**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**KHOA MẠNG MÁY TÍNH VÀ TRUYỂN THÔNG**

**LÊ VIẾT TÀI MẪN – 20521593**

**HOÀNG THANH LÂM - 20521513**

**VŨ HOÀNG THẠCH THIẾT -20521957**

**PHẠM VĂN XUÂN  - 20522184**

**BÁO CÁO ĐỒ ÁN CUỐI KỲ**

**ETIP: An Enriched Threat Intelligence Platform for improving OSINT correlation, analysis, visualization and sharing capabilities**

**GIẢNG VIÊN HƯỚNG DẪN**   
**Lê Đức Thịnh**

**TP. HỒ CHÍ MINH, 2023**

Contents

[I. Giới thiệu: 2](#_Toc137331948)

[II. Enriched Threat Intelligence Platform (ETIP) 3](#_Toc137331949)

[1) Input module: 4](#_Toc137331950)

[2) Operational module 6](#_Toc137331951)

[3) Output module: 7](#_Toc137331952)

[III. Threat score evaluation 9](#_Toc137331953)

[1) Source Identification: xác định nguồn dữ liệu 9](#_Toc137331954)

[2) Heristic Identification: Xác định thuộc tính Heuristic 9](#_Toc137331955)

[3) Threshold Definition: định nghĩa ngưỡng 9](#_Toc137331956)

[4) Score Computation: Kết quả thể hiện mức độ nghiêm trọng và cần được ưu tiên 9](#_Toc137331957)

[5) Training Period: Cần một bộ test để đánh giá hiệu suất của engine. 9](#_Toc137331958)

[6) Engine Calibration: giảm độ biến thiên (giảm độ sai sót) 9](#_Toc137331959)

[7) Final test: kết quả 9](#_Toc137331960)

[IV. Demo 12](#_Toc137331961)

[1. Hướng dẫn cách cài đặt 12](#_Toc137331962)

[2. MISP 12](#_Toc137331963)

[3. Feeds 12](#_Toc137331964)

[4. Event 12](#_Toc137331965)

[5. Bên trong 1 event 13](#_Toc137331966)

[V. Tài liệu tham khảo: 15](#_Toc137331967)

1. Giới thiệu:

OSINT ( Open Source Intelligence) data là dữ liệu liên quan đến bảo mật thu thập được từ những nguồn mở có sẵn.

TIPs (Threat Intelliegence Platforms). Những platform này giúp tập hợp, liên kết và phân tích những dữ liệu bảo mật từ nhiều nguồn khác nhau để hợ trỗ xây dựng hệ thống phòng thủ.

Nhưng những dữ liệu từ nhiều nguồn khác nhau không có đồng nhất về mặt kiến trúc (unstructured data). TIPs lọc, xử lý dữ liệu trước khi được phân tích và chia sẻ.

Bài báo đề xuất mô hình Enriched Threat Intelligence Platform (ETIP) với chức năng kiến trúc hoá được các dữ liệu thu thập được từ nhiều nguồn, sau đó tìm các điểm liên kết (correlation) giữa các dữ liệu này và cuối cùng đánh giá (threat score) trên lỗ hổng mà có thể gây ra cho hệ thống và có thể chia sẻ kết quả cho các thực thể (entities) khác.

Những đóng góp của bài báo:

