Identification of Potential Threats in the Critical Path for Defense Operations by Cross Reference Features Clustering

Antonio Horta[1,3], **Renato Marinho**[1,2], **Raimir Holanda**[1,2]

[1]Morphus labs, Morphus Segurança da Informação
R. Carolina Sucupira, 1368 - Aldeota, Fortaleza - CE, Brazil

[2]University of Fortaleza
Av. Washington Soares, 1321 - Edson Queiroz, Fortaleza - CE, Brazil

[3]Military Institute of Engineering - IME
Praça Gen. Tibúrcio, 80 - Urca, Rio de Janeiro - RJ, Brazil

antonio@horta.net.br, rmarinho@morphus.com.br, raimir@unifor.br

Abstract. *Cyber attacks are a threat to the security of the most diverse types of organizations. To mitigate the risk of suffering successful attacks, organizations use different types of assessments. The research problem addressed in this study is to present, among the behaviors of known threats, those that are similar to the assessment campaign carried out and consequently represent greater risk when attempting attacks using the more exploitable critical path. The purpose of this research is to present a method for identifying the critical path of the threat and the similarity factor by Cross Reference Features (CRF) method to identify the opponents most similar to the procedures used in the assessment campaign carried out. The CRF was used as a baseline for comparison between 2 unsupervised learning algorithms in the threat clustering task. For essays that considered only groups as threats, K-means outperformed the Hierarchical Agglomerative Clustering by 2.4 percentage points, while in the essay with all threats, Hierarchical Agglomerative Clustering surpassed K-means by 2.3%.*

1. Introduction

Nowadays, cyber attacks are a threat to the security of the most diverse types of organizations. To mitigate the risk of suffering successful attacks, organizations use different types of assessments in search of a set of practices and techniques that form the most likely critical path to be successfully exploited by hackers and malware groups in achieving their goals [Atiku et al. 2020].

Regardless of the type of assessment carried out, such as war games [Moumouh et al. 2023], penetration testing [Mayukha and Vadivel 2023] or breach attack simulations [Jaber and Fritsch 2023], the results of the execution of your campaigns are presented in reports that will serve as a basis for preparing the action plan to mitigate the weaknesses found. As assessment campaigns perform various procedures that simulate the behavior of known hackers or malware groups in a generalized way. The research problem addressed in this study is to present, among the behaviors of known threats, those that are similar to the assessment campaign carried out and consequently represent greater risk when attempting attacks using the more exploitable critical path. Knowing the threats similar to the campaigns carried out, it is possible to convey to organizations in a clearer

and more objective way, those to which they are vulnerable and help in the elaboration of a mitigation plan that is more adequate to their needs.

Most research found in the state of the art has the objective of detecting and recognizing patterns in combating malicious adversaries [Ahn et al. 2022, Lin et al. 2022, Noor et al. 2019, Shin et al. 2022]. However, these works have addressed this problem through proposals for a cyber threat attribution using unstructured reports in cyber threat intelligence [Irshad and Siddiqui 2022] and an automated reclassification for threat actors [Shin et al. 2021]. Also, the study in [Park et al. 2023] compares the TTPs of different Advanced Persistent Threat (APT) groups through vectorizing the ATT&CK matrix and calculating the cosine similarity.

The purpose of this research is to present a method for identifying the similarity between the procedures used into a specific threat campaign carried out by security analysts and the procedures adopted by attacking groups. A similarity factor will be used as a baseline for comparison between 2 unsupervised learning algorithms in the threat clustering task. For this comparison, 4 experiments were carried out that used the results of the OilRig attack simulation exercise. These results are publicly available by MEAE[1] (MITRE Engenuity ATT&CK Evaluations) to present the results achieved on the similarity of threats through the Cross Reference Features (CRF) method proposed here, Hierarchical Agglomerative Clustering (HAC) and K-means. Although K-means is a robust and widely used method, it was necessary to change some parameters to obtain better results throughout the experiments, unlike CRF and HAC, which maintained similar results with the initial parameters used. Considering the CRF as a baseline for the similarity factor, for the tests that consider only groups as threats, K-means outperformed the HAC by 2.4 percentage points, while in the test with all threats, that is, groups, malware and tools, HAC surpassed K-means by 2.3%

Finally, this article is structured starting with this introduction, then a fundamentals section, which will present the concepts necessary to understand the research, related works, the proposed method for clustering threats, the experiments with their results and the conclusions.

