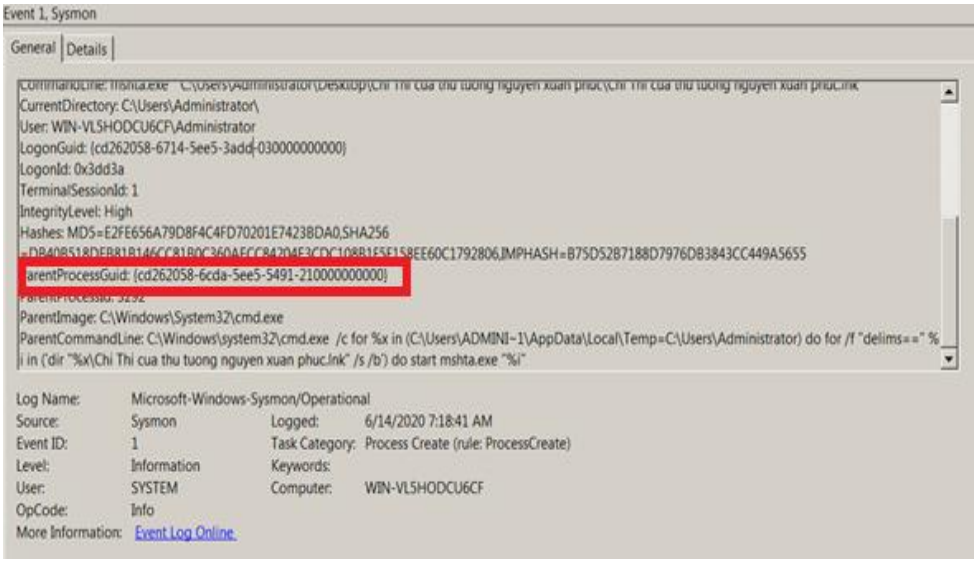
Tập tin mã độc	bbbeb1a937274825b0434414fa2d9ec629ba846b1e3e33a59c613b54d375e4d2.rar
Hash	60C89B54029442C5E131F01FF08F84C9
Định dạng tập tin	Tập tin rar sau khi giải nén ra tập tin "Chi Thi cua thu tuong nguyen xuan phuc.lnk"
Địa chỉ tải mã độc	https://app.any.run/tasks/dd877b4d-8b36-48c0-af07-ce37fd9fee7b/

Tiến hành tải tập tin mã độc và giải nén trên máy tính client, luận văn nhận được tập tin "Chi thi cua thu tuong nguyen xuan phuc.lnk" Phân tích sơ bộ tập tin mã độc, luận văn nhận định tập tin File.lnk là định dạng file lnk của Windows Fake icon và tên file để đánh lừa người dùng. Sau khi thực hiện chạy tập tin mã độc, luận văn truy cập máy chủ phân tích dữ liệu để tìm kiếm thông tin về tiến trình "Chi Thi cua thu tuong nguyen xuan phuc .lnk", luận văn thu được thông tin về tiến trình mã độc được khởi tạo như sau:



Hình 3.13. Tiến trình mã độc được khởi tạo

Kiểm tra dữ liệu event log do Sysmon lưu dữ lại có thông tin về tiến trình như sau: ParentProcessGuild: {cd262058-6cda-5ee5-5493-210000000000}, thực hiện chạy commandline: C:\Windows\system32\cmd.exe /c for %x in (C:\Users\ADMINI~1\AppData\Local\Temp=C:\Users\Administrator) do for /f "delims==" %i in ('dir "%x\Chi Thi cua thu tuong nguyen xuan phuc.lnk" /s /b') do start mshta.exe "%i"



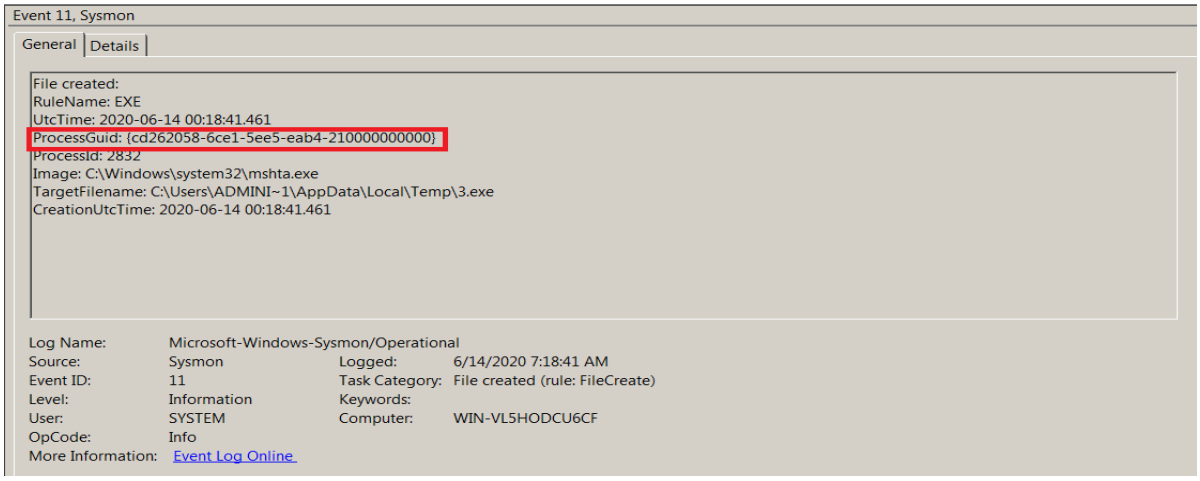
Hình 3.14. Event tiến trình mã độc chạy Sysmon thu thập

