


THÔNG TIN CHUNG CỦA BÁO CÁO

- Link YouTube video của báo cáo (tối đa 5 phút):
<https://www.youtube.com/watch?v=6WQfGMJtOuE>
- Link slides (dạng .pdf đặt trên Github):
<https://github.com/Tothimyth/CS2205.CH181/blob/main/A%20MACHINE%20LEARNING-BASE%20KNOWLEDGE%20GRAPH%20CONSTRUCTION%20FROM%20CYBER%20THREAT%20INTELLIGENCE%20FOR%20APT%20ATTRIBUTION.pdf>
- Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới
- Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in

<ul style="list-style-type: none">• Họ và Tên: Tô Thị Mỹ Âu• MSSV: 230202002 	<ul style="list-style-type: none">• Lớp: CS2205.APR2023• Tự đánh giá (điểm tổng kết môn): 9/10• Số buổi vắng: 1• Số câu hỏi QT cá nhân: 3• Link Github: https://github.com/mynameuit/CS2205.APR2023/
---	---

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

XÂY DỰNG ĐỒ THỊ TRI THỨC DỰA TRÊN HỌC MÁY TỪ THÔNG TIN ĐE DỌA MẠNG ĐỂ XÁC ĐỊNH NGUỒN GỐC APT

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

A MACHINE LEARNING-BASE KNOWLEDGE GRAPH CONSTRUCTION FROM CYBER THREAT INTELLIGENCE FOR APT ATTRIBUTION

TÓM TẮT (Tối đa 400 từ)

Với tình hình diễn biến của không gian mạng hiện nay, các cuộc tấn công mạng trở nên đa dạng và tinh vi hơn. Hơn thế nữa, các mối đe dọa liên tục tiến tiến (APTs - Advanced Persistent Threats) đã gây ra những mối đe dọa an ninh nghiêm trọng trên toàn thế giới. Do đó, các tình báo mối đe dọa mạng (CTI - Cyber threat intelligence) đang ngày càng có ảnh hưởng trong việc thu thập thông tin bảo mật mạng hiện tại. Đồ thị tri thức về an ninh mạng nhằm thay đổi cách biểu đạt kiến thức về mối đe dọa để các nhà nghiên cứu bảo mật có thể thu thập chính xác và hiệu quả các loại thông tin đe dọa khác nhau để đưa ra quyết định nhanh chóng và hiệu quả. Công nghệ truy tìm nguồn gốc không chỉ hỗ trợ các chuyên gia bảo mật trong việc phát hiện các APT, mà còn có thể xác định cùng một mối đe dọa từ các sự kiện tấn công khác nhau. Do đó, việc truy tìm tác nhân đe dọa tấn công là quan trọng. Trong **hình 1**, bằng việc sử dụng công nghệ đồ thị tri thức, xem xét các nghiên cứu mới nhất về truy tìm nguồn gốc tấn công mối đe dọa mạng và tìm hiểu các công nghệ và lý thuyết liên quan trong quá trình xây dựng và áp dụng đồ thị tri thức APT từ CTI. Nghiên cứu này thực hiện đề xuất việc “**Xây dựng đồ thị tri thức dựa trên học máy từ thông tin mối đe dọa để xác định nguồn gốc APT**”. Là một mô hình đồ thị tri thức về APT dựa trên các kịch bản tấn công APT thực tế. Sau đó, thực hiện thiết kế một thuật toán trích xuất kiến thức về mối đe dọa APT để hoàn thiện và cập nhật đồ thị tri thức bằng cách sử dụng học sâu. Cuối cùng, đề xuất một phương pháp thực tế về truy tìm nguồn gốc tấn công

APT với sự truy tìm nguồn gốc và các biện pháp phòng ngừa. Đồ thị này không phải là một phương pháp phòng thủ bị động trong mạng truyền thống mà là một phương pháp tích hợp một lượng lớn thông tin tình báo phân tán và có thể điều chỉnh chủ động chiến lược phòng thủ của mình. Nó đặt nền tảng trong việc phòng thủ mạng.



