Extending Threat Playbooks for Cyber Threat Intelligence: A Novel Approach for APT Attribution

Kelsie Edie
United States Military Academy
West Point, NY, USA
kelsie.edie@westpoint.edu

Cole Mckee
United States Military Academy
West Point, NY, USA
cole.mckee@westpoint.edu

Adam Duby
United States Military Academy
West Point, NY, USA
adam.duby@westpoint.edu

***Abstract*—Existing malware classifiers focus on classification accuracy. Additional goals of malware analysis include acquiring tactical intelligence (mitigations and detections). As such, the full utility of ML in malware classification is not realized in current techniques. We engineer STIX compliant attack pattern objects throughout the feature space, which enables pivoting into other frameworks for data fusion and improved situational awareness.**

***Index Terms*—Malware analysis, cyber threat intelligence, feature engineering**

I. Introduction

In recent years, advanced persistent threats (APTs) have carried out highly coordinated multi-stage cyber attacks with increasing sophistication and frequency. The evolving cyber threat landscape and the complexity of attacks requires expert domain knowledge to process and synthesize cyber threat intelligence (CTI). This challenge is amplified by the unstructured nature of CTI reporting on attack patterns and threat-related intelligence. A comprehensive CTI program that uses structured intelligence representation and data analytics to augment manual analysis enables organizations to more rapidly identify and assess potential cyber threats.

In this paper, we propose an analytic approach toward cyber attack attribution to identify, assess, and visualize the relationship between APTs and their associated tactics, techniques, and procedures (TTPs). Our approach uses association rule mining to identify statistical TTP relationships, and a weighted Jaccard similarity index for threat attribution. Specifically, we create extended threat playbooks comprised of weighted associations between techniques used by APT groups to accomplish a tactical goal which can then be used to identify "unknown APTs" based on the TTPs they employ during a cyber attack.

We utilize the widely accepted MITRE Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) framework [1] to gather real-world information on Advanced Persistent Threat (APT) groups and their associated tactics, techniques, and procedures (TTPs). The framework defines common threat group names, their motivations, common software used, and known techniques associated with each group. By adopting a standard naming convention for techniques, ATT&CK ensures consistency in structured intelligence reporting and analytics.

Our implementation adopts the Structured Threat Information eXpression (STIX) framework [2] to standardize the intelligence data. Standardization advances threat intelligence data collection and synthesis, leading to more precise and meaningful results. Further, it allows for analytical pivoting, an intelligence process that involves merging data to gain new insights about a subject. By using STIX objects, our threat intelligence models can be conveniently integrated into existing CTI processes or frameworks.

We visualize our STIX objects to lift the tactical attack data into strategically relevant and interpretable intelligence using Neo4j, which is a graph database management system that efficiently stores, queries, and retrieves complex graph data. As such, we can represent the relationships between APT groups and their TTPs in a knowledge graph format that provides interpretable visualization and analysis.

A CTI program that includes attribution exposes threat groups that are actively targeting and engaging the defended asset. Our approach enables analysts to identify, assess, and visualize the relationships between APTs and their TTPs based on association rule mining and a weighted Jaccard similarity index. By creating extended threat playbooks using the already weighted associations between techniques used by APT groups, our approach can be used to identify "unknown APTs" based on TTPs.

In this paper, we make the following contributions:

- create a dataset of APT threat playbooks that incorporate the statistical analysis of TTP associations;
- propose a novel approach to APT attribution;
- our data and implementation are made publicly avail-

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Military Academy, the United States Army, the Department of Defense, or the United States Government.

able[1].

The rest of this paper is organized as follows. Section II discusses background and related work. Section III describes our approach. In Section IV-B, we present our evaluation and results. Section V provides a discussion on visualizing threat playbooks, limitations, and future work. Finally, we conclude in section VI.

II. Background

A. Cyber Threat Intelligence

Cyber threat intelligence (CTI) involves collecting, analyzing, and synthesizing data to create situational awareness [3]. This informs decision makers on threat related risks to their operating environment. There are several widely adopted frameworks used to describe and communicate CTI. For example, the cyber kill-chain is used to describe an offensive attack in phases [4]. This concept was included in MITRE'S ATT&CK framework [5], which proposes an ontology of TTPs organized by phase. These TTPs are used to describe the capabilities deployed by an adversary over a multi-phase campaign [6]. TTP intelligence provides an abstraction to describe adversary actions and has been shown effective in advancing threat hunting operations [7].