Sau event Process Create, Sysmon lưu trữ được thêm hai event liên quan tới File Created như sau:

Level	Date and Time	Source	Event ID	Task Category
Information	6/14/2020 7:18:41 AM	Sysmon	11	File created (rule: FileCreate)
Information	6/14/2020 7:18:41 AM	Sysmon	1	Process Create (rule: Proces...
Information	6/14/2020 7:18:41 AM	Sysmon	1	Process Create (rule: Proces...
Information	6/14/2020 7:18:41 AM	Sysmon	1	Process Create (rule: Proces...
Information	6/14/2020 7:18:41 AM	Sysmon	1	Process Create (rule: Proces...
Information	6/14/2020 7:18:41 AM	Sysmon	11	File created (rule: FileCreate)
Information	6/14/2020 7:18:41 AM	Sysmon	11	File created (rule: FileCreate)
Information	6/14/2020 7:18:41 AM	Sysmon	11	File created (rule: FileCreate)

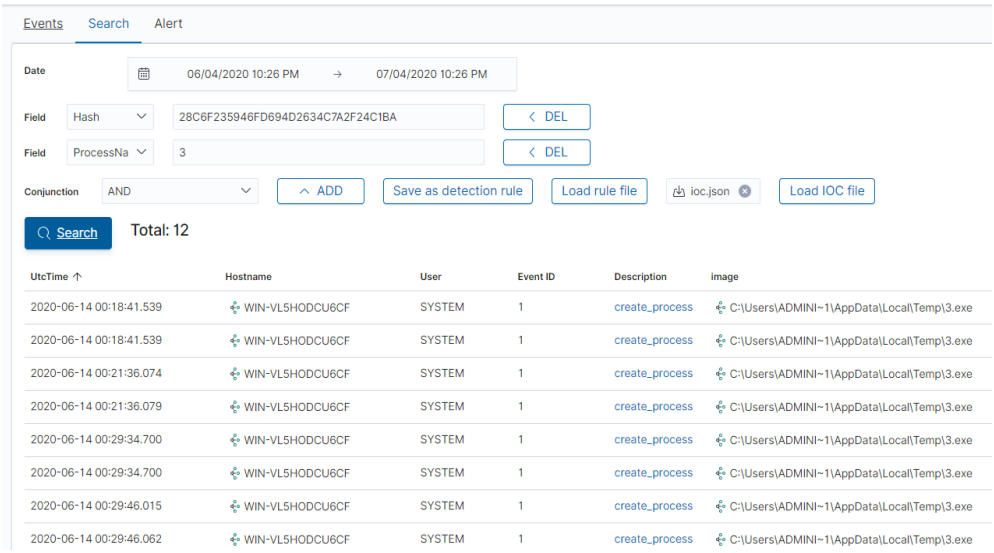
Hình 3.15. Event liên quan tiến trình mã độc

Kiểm tra thông tin về hai event này ta có được từ tập tin doc, mã độc đã thực hiện tạo ra hai tập tin có tên là 3.exe và http_dll.dll tại đường dẫn: C\users\Admin~1\AppData\Local\Temp\. Event có giá trị ProcessGuild trùng với giá trị ParentProcessGuild của tập tin mã độc.



Hình 3.16. Event liên quan tới tạo tập tin của tiến trình mã độc

Thực hiện tìm kiếm dữ liệu trên hệ thống, luận văn nhận thấy thông qua tập tin ioc về dấu hiệu nhận biết mã độc, hệ thống đã phát hiện được rất nhiều sự kiện liên quan tới tiến trình mã độc này.