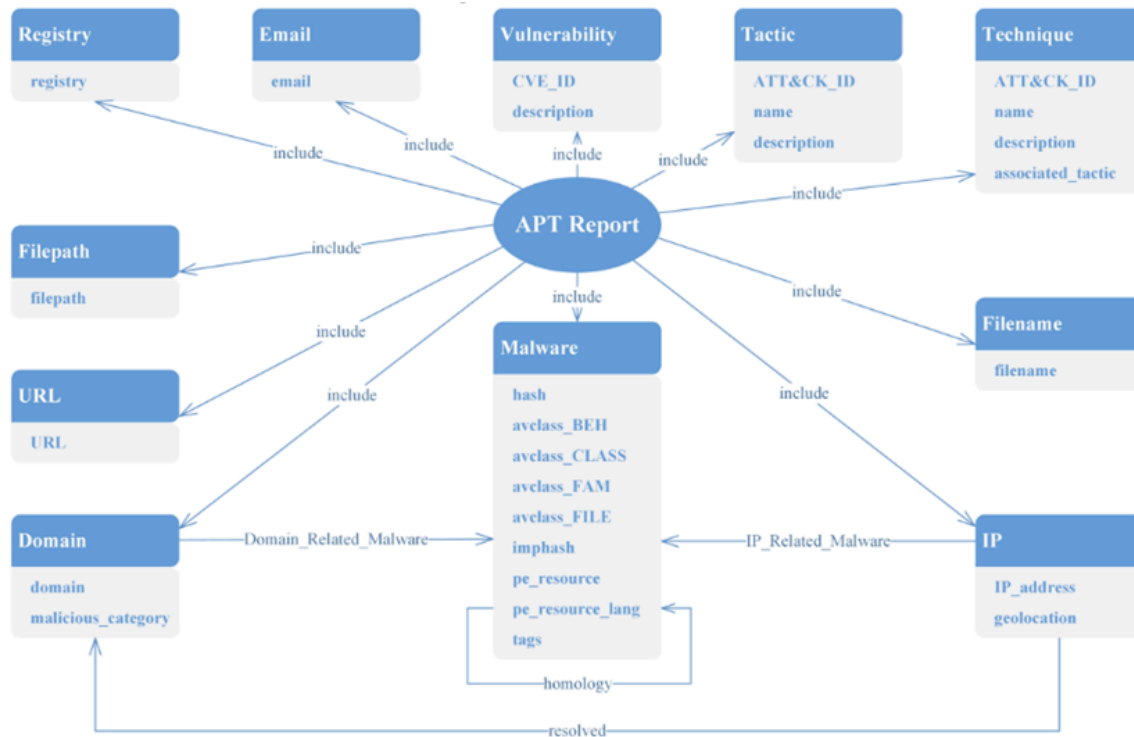
Hình 1 Minh họa đồ thị xác định nguồn gốc APT [1].

GIỚI THIỆU (Tối đa 1 trang A4)

Các cuộc tấn công mạng ngày càng sử dụng các chiến thuật ngày càng tinh vi và các kỹ thuật đa dạng. Chẳng hạn như các APT, làm cho việc phát hiện trở nên khó khăn hơn bao giờ hết.

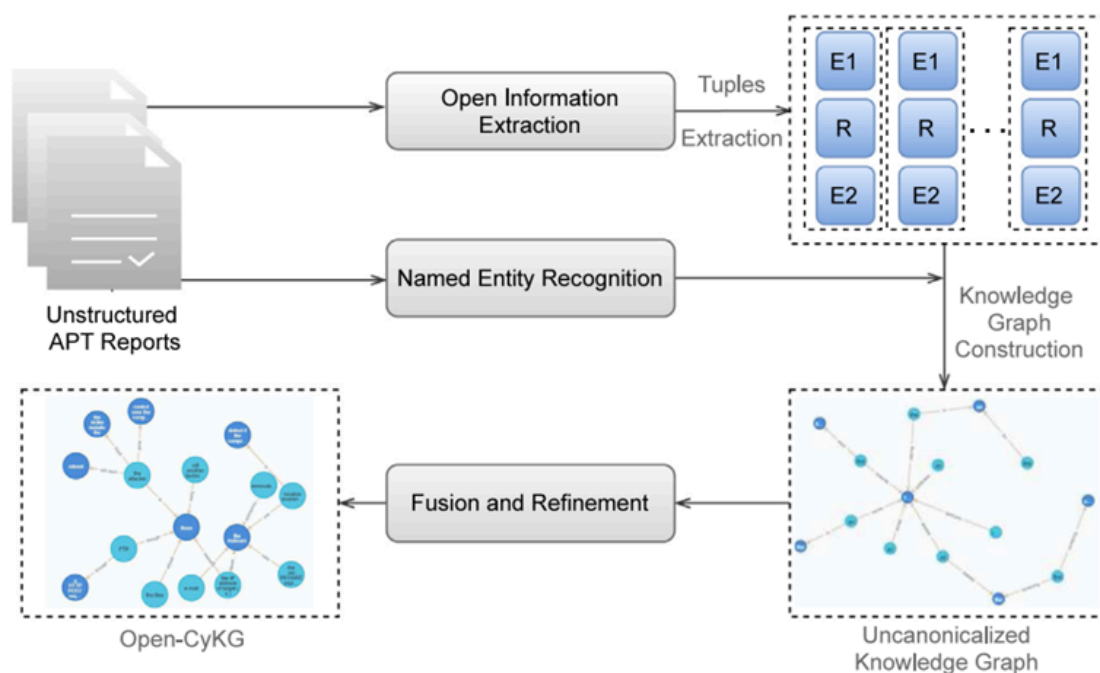
Vì thế nên, CTI cũng trở nên quan trọng trong việc nâng cao khả năng chống lại các mối đe dọa mạng [5]. Việc chia sẻ thông tin tình báo kịp thời và phân tích thông tin tình báo chính xác **hình 3** là chìa khóa để đối phó hiệu quả đối với các mối đe dọa mạng, nó giúp rút ngắn thời gian đối phó với mối đe dọa [4]. Một trong những