2. Fundamentals

In this section, an outline of the concepts necessary for understanding this research and the experiments carried out will be presented. The section is organized into 2 fundamental topics: kill-chains, related to offensive security and the other with machine learning, more specifically with unsupervised learning.

2.1. Kill-Chains

According Kim [Kim et al. 2019], the term kill-chain comes from the military area to define all stages of an attack until reaching its objectives. Therefore, cyber kill-chain models intend to explain adversary behavior to carry out APT attacks. APT is a concept for an attack carried out in several steps with well-defined objectives, usually found in malware and hacker groups.

The cyber kill-chain (CKC) concept was coined by Lockheed Martin [Martin 2014] in 2014 who introduced a 7-step kill-chain. Later, in 2017, Van Den

[1]https://mitre-engenuity.org/cybersecurity/attack-evaluations

Berg presented the *unified kill-chain* [Van Den Berg 2017], to resolve some issues not addressed in the original CKC and currently there are proposals to make kill-chains more dynamic, such as *polymer kill-chain* [Neto et al. 2021]. In parallel, MITRE has been organizing since 2018 the techniques used in these attacks, which gave rise to the ATT&CK [2] knowledge base.

MITRE ATT&CK is a public knowledge base, organized in adversary tactics, techniques, procedures and campaigns based on real-world observations. The ATT&CK knowledge base is used as a foundation for the studies in cyber-security area, development of specific threat models and methodologies in the all kinds of sectors, organizations and academic community. It is organized into 3 matrices: enterprise, describing tactical techniques and procedures (TTP) used by threats against organizations; mobile, which contains TTP for mobile systems and ICS focuses on industrial control system environments.

Leszczyna's [Leszczyna 2021] review about cyber-security assessment methods, featured 32 methods for evaluating organisations to help them improve their cyber-security. Among them, we can mention: penetration tests, war games and breach attack simulations. These assessments can be organized into campaigns that attempt to simulate the behavior of adversaries performing ethical hacking through APT attacks against organizations.

Usually, the result of these evaluations, depending on the executor, is a report containing the kill-chains carried out, with all the information from the MITRE ATT&CK TTP used during the campaigns.

2.2. Machine Learning

According to Alloghani [Alloghani et al. 2020], machine learning can used in different problem-solving paradigms through supervised and unsupervised learning. Supervised learning consists of a training phase and a test phase to assess whether the results obtained by the algorithm used are in accordance with the expected result. In this supervised learning, there is a set of known input data and output data that will be used in training and testing. After these phases, the trained model can predict results for regression, selection, classification problems based on what it has learned. On the other hand, unsupervised learning does not have a labeled data-set, so this type of machine learning is commonly used for problems where there is no answer checking in learning, leaving it to the algorithm, according to its fundamentals, to carry out the clustering process and estimate an outcome for a given problem.

Based on the characteristics of the problem studied in this research, unsupervised learning was considered the most suitable. We used this approach for the development of the proposed method and for the proof of concept experiments. Therefore, we focused on learning the structure of data without making use of labels.

The commonly used unsupervised learning technique is cluster analysis, which is massively utilized for exploratory data analysis to determine the hidden patterns and to group the data. According to Kinge [Kinge et al. 2023], there are many unsupervised learning algorithms for clustering, such as: Hierarchical, K-means, DBSCAN and Gaus-

[2] https://attack.mitre.org

sian Mixture Model Clustering. Among these, hierarchical clustering and K-means are very popular.

3. Related Works

Several studies have used machine learning techniques to detect and recognize attack patterns. However, the vast majority of the studies are focused on threat detection and on the offensive point of view, which use binary analysis logs for malware recognition or network dumps for attack detection as data-sets.

The main aim of this study from [Ahn et al. 2022] was to perform a dynamic analysis of malicious and suspicious files and apply the results to the MITRE ATT&CK framework to visually present the attack tactics and detailed techniques used for the files. The proposal was to make it easier for agents to identify and respond to threats. In addition, it utilizes the advantages of dynamic-analysis based malicious file detection to increase detection accuracy.

In [Lin et al. 2022], the authors propose a mechanism for labeling attack tactics of network intrusion detection system (NIDS) rules on the basis of text mining and machine learning. The proposed approach can help to determine a current attack state and infer its purpose, making it possible to detect complex attacks (e.g., Advanced Persistent Threat). Besides, the authors refer to the ATT&CK framework to strengthen the reliability of labeling results. The experiment result shows that the accuracy of the proposed mechanism can effectively boost the performance of the labeling attack tactic.