An adversary can be defined as a function of capability and intent [8]. While intent informs their targeting process (who they attack), capabilities influence how they go about advancing their tactical and strategic objectives. An adversary's capability arsenal, or playbook, is often represented as a set of TTPs that the adversary has been known to deploy. Historical incident response analysis and forensic investigations inform current understandings on the adversary's playbook [9].

B. Attribution

Cyber attack attribution is the process of determining an attacker's identity [10]. Attribution can be conducted at varying levels of granularity, from identifying the human operators, to attributing criminal or nation state APT groups [11], [12]. For our purposes, we focus on APT attribution. Correctly attributing an attack to an APT is critical to maintain an organization's defensive posture, deter future attacks, and enables a nation state's right to self-defense [13].

Traditional data points used to inform attribution include indicators of compromise (IOCs) and code similarity in malware. IOCs, such as IP addresses and network based artifacts can be reused by an adversary and help identify physical and logical origins. Despite this, IP spoofing and dynamic domain generation algorithms (DGA) make these IOCs brittle without corroborating evidence [14]. Attacks involving malware can be attributed using code similarity. Indeed, attackers reuse malware payloads across different campaigns. As such, code reuse detection is a promising technique. However, adversarial malware introduces randomized diversity, amplifying the challenge of malware classification [15].

[1]https://github.com/KelsieEdie/Extending-Threat-Playbooks-for-APT-Attribution

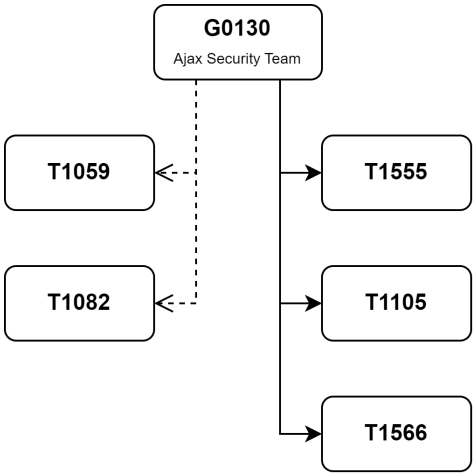

Fig. 1: Excerpt from threat and extended threat playbooks created on G0130. Solid arrows represent seed TTPs and dashed arrows represent hypothesized TTPs.

Popular frameworks for threat analysis incorporate more comprehensive features. The diamond model of intrusion analysis describes adversary activity in terms of APT, capabilities (TTPs), victim, and infrastructure [8]. Similarly, the triangle model of attribution proposes a framework that focuses on TTPs and the relationship between the adversary and victim [16].

C. Motivation

Despite promising advances in standardized threat intelligence and threat modeling, data analytics that inform attack attribution remains an open problem. Our research focuses on using TTP-based CTI to learn statistical relationships between TTPs. Using these learned relationships, we augment known threat TTP playbooks with additional TTPs that an adversary may reasonably deploy. This can anticipate adversary maneuvers and help identify unknown unknowns [17]. We focus on TTPs because they operate at the upper layers of the pyramid of pain [18], which suggests that it is more challenging for an adversary to change their TTPs instead of tactical IOCs [16], making them more robust to adversarial shifts.

III. Approach

In this section, we describe our approach to TTP-centric APT attribution. An overview is visualized in figure2 , and involves the following steps:

1) learning TTP associations;
2) develop a dataset of extended threat playbooks;
3) seed the algorithm with ground truth TTPs observed during the attack;
4) overlay known TTP associations to identify possible unobserved TTPs;
5) use a weighted Jaccard similarity to predict likely adversaries (attribution).