hình thức CTI là các bản báo cáo tình báo mỗi đe dọa mạng không cấu trúc dựa trên văn bản nguồn mở (OSCTI). Các báo cáo OSCTI là các phân tích sâu về các sự kiện tấn công APT.



Hình 3 Lược đồ của việc xác định nguồn gốc APT [1].

Bên cạnh đó, các báo cáo OSCTI được thu thập từ nhiều nguồn dữ liệu, có chất lượng không đồng đều, các kịch bản ứng dụng phức tạp và không có cấu trúc. Vậy nên, việc trích xuất thực thể và mối quan hệ từ các báo cáo OSCTI là một bước quan trọng. Do đó, các công cụ đã được thiết kế giúp trích xuất thông tin tấn công tự động từ các báo cáo OSCTI **hình 4**. Trong nghiên cứu về các kịch bản tấn công APT, phát hiện tấn công dựa trên theo dõi bằng đồ thị (bằng cách học hành vi tấn công) dần thay thế phát hiện bất thường (bằng cách học hành vi bình thường thông qua kiểm tra nhật ký) như một phương pháp hiệu quả hơn. Việc nghiên cứu về việc xác định nguồn gốc của các sự kiện tấn công so với việc phát hiện và suy luận tấn công bằng các nhật ký trong lĩnh vực phát hiện tấn công APT cũng rất quan trọng. Nếu không xác nhận được danh tính và hành vi của các mối đe dọa, không thể dự đoán một sự kiện tương tự.



Hình 4 Quy trình để xây dựng đồ thị tri thức từ CTI [2] .

Để giải quyết các vấn đề trên, thực hiện đề xuất "**Xây dựng đồ thị tri thức dựa trên học máy từ thông tin mối đe dọa để xác định nguồn gốc APT**", một nền tảng mô hình hóa mối đe dọa không gian mạng dựa trên biểu đồ tri thức. Mô hình trích xuất các thực thể thông tin mối đe dọa liên quan bằng cách thu thập và phân tích thông tin mô tả và mối quan hệ logic của tổ chức tấn công APT trong OSCTI. Bằng cách kết hợp STIX, CYBOX và các tiêu chuẩn tình báo mối đe dọa khác, thực hiện xây dựng một biểu đồ tri thức về APT. Thiết kế thuật toán trích xuất thực thể từ các báo cáo CTI để cập nhật kiến thức biểu đồ tri thức trên. Cuối cùng, áp dụng mô hình kim cương và săn mối đe dọa để đề xuất một kế hoạch theo dõi tổ chức tấn công APT cũng như xác định nguồn gốc APT.

Input: Các nguồn **thông tin CTI**.

Output: **Đồ thị tri thức từ CTI** có phân tích các tác nhân tấn công nhằm xác định nguồn gốc tấn công.