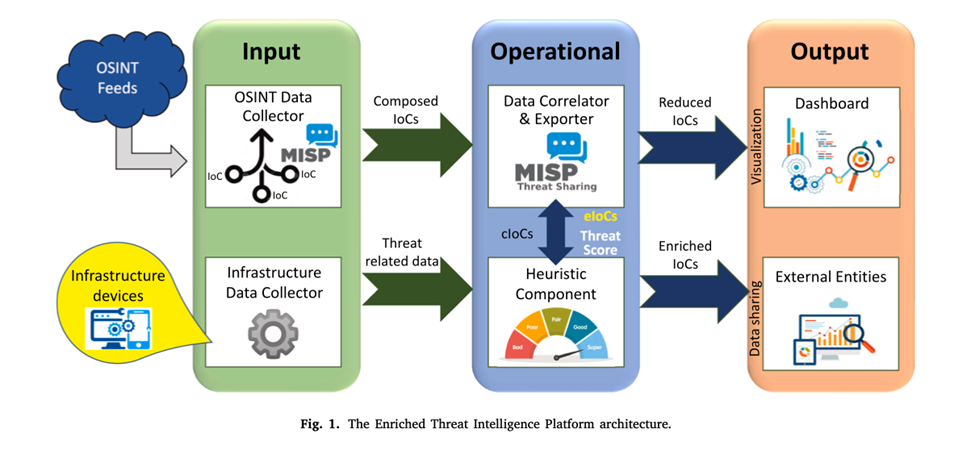
* Nền tảng ETIP mở rộng khả năng nhập và chia sẻ thông tin của các hệ thống giám sát và phát hiện nội bộ
* Một phương pháp cải thiện đánh giá (công thức tính toán threat score) để gán mức độ nguy hiểm cho dữ liệu nhận được.
* Có thể tích hợp những thông tin bảo mật liên quan đến nhau từ những nguồn public.
* Một database gồm những heuristic features để (áp dụng cho công thức tính threat score) để đánh giá mức độ nguy hiểm
* Một ví dụ ngoài thực tế.

1. Enriched Threat Intelligence Platform (ETIP)

Một số thuật ngữ chính:

* Indicator of Compromise (IoC): là dữ liệu pháp chứng, ví dụ là data thu thập được từ các system logs hoặc file, được xác định có khả năng độc hại đến mạng hoặc hệ thống
* Composed Indicator of Compromise (cIoC): là sự tổng hợp, quan hệ hoá giữa những dữ liệu OSINT ( được thu thập từ nhiều nguồn, plaintext, csv)
* Enriched Indicator of Compromise (eIoC): là những dữ liệu được làm phong phú hơn từ sự kết hợp từ nguồn tĩnh (OSINT Feeds) và nguồn thông tin động (infrastructure).   
  Kết quả của bước này là gán threat score vào cIoC.
* Reduced Indicator of Compromise (rIoC): là trích xuất những thông tin quan trọng nhất, lọc thông tin liên quan tới bảo mật từ eIoC

Mô hình:

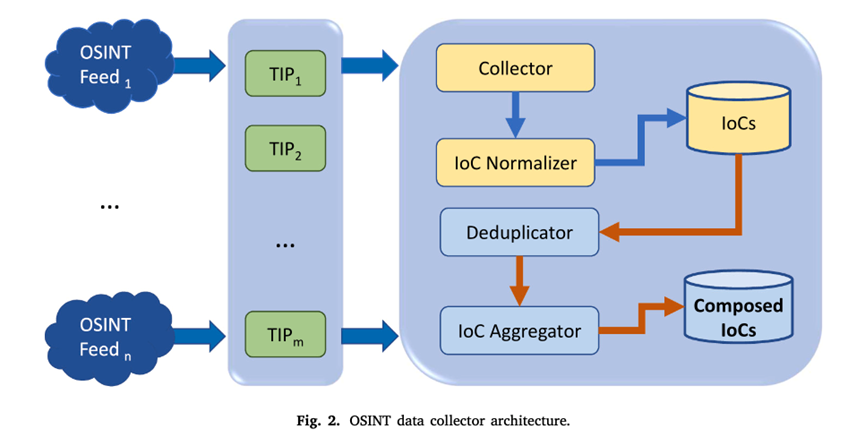


Tổng quan: Input Module bao gồm những IoC thu thập từ OSINT/ infrastructure tools. Operational Module phân tích dữ liệu tính toán threat score và tìm sự liên kết của các dữ liệu. Output module bao gồm dashboard để trực quan hoá dữ liệu và có thể chia sẻ thông tin với các thực thể khác.