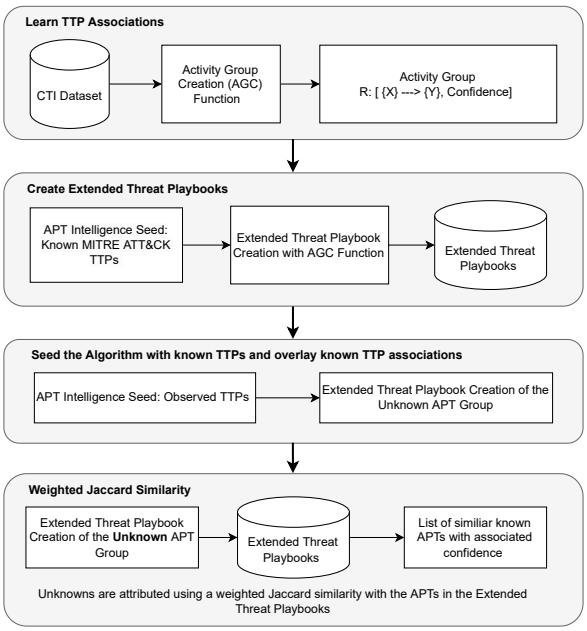

Fig. 2: Overview of approach to APT Attribution.

A. Learning TTP Associations

We adopt the approach proposed by Mckee et al. to learn the associations between TTPs [19]. This technique utilizes association rule mining with the aid of the apriori algorithm to learn the statistical relationships between TTPs. A relationship or grouping of statistically related TTPs is known as an Activity Group (AG). Here, let T be the set of all known ATT&CK TTPs. An AG is formally defined as the association between two sets of techniques, $AG : X \rightarrow Y$, where $x, y \in T$. The activity group is a representation of the TTPs in Y occurring given the behaviors in X are observed. An example activity group is shown here:

$$AG : T1027, T1059 \rightarrow T1105$$

In this example, if T1027 and T1059 are observed, then T1105 is hypothesized to exist in the same attack campaign. This *hypothesized TTP* helps analysts hunt for unobserved TTPs and draw analytical conclusions around attack scenarios. The apriori algorithm generates activity groups by iteratively finding frequent item sets and then using them to generate the association rules, or activity groups. These groups can be used to anticipate the realization of otherwise unobserved TTPs based on observed behavior.

The confidence of an AG is defined in 1 and represents the probability of the set of techniques in Y being observed under the condition that all the techniques in X are observed. Threat playbooks were generated by querying the MITRE ATT&K Framework and then storing APT G-codes and their associated TTPs to be parsed for APT attribution.

$$Conf(AG : X \rightarrow Y) = \frac{P(X \cup Y)}{Sup(X)} \tag{1}$$

B. Generating Threat Playbooks

A threat playbook is the set of TTPs that an APT group regularly uses. To generate standard threat playbooks, we query the MITRE ATT&CK framework to produce a set of all known APT groups and their respective TTPs. We associate a confidence level for each TTP in a playbook. For standard threat playbooks, all confidence levels are 1 as these are known and observed TTPs.

To generate *extended threat playbooks*, we use the activity-attack graph creation function proposed by Mckee et al. and seed it with the known TTPs in an APTs standard playbook [19]. We add each hypothesized TTP from the function to the standard playbook to form the extended playbook. Fig. 1 displays an excerpt from the extended playbook generated for the Ajax Security Team (G0130). The playbook includes both their known associated TTPs and *hypothesized* TTPs. A confidence level of 1 is associated with known TTPs and hypothesized TTPs are associated with their respective confidence levels generated by the activity-attack graphs.

C. Seeding the Algorithm

In this approach, we generate extended threat playbooks to provide a systematic and automated approach to identifying potential APTs that may be operating within an organization's threat space. To achieve this, our algorithm requires an initial intelligence seed, which is a set of observed TTPs associated with a known or suspected APT. Formally, the intelligence seed is represented by a set S that contains at least one TTP. The set S is a subset of the set T of all possible TTPs, but it is not empty. That is, $S = \{s : s \in T\}$, where $S \subseteq T$ and $S \neq \emptyset$.

The intelligence seed serves as the starting point for the algorithm to generate hypothesized techniques and relationships based on the techniques of interest to the analyst. For example, suppose an incident responder discovered a TTP used by an "unknown" APT. That TTP is used to seed the process that identifies other potential TTPs that have a high degree of co-occurrence.

Using the intelligence seed, our algorithm generates a list of potential APTs that frequently utilize the same TTPs that were produced. Next, these TTPs are compared with TTPs listed in already crafted threat playbooks. This comparison helps identify any matches between the TTPs generated by the intelligence seed and those listed in the existing threat playbooks. Based on these matches, a ranked list of APTs is generated. This ranked list indicates which APTs are most likely to be the "unknown" APT operating within the organization's threat space. This approach allows analysts to quickly and effectively identify potential APTs operating within their threat space and allows them to focus their investigation efforts on the most likely APTs and develop targeted responses to mitigate the threat.