MỤC TIÊU

(Viết trong vòng 3 mục tiêu, lưu ý về tính khả thi và có thể đánh giá được)

- Nghiên cứu về phương pháp để xây dựng đồ thị tri thức trong an ninh mạng, các nguồn dữ liệu để xây dựng đồ thị tri thức trong an ninh mạng và ứng dụng của đồ thị tri thức trong việc xác định nguồn gốc tấn công.
- Nghiên cứu về thuật toán để thực hiện thu thập, phân tích và trích xuất thông tin về mối đe dọa. Từ đó tạo ra đồ thị tri thức thông tin về đe dọa mạng và mối quan hệ của chúng.
- Dựa trên đồ thị tri thức và thông tin thu thập được thực hiện xác định nguồn gốc của APT. giúp tăng cường khả năng phát hiện, ngăn chặn kịp thời trước các cuộc tấn công APT.

NỘI DUNG VÀ PHƯƠNG PHÁP

(Viết nội dung và phương pháp thực hiện để đạt được các mục tiêu đã nêu)

Nội dung 1: Nghiên cứu tìm hiểu về CTI, các tiêu chuẩn của CTI, phương pháp phân tích CTI bằng cách tự động hóa.

Phương pháp:

- Thu thập các nguồn tài liệu về các tiêu chuẩn CTI, tìm hiểu về các tiêu chuẩn phổ biến như: STIX, TAXII, CyBox, MISP và OpenIOC.
- Tìm hiểu các phương pháp trong việc tự động hóa trích xuất và phân loại thực thể như [6]: **phương pháp học máy thống kê**: Sử dụng các bộ phân loại máy vector hỗ trợ (support vector machine), phương pháp CRF; **phương pháp học sâu**: Sử dụng mô hình CNN-CRF, mô hình nhận dạng thực thể BiLSTM-CRF và mô hình LSTM-CRF; **phương pháp lý ngôn ngữ tự nhiên (NLP)**: Sử dụng các phương pháp biến đổi nhiệm vụ trích xuất thông tin từ văn bản thành các bài toán tương tự đồ thị. Sử dụng cơ chế chú ý (attention mechanism), công cụ truy xuất thông tin để trích xuất và phân loại các hành động đe dọa, các chỉ số IOCs và các khái niệm liên quan trong mạng bảo mật.

Nội dung 2: Nghiên cứu, tìm hiểu về việc sử dụng đồ thị tri thức về an ninh mạng.

Phương pháp: Dựa vào các nghiên cứu đã công bố, thu thập thông tin về GNN và quy trình

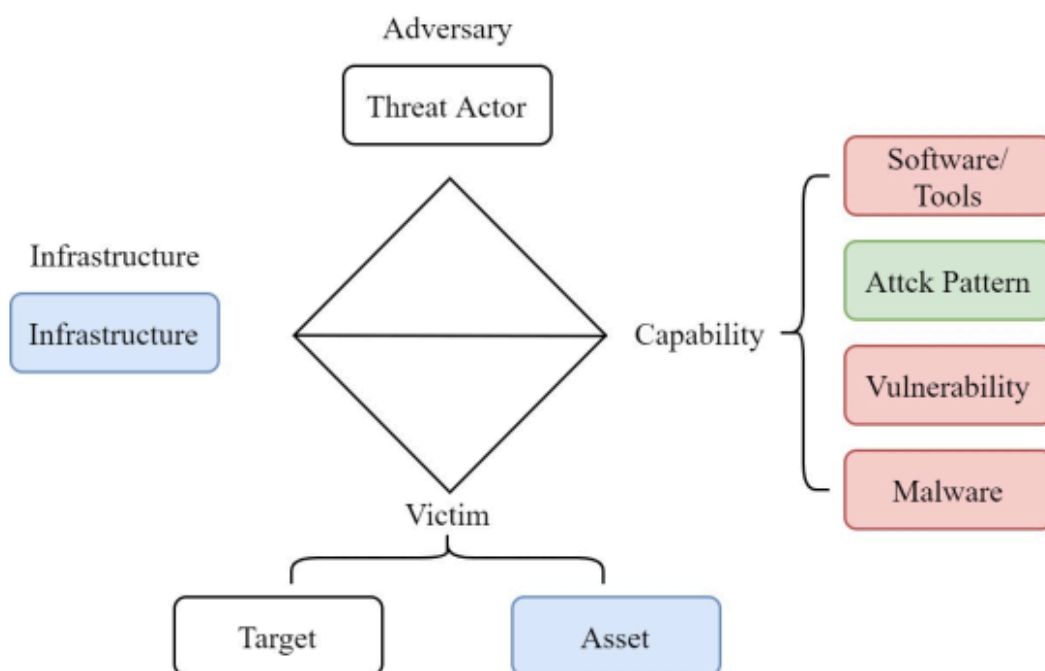
thiết kế mô hình GNN, KG và ứng dụng của nó trong lĩnh vực an ninh thông tin và trong việc xác định nguồn gốc APT.