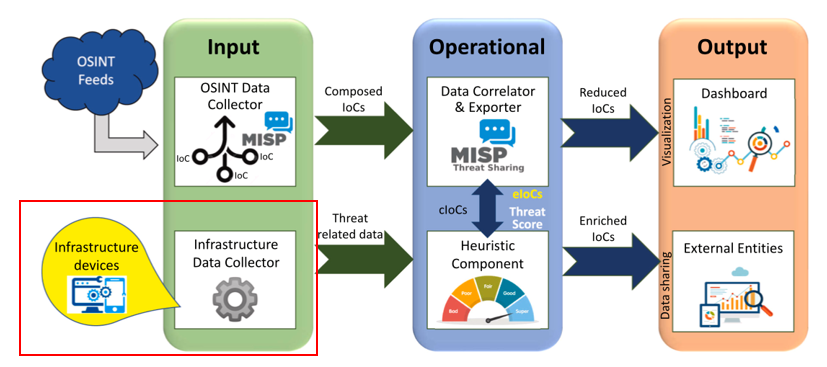
Chi tiết của mô hình

1. Input module:

* OSINT data collector:



* OSINT Feeds: được thu thập từ nhiều nguồn (cyber-attacks, malware domains, vulnerability exploitation, IP blacklist)
* Threat Intelligence Platform (TIPs): thu thập dữ liệu OSINT
* Collector: kết quả Của TIPs, hình thành nên các IoC
* IoC normalizer: Các IoC được chuẩn hoá lại theo một cấu trúc định nghĩa chung (MISP format)
* Deduplicator: Loại bỏ những IoCs bị trùng
* IoC Aggregator: tìm sự tương quan giữa IoCs và tích hợp những IoCs lại. Đồng thời, loại bỏ những IoCs giống nhau. Kết quả là 1 tập hợp gồm những IoCs. Tập hợp những IoC có liên quan với nhau nhất, kết hợp thành một tập hợp những IoC riêng lẻ, và tạo thành Composed IoCs.
  + Infrastructure data collector



Thu thập từ dữ liệu (hashes, signatures, IPs, domains, URLs) từ system log,… của những monitor infrastructure ( firewall, intrusion detection, prevention system,…)

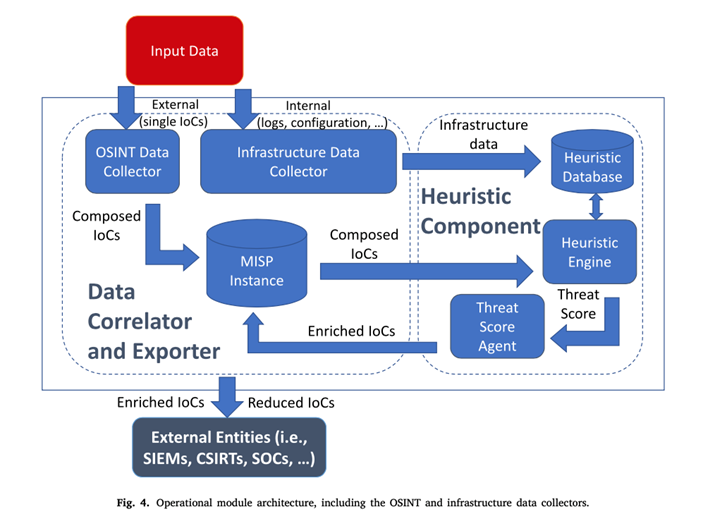
* Tạo composed IoCs từ OSINT và Infrastructure data

OSINT và infrastructure data được xây dựng dựa trên nền tảng MISP. Các IoC deduplicator và IoC aggregator từ OSINT collector được phát triển từ module MISP. MISP đóng vai trò là bộ thu thập và bộ chuẩn hóa IoC để nhận và chuẩn hóa dữ liệu có thể được cung cấp từ các TIP khác nhau (trong trường hợp dữ liệu OSINT) hoặc từ infrastructure device. Từ đó xử lý dữ liệu đã nhận, tìm sự liên kết, loại bỏ các IoC trùng lặp và cho ra kết quả composed IoCs