The paper proposed by [Noor et al. 2019] presents a framework to automate cyber threat attribution. Specifically, the authors profile cyber threat actors (CTAs) based on their attack patterns extracted from cyber threat intelligence (CTI) reports, using the distributional semantics technique of Natural Language Processing. Using these profiles, they train and test five machine learning classifiers on 327 CTI reports collected from publicly available incident reports. Findings from the five machine learning models evaluated in this paper suggested that the Deep Learning Neural Network (DLNN) model is more effective than the remaining four models. They also compared the effectiveness of the high-level IOC profiled CTA dataset with the dataset provided by ATT&CK MITRE. The findings demonstrated that the machine learning models trained with the proposed dataset attribute cyber threats with high precision, recall, f-measure, and a low FPR, as compared to the ATT&CK dataset.

In [Shin et al. 2022], the study proposes a method to select the most effective strategy for responding to threats by analyzing the similarity of TTPs (Tactics, Techniques, and Procedures) of cyber campaigns. The similarity analysis method presented in this paper is a practical approach to cyber threat analysis that can effectively respond to cyber threats and reclassify the names of threat actors. In order to show that this study can effectively suggest countermeasures for cyber campaigns, the authors collected phishing incidents by three country-based threat groups, classified them into 16 campaigns, and expressed them as ATT&CK matrix. When the similarity of each tactic of the phishing campaigns expressed by the ATT&CK matrix was calculated, the tactic with a high average similarity could identify the stage where mitigation should be focused on during the phishing campaign through the approach that the technique used had little diversity.

Irshad and Siddiqui [Irshad and Siddiqui 2022] developed a mechanism to at-

tribute or profile cyber threat actors (CTA) by extracting features from cyber threat intelligence (CTI) reports. They defined a methodology to extract features from unstructured CTI reports by using natural language processing (NLP) techniques and then attributing cyber threat actor by using machine learning algorithms. Machine learning algorithms such as decision tree, random forest, support vector machine were used for classification of CTA.

Shin [Shin et al. 2021] proposes an Automated Reclassification for Threat Actors (ART) that quantitatively compares the TTPs (Tactics, Techniques, and Procedures) from different APT (Advanced Persistent Threat) groups. This work crawls cyber threat reports (CTA) and retrieves the ATT&CK matrix of APT groups. Then, it vectorizes the ATT&CK matrix and calculates the cosine similarity. By reexamining the various aliases of the CTAs with the ATT&CK framework, ART can help to classify the indiscriminately established APT groups.

Although the offensive point of view has been discussed in many works, the defensive one presents an important lack of research. Therefore, we have addressed, in this paper, an approach not largely analyzed: to prevent but, especially, to detect and to neutralize security threats. Specifically, our proposal differs from the works listed above by proposing a simplified method for calculating the similarity between a campaign and all types of threats, such as groups, malware and tools. Additionally, we present the critical path of the attack, which is the one that represents the greatest risk of being successfully exploited in the target environment.

4. Threat Similarity Clustering Process

In this section, the elements that make up the method used in this research to carry out the similarity grouping of potentially dangerous threats to the cyber-defense of a given organization will be explained.

As can be seen in Figure 1, the proposed method has as input a table with the results of kill-chain previously performed in the offensive security exercise. Afterwards, this results table is used to make another table that represents the offensive procedures, represented by crossing the groups, malware and tools with the techniques used by them. In procedures table, the techniques used in the campaign and the results of the operations of the cyber-security teams are inserted from results table. Lastly, grouping algorithms are applied in order to identify which threats are part of the same group as the campaign carried out. Furthermore, threats belonging to the same group as the campaign's kill-chain performed, as they are considered similar, represent a potential threats similar to the campaign carried out.

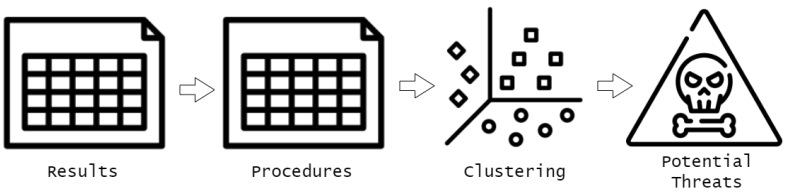


Figure 1. Threat clustering process used in this research.