Nội dung 3: Nghiên cứu các phương pháp xác định nguồn gốc mối đe dọa

Phương pháp: Nghiên cứu các phương pháp xác định nguồn gốc tấn công APT gồm: trích xuất thông tin quan trọng (IoC) từ tài liệu OSCTI, cách xây dựng ma trận tương quan giữa các thông tin quan trọng để huấn luyện mô hình phân loại, cách thực hiện dự đoán và xác định nguồn gốc.

Nội dung 4: Từ những phương pháp đã thực hiện nghiên cứu, đề xuất kiến trúc cụ thể cho việc xây dựng đồ thị tri thức để xác định nguồn gốc APT. thực hiện ứng dụng vào xác định nguồn gốc APT như **hình 5**.

Phương pháp: Đề xuất một mô hình biểu diễn dữ liệu chung để mô tả hành vi và đặc điểm của các tổ chức APT, thu thập thông tin, phân tích thông tin trí tuệ mối đe dọa, xây dựng mô hình đồ thị tri thức mối đe dọa và áp dụng nó cho việc xác định nguồn gốc, tổ chức tấn công APT.



Hình 5 Mô hình kim cương để ứng dụng trong việc xác định nguồn gốc tấn công APT [3]

Nội dung 5: Xây dựng và thực nghiệm một hệ thống hoàn chỉnh, dựa trên các phương pháp đã đề xuất.

Phương pháp:

- Thu thập thông tin về mối đe dọa APT từ các nguồn khác nhau và xây dựng một tập dữ liệu tri thức về mối đe dọa APT để hành thực nghiệm.
- Từ kiến trúc đã đề xuất, thực hiện tiến hành xây dựng hệ thống trích xuất dữ liệu.
- Liên kết với mô hình tấn công diamond với các sự kiện tấn công đã trích xuất ra được để biểu diễn rõ ràng hơn về nguồn gốc tấn công APT.

Nội dung 6: Đánh giá kết quả của hệ thống vừa đề xuất với các hệ thống khác.

Phương pháp: Thực hiện tổng hợp, so sánh mô hình nhận dạng thực thể đề xuất với các mô hình khác với cùng một bộ dữ liệu.

KẾT QUẢ MONG ĐỢI

(Viết kết quả phù hợp với mục tiêu đặt ra, trên cơ sở nội dung nghiên cứu ở trên)

- Hiểu về các phương pháp xây dựng và ứng dụng đồ thị tri thức trong lĩnh vực an ninh mạng.
- Tài liệu nghiên cứu khảo sát và thiết kế hệ thống đồ thị tri thức biểu diễn các mối đe dọa, xác định nguồn gốc tấn công APT.
- Báo cáo về xây dựng và thực nghiệm đồ thị tri thức dựa trên học máy từ thông tin mối đe dọa để xác định nguồn gốc APT.

TÀI LIỆU THAM KHẢO (Định dạng DBLP)

- [1]. Xiao, N., Lang, B., Wang, T., & Chen, Y. (2024). APT-MMF: An advanced persistent threat actor attribution method based on multimodal and multilevel feature fusion. *arXiv preprint arXiv:2402.12743*.
- [2]. Sarhan, I., & Spruit, M. (2021). Open-cykg: An open cyber threat intelligence knowledge graph. *Knowledge-based systems*, 233, 107524.
- [3]. Ren, Y., Xiao, Y., Zhou, Y., Zhang, Z., & Tian, Z. (2022). CSKG4APT: A cybersecurity

knowledge graph for advanced persistent threat organization attribution. *IEEE Transactions on Knowledge and Data Engineering*.

[4]. Yue, L., Liu, P., & Wang, H. (2020). Overview of network security threat intelligence sharing and exchange. *Computer research and development*, 57(10), 2052.

[5]. Tian, Z. (2020). Detection and traceability of high covert unknown threats in cyberspace. *Information and communication technology*, 14(06), 4-7.

[6]. Joshi, A., Lal, R., Finin, T., & Joshi, A. (2013, September). Extracting cybersecurity related linked data from text. In *2013 IEEE Seventh International Conference on Semantic Computing* (pp. 252-259). IEEE.