Input module chia nguồn cấp OSINT thành hai loại: (1) low-level, bao gồm chủ yếu là địa chỉ IP và URL; và (2) high-level, chứa phân tích nâng cao hơn với thông tin về các network artifacts, để cung cấp TIP (ví dụ: CRIT, MISP). Đầu tiên query tới database để xác định target, rồi tổng hợp để tạo thành một IoC mới và đưa vào database, được gắn nhãn bằng thẻ cho phép xác định nó là một enriched IoC và tránh việc tạo của các vòng lặp.

1. Operational module

Module này được dựa trên MISP, có khả năng liên kết hoá nguồn thông tin tĩnh ( static) và thực (dynamic). Kết quả của bước này là enriched IoC (eIoC) và đồng thời tính toán threat score.



* Data correlator & exporter:

OSINT and Infrastructure data collectors thu thập dữ liệu IoC từ OSINT và cơ sở hạ tầng được giám sát, tạo cIoC để đánh giá một cIoC để tính toán threat score. Composed IoC được lưu vào MISP database, còn dữ liệu từ infrastructure sẽ được lưu trữ trong Heuristic Component Database.

Dữ liệu được lưu trong MISP được lưu trữ dưới dạng JSON formats hoặc là một dạng doc cơ bản. MISP còn có cơ chế tự động chia sẻ những thông tin này với những thực thể khác bằng cách đồng bộ hoá các MISP instance.

* Heuristic component:

Heuristic component nhận thông tin từ nhiều nguồn ( OSINT data, infrastructure, IoCs,…) để sử dụng tính toán threat score được thực hiện bởi heuristic engine. Điểm đánh giá sẽ được gán cho từng heuristic feature. Và tổng của từng score sẽ là Threat Score cho từng dữ liệu được phân tích.

1. Output module:

Kết quả của module này là trực quan hoá những thông tin liên quan và quan trọng nhất chứa trong rIoC. Các eIoC có thể chứa một lượng rất lớn những thuộc tính để tính threat score, và điều này gây khó khăn cho việc vẽ graph nên chắc lọc lại để ra reduced IoC.

* Dashboard

A screenshot of a computer

Description automatically generated with medium confidence

Biểu đồ cung cấp biểu diễn đồ họa về cấu trúc liên kết infrastructure bằng cách làm nổi bật các alarm và rIoC cho một node.

Ở một node, thể hiện cho số lượng alarm và màu sắc cho mức độ nghiêm trọng ( hình tròn nhỏ) và số lượng rIoC (hình ngôi sao)

Alarm thể hiện số lượng issues, IP source và destination, và bảng miêu tả những issue. rIoC thể hiện số lượng lỗ hổng, threat score, và bảng miêu tả các lổ hổng và những chương trình có thể bị ảnh hưởng.

* External entities

Việc trao đổi eIoC được thực hiện thông qua MISP, tự động chuyển đổi tất cả thông tin nhận được sang định dạng MISP JSON và lưu trữ nó trong MISP database. Định dạng JSON luôn được sử dụng bất cứ khi nào MISP instance trao đổi thông tin. Tuy nhiên, khi chia sẻ với các thực thể bên ngoài không sử dụng MISP, cũng như các hệ thống không thể xử lý trực tiếp định dạng MISP, nên sử dụng các tiêu chuẩn khác, cũng để mô tả một bộ TI rộng hơn.