D. Similarity Metric for Attribution

When comparing TTPs to the APT playbook, the Jaccard index [20] is a commonly used similarity metric. It measures

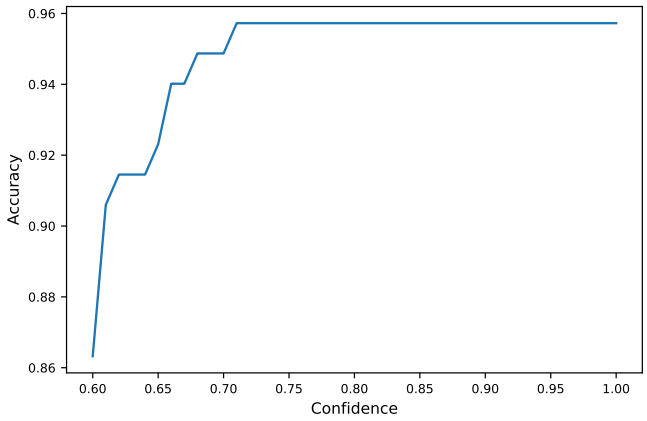

Fig. 3: Elbow plot of minimum confidence threshold for AAGs versus accuracy in identifying APT groups.

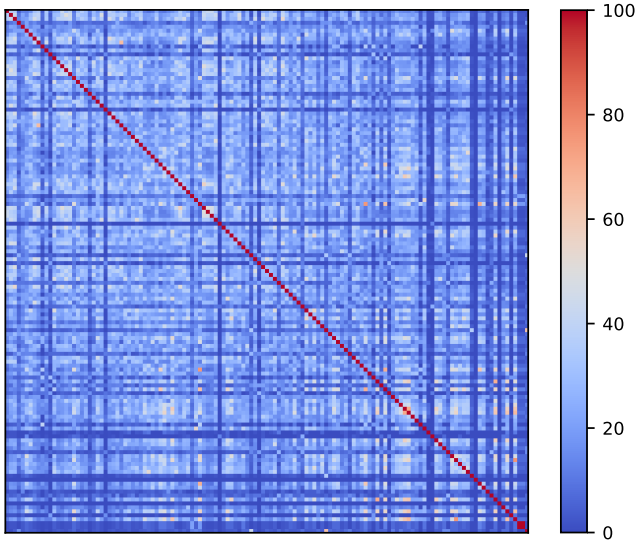

Fig. 4: Heatmap of similarities between APT groups. Each axis contains APT groups and colors indicate their degree of similarity.

the similarity between two sets by comparing the size of their intersection to the size of their union, as shown in 2.

$$Jaccard(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

Although useful for set similarity where all elements are equally weighted, the standard Jaccard is not an appropriate similarity metric for our research due to the varying confidence of different TTPs. To address this issue, a weighted Jaccard index [21] is used to account for the confidence value of each TTP when calculating the similarity between the sets. We assign weights to each TTP based on their relative confidence values. The weighted Jaccard index is calculated by multiplying the weight of each TTP by the number of times it appears in the intersection and then dividing by the sum of the weights of all TTPs in the union, as shown in 3.

$$Jaccard_{weighted}(A, B) = \frac{\sum_i \min(A_i w_{Ai}, B_i w_{Bi})}{\sum_i \max(A_i w_{Ai}, B_i w_{Bi})} \tag{3}$$

Our research focuses on the use of a weighted Jaccard to produce a similarity metric that we use to compare threat playbooks. However, a weighted Jaccard cannot capture the relationship between techniques and their associated sub-techniques. For example, T1087 (Account Discovery) and T1087.001 (Local Account Discovery) are closely related but would have a weighted Jaccard index of 0. As such, it is important to note that our research relies on abstracting sub-techniques into their parent techniques.

IV. Evaluation and Results

In this section, we describe our experimental design and present our results.