4.1. Offensive Campaign Results Table

After an offensive security assessment campaign is conducted, its results need to be consolidated and tabulated to serve as input into the threat clustering process. The *Results Table* 1 is organized into columns that represent the weighted average of the metrics achieved by the defense layers for each time the campaign technique was detected or responded to, such as incident response activities performed by security operations center. Where the columns *tactic* and *technique* denote the MITRE ATT&CK tactic and technique respectively used; *udtc* represents undetected events; *dtc* is the number of times the event was detected, *total* is the sum of detected and undetected events; and *SVI* (Security Visibility Index) is the percentage of events detected by each technique. Finally, the total visibility index for the entire table can be calculated by weighting the columns *SVI* by *total*.

Table 1. Example of results table used as input.

tactic	technique	udtc	dtc	total	SVI
TA0002	T1033	0	16	16	1.0
TA0005	T1027	13	19	32	0.6
TA0005	T1036	2	30	32	0.9
...	...	...	...	...	...
TA0011	T1105	18	109	127	0.9
TA0011	T1572	1	15	16	0.9

4.1.1. Threat Critical Path

The SVI is an indicator that reflects the performance of the security operations centre in terms of detection and response to the techniques observed in the offensive campaign carried out to assess the environment through BAS, war-games, penetration or resilience testing exercises. As a kill-chain is the set of techniques used along the employed tactics, the threat critical path can be traced through the selection of the set of techniques with the minimum SVI values found in kill-chain performed. The critical path is represented by equation 1, where the Threat Critical Path (*TCP*) is the set of techniques (*T*) with minimal Security Visibility Index (*SVI*) for each tactic (τ).

$$TCP = \sum_{\tau} \left\{ \left\{ T_{min(SVI)} \right\} \in \tau \right\} \tag{1}$$

4.2. Threat Procedures Table

The table of threat procedures brings together all groups, malware and tools published in the latest version of ATT&CK available in a public repository[3] made available by MITRE itself, as well as the techniques used in the campaign and the actions of the security operations center extracted from the results table. The threat procedures table has each MITRE technique as columns and the name of the threats as index. As can be seen in the example

[3] https://github.com/mitre/cti

Table 2, each technique found in the kill-chain of a given threat is filled with the value 1 and 0 if it is not part of it. In addition to the MITRE threats, the *Campaign* row is inserted, which are the techniques that were part of the kill-chain of the offensive exercise performed and the row *OPS* also is inserted representing the SVI results of security operations center for each technique.

Table 2. Example of threat procedures table.

procedure	T1566.002	T1033	...
APT1	1.0	0.0	...
...	...	...	...
Campaign	1.0	1.0	...
OPS	0.0	0.7	...

4.3. The Clustering Process

The clustering process proposed in this research aims to identify, among all the threats cataloged in MITRE ATT&CK, those that contain in their kill-chains the same techniques used in the offensive security campaigns carried out. This clustering procedure considers the campaign techniques carried out as a reference kill-chain with 100% similarity. Therefore, threats with a greater intersection of techniques are considered more similar than those with less intersection.

4.3.1. Plotting Requirements

As the proposed process consists of the intersection of threat features with the features of the reference campaign through the intersection of 594 techniques present in the threat procedure table. Therefore, it is necessary to reduce the dimensions so that a scatter plot can be plotted.

According Ji [Ji et al. 2023], PCA (Principal Component Analysis) is useful to reduce dimensions. Therefore, it was used to reduce the threat procedures table and allowed the kill-chains of each threat, campaign and security operations centre to be plotted in a scatter with two dimensions. Due to this fact, the first and second components were used in this research exclusively to plot the scatter and were not used in the proposed clustering procedure or in the performed experiments.

4.3.2. The Range of Threat Clusters

The proposed cross reference features clustering process requires that the maximum number of clusters be informed. This maximum number of clusters represents the percentage divisor for establishing the ranges in which threat similarity factors will fit.

In contrast other methods, such as HAC and KMEANS, which group the sample into the informed number of clusters, in the method proposed here, the maximum number of clusters indicates that the sample will be grouped into up to the maximum number of clusters specified. Therefore, the greater the entropy of the threat similarity factors, the closer to the maximum number of clusters the sample will be clustered.

4.3.3. Cross Reference Features

The cross reference features clustering process consists of generating a reference table that contains only the features (columns of techniques) found in the reference campaign extracted from procedures table without the security operations centre row (OPS). Therefore, in this reference table, the total column is obtained from the sum of the columns of each threat line and the similarity factor is obtained by the quotient of dividing the total by the reference campaign total. In this way, the quotient closer to 1 indicates a greater similarity with the techniques used in the reference campaign.