1. Threat score evaluation

Quá trình tính toán threat score:

1. Source Identification: xác định nguồn dữ liệu
2. Heristic Identification: Xác định thuộc tính Heuristic
3. Threshold Definition: định nghĩa ngưỡng
4. Score Computation: Kết quả thể hiện mức độ nghiêm trọng và cần được ưu tiên
5. Training Period: Cần một bộ test để đánh giá hiệu suất của engine.
6. Engine Calibration: giảm độ biến thiên (giảm độ sai sót)
7. Final test: kết quả

Công thức tính threat score

A black text on a white background

Description automatically generated with low confidence

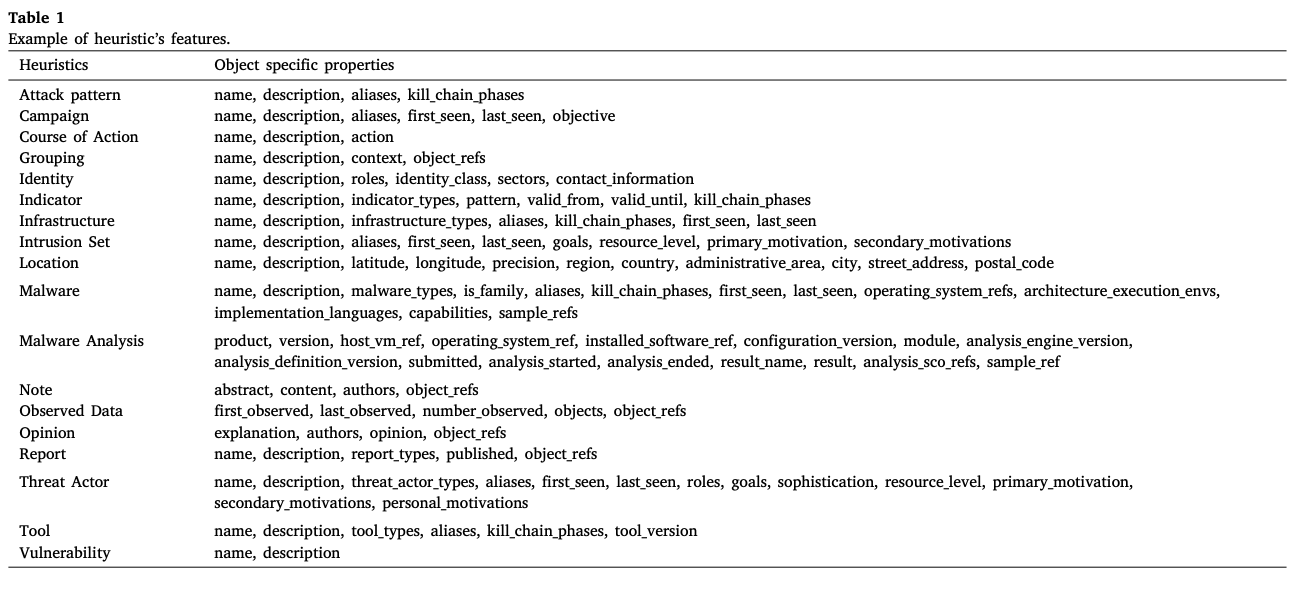
Cp  : completeness criterion

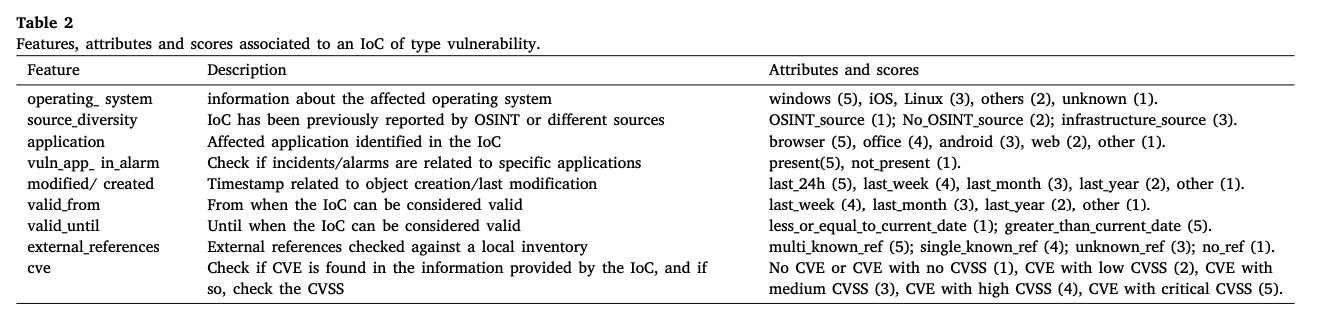
Xi : giá trị của từng heuristic’s feature