A. Experiment Design

We evaluate our approach to attribution using accuracy, which measures the proportion of attacks that are correctly attributed out of the total number of attacks. Specifically, we

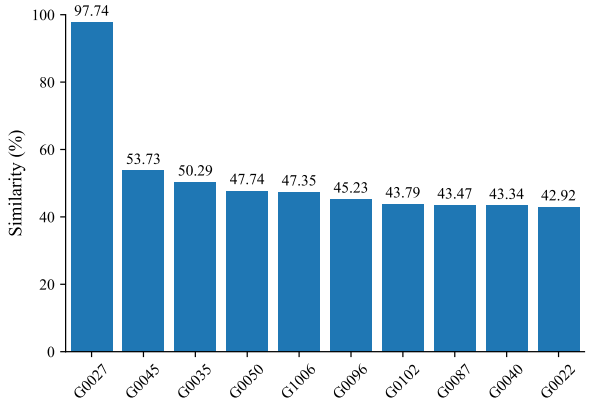

Fig. 5: Top 10 similar APT groups to an observation in the test dataset.

use micro-averaging, which aggregates the counts of correct classifications across all instances. This treats each prediction and APT as equally important when measuring the accuracy of the entire system.

We split the dataset into a training set and a testing set. we reserve 117 observations for testing, while the remaining 1199 observations are used to create the activity groups (association rule mining) and extended threat playbooks. Each observation in the test set includes ground truth threat groups and the associated TTPs they deployed during a given attack.

To prepare the test dataset, each observation is run on our Activity Attack Graph generation function to generate a testing playbook that includes seeded and hypothesized TTPs. Next, for each testing playbook, we use a weighted Jaccard index to calculate the most similar APT group using our threat playbooks and compare it to the APT group recorded for the

TABLE I: Table of similarity evaluation results. Accuracy is the number of correctly identified observations in the test dataset divided by the total number of observations. Similarity is the metric used to compare threat playbooks.

Playbook Type	Accuracy	Accuracy Std Dev	Avg Similarity	Similarity Std Dev
Standard	0.957	0.203	0.877	0.121
Extended	0.957	0.203	0.988	0.043

testing playbook.

B. Results

When constructing activity groups and extended threat playbooks, a confidence threshold must be selected as discussed in Section III-A. Using our accuracy metric, we were able to successfully determine that a minimum confidence threshold of 0.71 using the elbow plot in Fig.3. The plot levels out at a confidence threshold of 0.71, indicating that the best confidence to use to achieve maximum accuracy when classifying APT groups is 0.71.

Furthermore, the similarity heatmap shown in Fig. 4 provides insight into the uniqueness of each extended threat playbook. The heatmap depicts the level of similarity between the extended threat playbooks and indicates that there is not much overlap between them. This finding suggests that similarity is a useful metric when matching playbooks to our extended threat playbooks. It also highlights the importance of incorporating a large dataset when constructing extended threat playbooks, as it allows for the identification of unique and specific TTPs used by APT groups.

As seen in table I, our weighted Jaccard index correctly identified 95.7% of the APT groups in the test dataset when comparing observations from the test dataset against both standard and extended playbooks. While using both standard and extended playbooks achieved the same results, the extended playbooks provided much more confident identifications. Using extended playbooks yielded an average similarity level of 0.988 for correct classifications while using standard playbooks for evaluation resulted in an average similarity level of 0.877. These results indicate extended playbooks provide important insight into how APT groups operate and also help to improve distinguishability between APT groups.

An example of an observation in our test dataset being compared with extended threat playbooks is provided in Fig. 5. This particular observation was a set of recorded TTPs from a cyber attack performed by Threat Group-3390. In this case, G0027 (Threat Group-3390) is identified as having a similarity score of 97.74%. After G0027, there is a drastic reduction in similarity scores, indicating that our method of identifying possible APT groups using a similarity metric is valuable.

V. Discussion

In this section, we discuss the benefits of visualization, discuss limitations of our approach, and propose future work.

A. On Visualization

Extended threat playbooks enable analysts to better determine what adversaries they are dealing with, identify a list of TTPs to defend against, and compare APT groups with each other to identify similarities. In Fig. 6, the extended threat playbook for the Lazarus Group is plotted with all of the group's known and hypothesized TTPs. An incident responder that has identified this group as a potential threat can use this playbook to prepare to defend against TTPs that the Lazarus Group employs.