When the individual similarity factors are calculated, the *range of threat clusters* is applied according to the number of clusters defined, each threat will be grouped according to its similarity factor.

Table 3 is an example of cross reference features clustering, in which the *Threat* column refers to the threat name; *T1033, ..., T1204* columns are a subset of MITRE ATT&CK techniques filtered by crossing the reference campaign with the other threats; *T* denotes the total, obtained by sum of the technique columns; *S* is the similarity factor; and *L* are the labels assigned after the clustering process according to the calculated range based on 10 as maximum cluster number.

Table 3. Example of cross reference features table.

Threat	T1033	...	T1204	T	S	L
APT37	1	...	1	3	0.7	7
OilRig	1	...	1	4	1.0	9
Campaing	1	...	1	4	1.0	9
...	...	...	...	...	...	...

As can be seen in the example of Table 3, the threat *OilRig* was labeled (*L*) as *9*, the same group label as the referenced campaign, a consequence of the high similarity factor *S*. The other threats were grouped in clusters with lower similarity factors.

5. Experiments and Results

The experiment carried out aims to serve as a proof of concept of the proposed method, in the identification of the critical path for the offensive campaign carried out, as well as to present a baseline to reference potential threats that, by similarity, represent a similar risk to the offensive exercise of evaluation of organizations' security operations centers.

The MITRE Engenuity ATT&CK® Evaluations program brings together cybersecurity solutions providers with MITRE experts to evaluate an organization's capabilities. Each evaluation follows a systematic methodology using a threat-informed purple teaming approach to capture critical context around a solution's ability to detect or protect against known adversary behavior as defined by the ATT&CK knowledge base. Results from each evaluation are thoroughly documented and openly published. The evaluations are measurable and repeatable, making them useful for continual assessments of incremental improvements.

Therefore, the MEAE was used as database for the offensive evaluation exercises, as well as a comparison of the results achieved by other clustering methods, such as HAC

and KMEANS. MEAE is a program that performed independent ATT&CK Evaluations for 16 managed security service providers in their ability to analyze and describe adversary behavior through emulation the tactics and techniques of OilRig[4].

Firstly, a web scraping process was carried out on the MEAE campaign website[5] to generate a table of results in the format shown in Table 1. Then, the total SVI were calculated for each of the providers, which represents, according to the previously explained methodology, the result of the security operations of each of the participants in relation to the campaign carried out by MEAE. The bar chart in Figure 2 presents the calculated SVI for each of the participants and reflects the result of the response and detection operations for the OilRig campaign carried out by the MEAE. Furthermore, considering the SVI of each participant for the OilRig campaign, the average SVI achieved by the group was 78.41%.

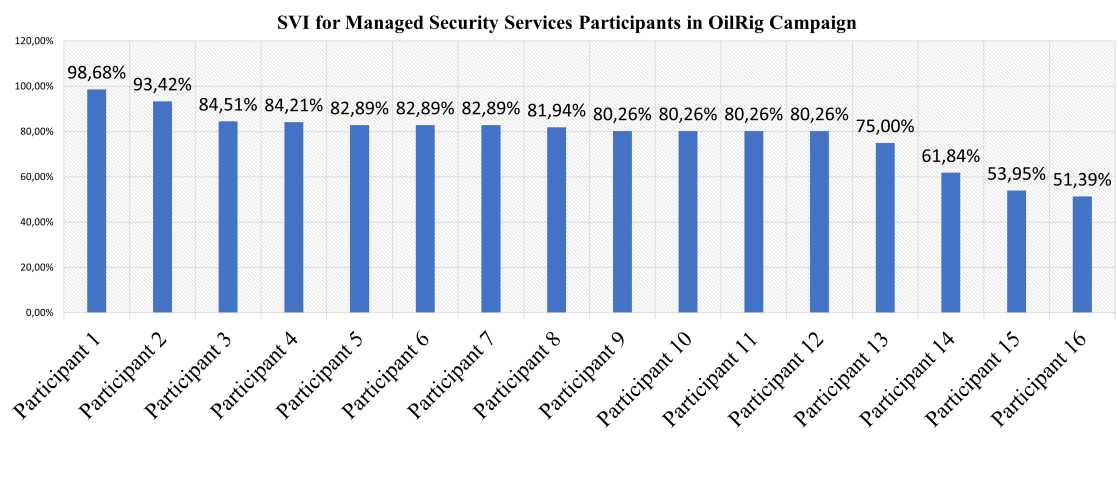


Figure 2. Security Visibility Index for MEAE participants.