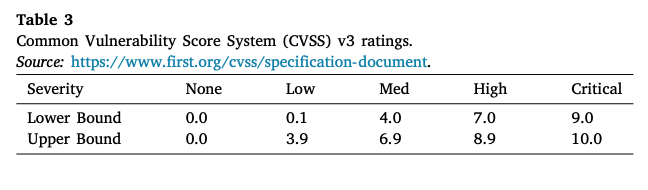
Pi : Weighting Criteria

Kết quả: 1 <= TS <= 5: thể hiện mức độ nghiêm trọng của IoC

Heuristic features và giá trị của heuristic feature







Khi một thông tin được đưa vào mô hình, dữ liệu sẽ được trích xuất và so sánh với các heuristic features mà tác giả đã định nghĩa sẵn và gán giá trị cho những dữ liệu đó để tính threat score.

Case study: MySQL server vulnerability  
  
- Một IoC được thu thập với trường thông tin: CVE-2019-2834: lổ hổng trong thành phần MySQL Server của Oracle MySQL, mức độ nghiêm trọng của lổ hổng là medium, giá trị CVSS v3.0 bằng 6.5.  
  
- Input module nhận và xử lý IoC này thông qua Data Correlator & Exporter ( module của MISP). Thông qua Operational module và tính toán threat score.

A screenshot of a computer

Description automatically generated with medium confidence

* Kết quả:
* - node bị ảnh hưởng
* - 2 alarm ( low severity)
* - 4 rIoC
* - node type, IP address, OS.
* - CVE-2019-2834 vulnerability
* - threat score

1. Demo
2. Hướng dẫn cách cài đặt

* Cách cài đặt có thể xem chi tiết ở video sau: [Link](https://www.youtube.com/watch?v=Etr9vVC6WYQ&t=752s)
* Tài liệu chi tiết về MISP: [Document](https://www.circl.lu/doc/misp/)

1. MISP

* Là giải pháp phần mềm mã nguồn mở để thu thập, lưu trữ, phân phối, chia sẻ các chỉ số và mối đe dọa về an ninh mạng, đặc biệt liên quan đến phân tích sự cố an ninh mạng và phân tích phần mềm độc hại
* MISP cung cấp các chức năng để hỗ trợ việc trao đổi thông tin và tiêu thụ thông tin đó bởi các hệ thống phát hiện xâm nhập mạng NIDS, hệ thống phát hiện xâm nhập cục bộ LIDS cũng như các công cụ phân tích nhật ký và hệ thống SIEM

1. Feeds

* Là các nguồn thông tin cung cấp dữ liệu về các mối đe dọa và các chỉ số an ninh mạng
* Các feeds này thường được tạo ra và duy trì bởi các tổ chức an ninh mạng, nhà cung cấp dịch vụ an ninh hoặc cộng đồng an ninh mạng
* Feeds cung cấp thông tin về các sự kiện, định danh và các thông tin liên quan đến các mối đe dọa, bao gồm IP độc hại, domain độc hại, mã độc hại, chữ ký tấn công, APT. Thông tin này được thu thập từ nhiều nguồn, bao gồm các trang web đáng tin cậy, tổ chức an ninh, cơ quan chính phủ, các nhóm cộng đồng an ninh và các công cụ phân tích an ninh
* Feeds được sử dụng để cung cấp thông tin về các mối đe dọa mới nhất và biểu hiện của các cuộc tấn công đang diễn ra