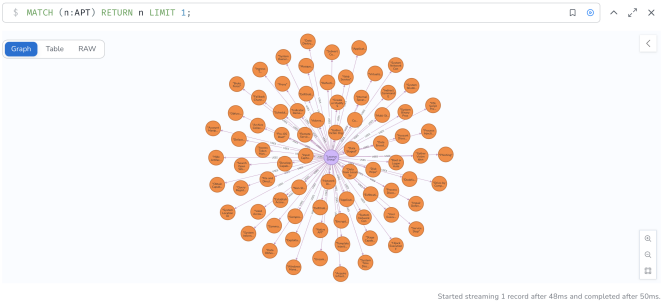

Fig. 6: Neo4j output showing an APT named "Lazarus Group" and associated TTPs.

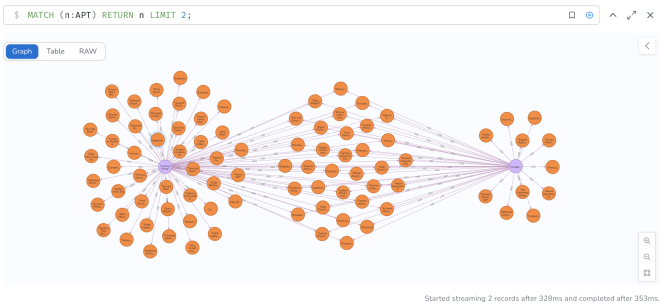

Fig. 7: Neo4j output showing the APTs "Lazarus Group" and "MuddyWater" and their associated TTPs as well the TTPs that they have in common.

In addition to incident response, extended threat playbooks can also be employed in threat hunting. In Fig. 7, the Lazarus group compared against MuddyWater and common TTPs can be identified. While this example provides two known APT groups, our weighted Jaccard similarity function can be employed to help identify a list of APT groups to investigate given a set of observed TTPs. Through visualization, a threat hunter or incident responder can study the similarities and differences between threat playbooks.

B. Limitations

Although our proposed approach for cyber attack attribution using association rule mining and a weighted Jaccard similarity index is promising, it has its limitations. One limitation is the availability of highly accurate and complete CTI data. Without accurate and current information on APT groups and

their associated TTPs, our approach may not be the most effective method to identify and assess potential cyber threats in real time. Another limitation is Concept Drift, which refers to the possibility that specific APTs could start to use different TTPs over time. As a result, the extended threat playbooks we create may become obsolete or less effective over time, requiring frequent updates and maintenance. To address these limitations, future work should focus on acquiring a larger and more diverse dataset to create more accurate relationships and extended threat playbooks. Additionally, we need to continuously monitor and update our models with the latest CTI data to ensure they remain effective and reliable.

C. Future Work

Our research can be extended to incorporate mitigations into the analysis process to help organizations prioritize defensive efforts against APTs. By mapping mitigations to specific techniques and TTPs, organizations can proactively defend against attacks and improve their overall cybersecurity posture.

Victimology analysis can be conducted to identify patterns and trends in APT attacks across different industries and countries, providing helpful information for analysts to prevent or recover from an attack. Further, including industry and country-specific threat intelligence may allow for a more comprehensive understanding of the threat landscape. Incorporating all features of the diamond model, which include motives, infrastructure, and capabilities, could be incorporated to provide a holistic view of the cyber threat landscape. This approach would involve examining political, economic, social, and technological factors, providing deeper insights into the behavior and tactics of APT groups. Ultimately, integrating these strategies would improve the accuracy and effectiveness of the proposed approach, enabling organizations to identify and mitigate cyber threats proactively.

VI. Conclusion

This paper proposes a novel approach to cyber attack attribution utilizing association rule mining and a weighted Jaccard similarity with an accuracy of 0.957 on our dataset. When coupled with STIX and visualization, our approach enhances situational awareness on threat actors and their TTPs. Further, we offer a dataset of APT threat playbooks that incorporate statistical analysis of TTP associations. We analyzed each playbook for uniqueness, showing that the weighted Jaccard similarity is a valuable metric for attributing attacks. For future work, we recommend acquiring more data to enhance our TTP activity groups and extended threat playbooks. Additionally, we need to continuously monitor and update our models with the latest CTI data to ensure their effectiveness and reliability. Despite these limitations, our approach represents a significant step forward in the field of cyber attack attribution and has the potential to make significant contributions to the development of practical and efficient cyber defense strategies.