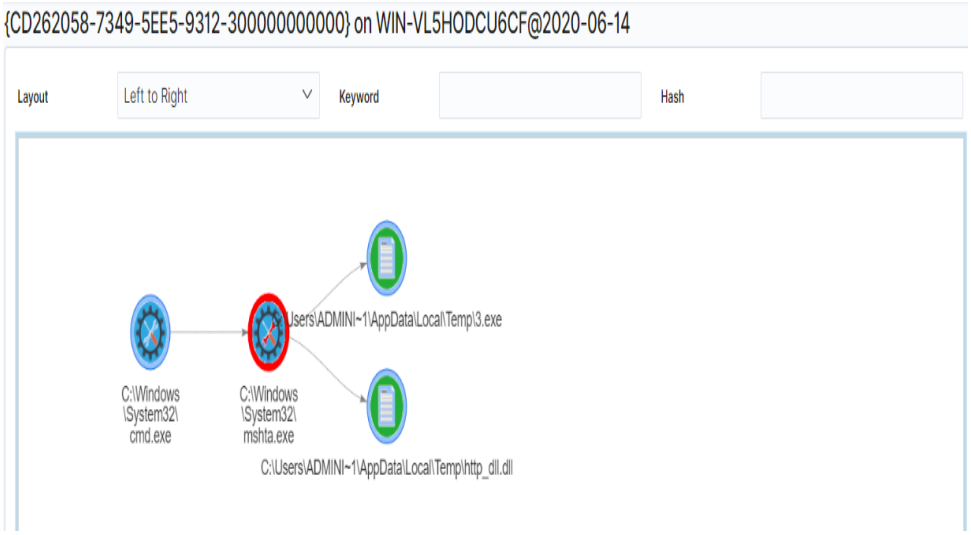


Hình 3.17. Xác định tiến trình độc thông qua tập tin IOC

Phân tích chi tiết vào sự kiện CREATE_PROCESS này cho thấy tiến trình sau khi gọi tập tin cmd.exe đã gọi them hai tiến trình khác đó là:

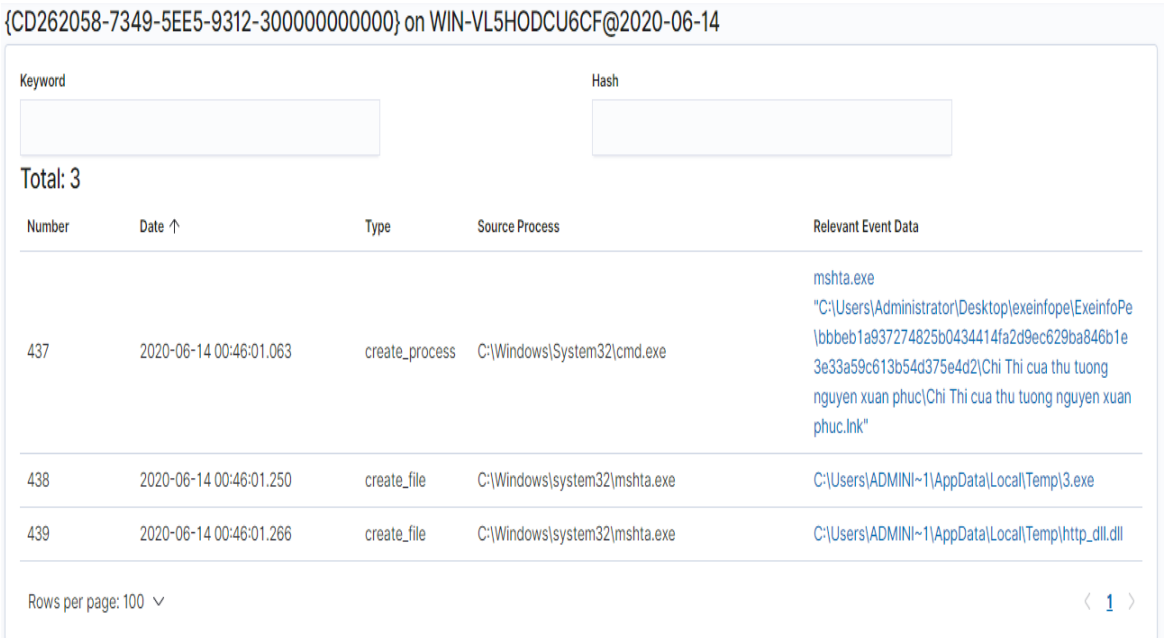
“C\users\Admin~1\AppData\Local\Temp\3.exe” và
“C\users\Admin~1\AppData\Local\Temp\http_dll.dll”

như hình dưới đây:



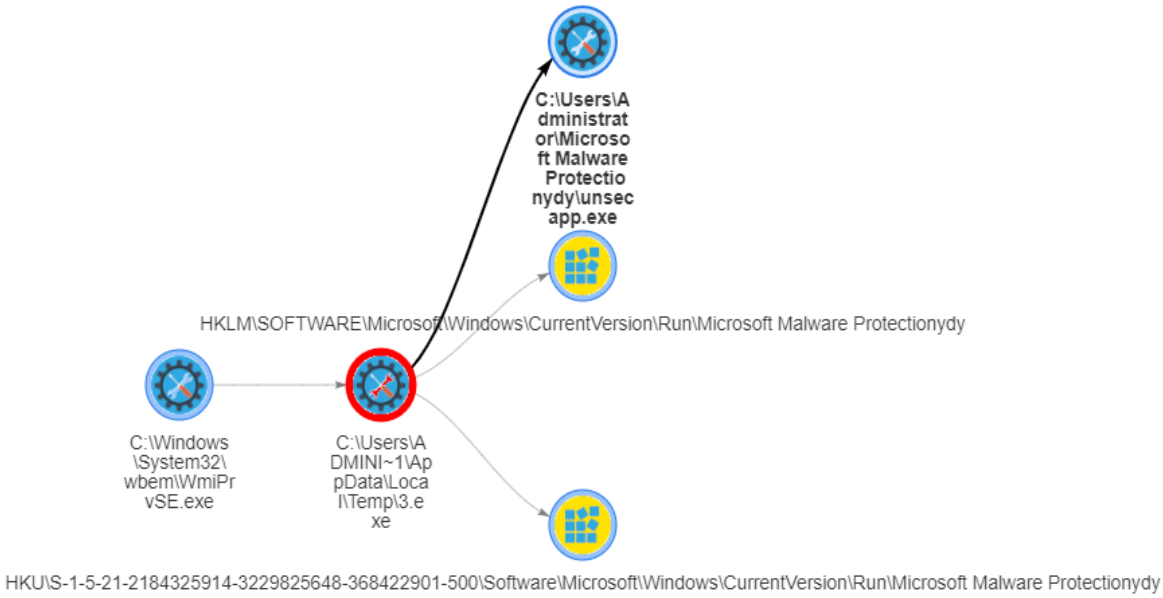
Hình 3.18. Hệ thống xây dựng lại tiến trình hoạt động mã độc

Kiểm tra thông tin chi tiết về sự kiện này cho thấy sau khi chạy tập tin mã độc, mã độc thực hiện tạo ra 2 tập tin mới là 3.exe và http_dll.dll tại đường dẫn C\users\Admin~1\AppData\Local\Temp\ như hình dưới đây:



Hình 3.19. Hệ thống hiển thị thông tin tiến trình mã độc tạo 02 tập tin

Kiểm tra thông tin chi tiết về tập tin 3.exe, luận văn nhận thấy tập tin 3.exe thực hiện tạo registry và gọi tập tin unsecapp.exe như sau:



Hình 3.20. Hệ thống hiển thị thông tin chi tiết đối với tập tin 3.exe

Thông tin chi tiết về sự kiện đối với tập tin 3.exe cho thấy rõ mã Hash của tập tin cũng như tiến trình cha đã gọi tập tin.

Hình 3.21. Hệ thống hiển thị thông tin về sự kiện tập tin 3.exe

Kiểm tra thông tin về kết nối network, luận văn không thu thập được thông tin về kết nối máy chủ điều khiển của mã độc này, qua phân tích so sánh với kết quả phân tích được công bố, nhận thấy tên miền máy chủ điều khiển đã không còn hoạt động do đó mã độc không thực hiện kết nối thành công dẫn đến hệ thống không ghi nhận được sự kiện này. Như vậy, qua việc sử dụng công cụ thu thập, phân tích tiến trình dựa trên sự kiện event của Sysmon, luận văn đã triển khai phân tích và nhận diện được tập tin mã độc.

3.4. Kết luận chương 3

Trình bày tổng quan về ứng dụng phát hiện tiến trình bất thường trên máy người dùng bao gồm: hệ thống thu thập trên máy người dùng và hệ thống giám sát tiến trình.

Mô tả chi tiết cách thức xây dựng và cài đặt hệ thống thu thập dữ liệu trên máy người dùng.

Thực hiện kiểm thử hệ thống phát hiện tiến trình bất thường sử dụng tập dấu hiệu.

KẾT LUẬN LUẬN VĂN

Trên những kết quả đã làm được luận văn có thể nghiên cứu và phát triển theo các hướng sau:

- Nghiên cứu và áp dụng một số phương pháp phân tích hành vi bất thường của tiến trình để phát hiện mã độc.

- Tìm hiểu một số công nghệ cho việc thu thập tiến trình trên máy người dùng sử dụng các hệ điều hành khác.