1. Event

* Event chứa các thông tin liên quan đến các yếu tố và chỉ số an ninh mạng, như các tấn công, mã độc, thông tin về kẻ tấn công, mẫu mã độc, chiến dịch tấn công, hoạt động độc hại, hoặc bất kỳ thông tin an ninh mạng nào khác.
* Được lấy từ feeds ta đã cung cấp, chứa meta data về các apt group, ransomeware, malware đã biết
* MISP phân loại cho ta các event để phân tích dễ dàng hơn như. Ví dụ như phân tích malware ta có thể xác định giá trị băm của các tập tin, tên domain, tên file, địa chỉ IP máy chủ,…

1. Bên trong 1 event

Ảnh có chứa văn bản, phần mềm, Biểu tượng máy tính, Trang web

Mô tả được tạo tự động

Ảnh có chứa văn bản, ảnh chụp màn hình, phần mềm, Trang web

Mô tả được tạo tự động

* Phân tích

- Một số đối số commandline có liên quan đến phần mềm độc hại này. Có thể đây là 1 commandline được dùng để leo thang đặc quyền.

+ Như là ping tới địa chỉ loopback tạo delay

+ Sau đó vô hiệu hóa tính năng của window defender

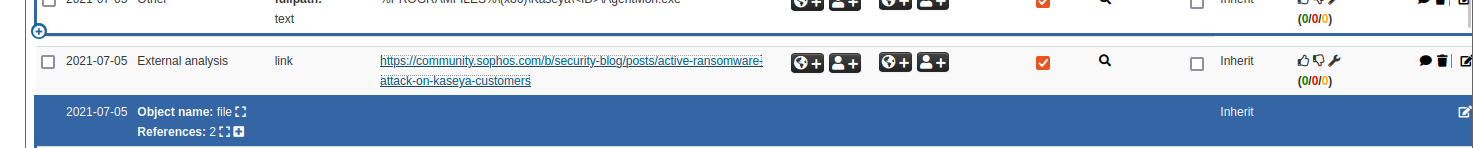
+ Sao chép tệp cerutil.exe từ system32 đến thư mục gốc của windows

+ Ghi giá trị ngẫu nhiên vào tệp cert.exe trong thư mục gốc của windows

+ Giải mã tệp agentcrt từ tư mục c:\kworing và lưu kết quả vào tệp agent.exe trong cùng thư mục

+ Xóa tệp agent.crt và cert.exe khỏi hệ thống

+ Thực thi tệp agent.exe từ thư mục c:\kworking



- Link trực tiếp dẫn đến bài blog nói về malware này

- Thay vì ta phải google hay tìm các thông tin liên quan thì MISP đã cung cấp cho ta đầy đủ thông tin liên quan và tập trung lại

Ảnh có chứa văn bản, Phông chữ, số, hàng

Mô tả được tạo tự động

- Dll độc hại ta cũng có thể tìm thấy khi phân tích 1 event trong MISP mà phần mềm độc hại cố tìm cách chiến quyền điều khiển

Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ

Mô tả được tạo tự động

- Ta cũng có thể xem các event có liên quan đến event này

Ảnh có chứa văn bản, ảnh chụp màn hình, phần mềm, số

Mô tả được tạo tự động

- Có thể attack event này liên quan đến event kia như là sử dụng các domain độc hại thông dụng, command thông dụng, payload tương tự chẳng hạn

- Ta có thể xem các mối liên hệ từ các thông tin mà MISP lấy được

- Và tất cả điều đó ta thông cần phải làm thủ công mà đã được MISP hỗ trợ 1 cách tự động

- Và tất nhiên khi xử lý 1 phần mềm độc hại thì việc nhận biết càng sớm thì càng dễ phát hiện và kịp thời ngăn chặn thay vì ta phải thủ công research các thông tin liên quan để nhận biết nó có phải phần mềm độc hại hay không

1. Tài liệu tham khảo:

* ETIP: An Enriched Threat Intelligence Platform for improving OSINT