References

[1] M. Ahmed, S. Panda, C. Xenakis, and E. Panaousis, "Mitre att&ck-driven cyber risk assessment," in *Proceedings of the 17th International Conference on Availability, Reliability and Security*, 2022, pp. 1–10.

[2] STIX. OASIS. [Online]. Available: https://oasis-open.github.io/cti-documentation/

[3] V. Mavroeidis and S. Bromander, "Cyber threat intelligence model: An evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence," in *2017 European Intelligence and Security Informatics Conference (EISIC)*, 2017, pp. 91–98.

[4] E. M. Hutchins, M. J. Cloppert, R. M. Amin *et al.*, "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains," *Leading Issues in Information Warfare & Security Research*, 2011.

[5] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, "Mitre att&ck: Design and philosophy," in *Technical report*. The MITRE Corporation, 2018.

[6] F. Maymí, R. Bixler, R. Jones, and S. Lathrop, "Towards a definition of cyberspace tactics, techniques and procedures," in *2017 IEEE International Conference on Big Data (Big Data)*. IEEE, 2017, pp. 4674–4679.

[7] P. Rajesh, M. Alam, M. Tahernezhadi, A. Monika, and G. Chanakya, "Analysis of cyber threat detection and emulation using mitre attack framework," in *2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*. IEEE, 2022, pp. 4–12.

[8] S. Caltagirone, A. Pendergast, and C. Betz, "The diamond model of intrusion analysis," Center For Cyber Intelligence Analysis and Threat Research Hanover MD, Tech. Rep., 2013.

[9] R. Howard and R. Olson, "Implementing intrusion kill chain strategies," *The Cyber Defense Review*, vol. 5, no. 3, pp. 59–76, 2020.

[10] L. Qiang, Y. Zeming, L. Baoxu, J. Zhengwei, and Y. Jian, "Framework of cyber attack attribution based on threat intelligence," in *Interoperability, Safety and Security in IoT: Second International Conference*. Springer, 2017, pp. 92–103.

[11] J. Ryu and J. Na, "Security requirement for cyber attack traceback," in *2008 Fourth International Conference on Networked Computing and Advanced Information Management*. IEEE, 2008, pp. 653–658.

[12] J. Hunker, B. Hutchinson, and J. Margulies, "Role and challenges for sufficient cyber-attack attribution," *Institute for Information Infrastructure Protection*, pp. 5–10, 2008.

[13] N. Tsagourias, "Cyber attacks, self-defence and the problem of attribution," *Journal of conflict and security law*, vol. 17, no. 2, pp. 229–244, 2012.

[14] F. Jaafar, F. Avellaneda, and E.-H. Alikacem, "Demystifying the cyber attribution: An exploratory study," in *2020 IEEE International Conference on Dependable, Autonomic and Secure Computing*. IEEE, 2020, pp. 35–40.

[15] N. Schultz and A. Duby, "Towards a framework for preprocessing analysis of adversarial windows malware," in *2022 10th International Symposium on Digital Forensics and Security (ISDFS)*. IEEE, 2022, pp. 1–6.

[16] A. Warikoo, "The triangle model for cyber threat attribution," *Journal of Cyber Security Technology*, vol. 5, no. 3-4, pp. 191–208, 2021.

[17] G. Sharkov, "From cybersecurity to collaborative resiliency," in *Proceedings of the 2016 ACM workshop on automated decision making for active cyber defense*, 2016, pp. 3–9.

[18] S. Chenette, "Emulating attacker activities and the pyramid of pain," https://www.attackiq.com/2019/06/26/emulating-attacker-activities-and-the-pyramid-of-pain/, Jun. 2019, accessed: 2022-12-9.

[19] C. Mckee, K. Edie, and A. Duby, "Activity-attack graphs for intelligence-informed threat coa development," in *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*. IEEE, 2023.

[20] P. Jaccard, "The distribution of the flora in the alpine zone. 1," *New phytologist*, vol. 11, no. 2, pp. 37–50, 1912.

[21] S. Ioffe, "Improved consistent sampling, weighted minhash and l1 sketching," in *2010 IEEE international conference on data mining*. IEEE, 2010, pp. 246–255.