Analogous to calculating the average SVI, it is also possible to determine the critical path of the campaign for all participants by following the steps presented in section 4.1.1. Therefore, the threat critical path calculated indicates as the main points of attention, with the lowest visibility indices by the group's average, the respective tactics and techniques: T1573.001) Command and Control - Encrypted Channel: Symmetric Cryptography with 19 %; T1083) Discovery - File and Directory Discovery with 29%; T1041) Exfiltration - Exfiltration Over C2 Channel with 32% and; T1497.001) Defense Evasion - Virtualization/Sandbox Evasion: System Checks with 50%.

In addition to the critical path presented, still based on the table of procedures, it is possible to apply clustering algorithms to identify threats that use techniques similar to those used in kill-chains of other threats.

For this experiment, the paradigm of unsupervised machine learning was chosen, because there are no previous labels for the reference campaign, which makes prior training by supervised learning algorithms complex. Considering that the objective of this experiment is to identify the threats that use the largest number of techniques equal to

[4]https://attack.mitre.org/groups/G0049
[5]https://attackevals.mitre-engenuity.org/managed-services/oilrig/

those used in the reference campaign kill-chain. The *cross reference features*, previously described in section 4.3.3, is a baseline to achieve this objective, by determining a *similarity factor* that can be ordered or applied to some classification range, serving as a simple technique for comparing the elements that form the clusters generated by different algorithms.

In this experiment, 3 clustering techniques were applied to the procedure table, the Hierarchical Agglomerative, KMEANS and the Cross Reference Features with the input parameters according to Table 4. Table 4 is organized into 6 columns, where: *Exp.* represents an identifier of the essays in the experiment; *Alg.* are the algorithms used, which can be *CRF* (Cross Reference Features), *HAC* (Hierarchical Agglomerative Clustering) or *KMNS* (K-means); *N* denotes the number of clusters used; *Linkage* represents which binding criteria was used; *Feat.* indicates whether *all* or *ref.* as a subset of features from the reference campaign performed and; *Sample* denotes the scope of the sample used, which can be only *groups* or *all* for groups, malware and tools.

Table 4. Initial parameters used in clustering essays.

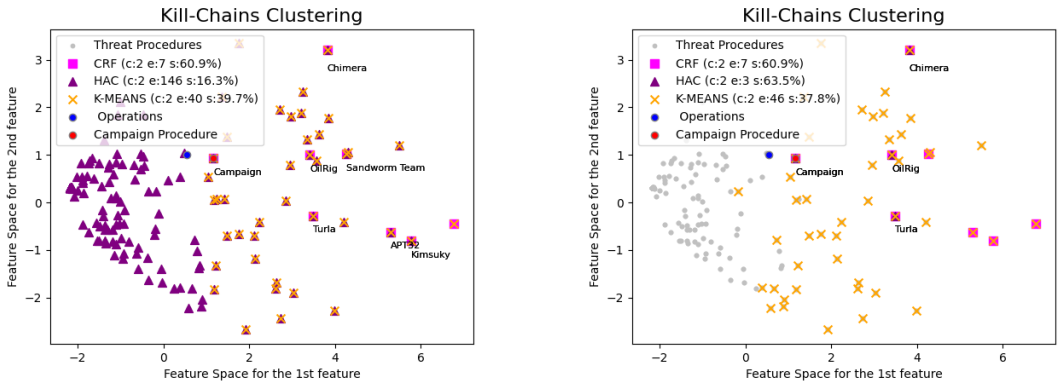
Alg.	Exp.	N	Linkage	Feat.	Sample	Exp.	N	Linkage	Feat.	Sample
CRF	E1	2	$> 50\%$	ref.	groups	E3	2	$> 50\%$	ref.	groups
HAC		2	average	all	groups		2	average	ref.	groups
KMNS		2	lloyd	all	groups		7	lloyd	ref.	groups
CRF	E2	2	$> 50\%$	ref.	groups	E4	2	$>50\%$	ref.	all
HAC		2	average	ref.	groups		2	average	ref.	all
KMNS		2	lloyd	ref.	groups		17	lloyd	ref.	all

According to the initial setup presented, the experiments were performed in order to compare the similarity of the elements grouped by the 3 algorithms with different parameters. Each experiment generated a scatter, which indicates an average kill-chain of the security operations of all participants, one that represents the MEAE OilRig campaign and, finally, the kill-chains of similar threats grouped by similarity by algorithms. In the legends of the scatters below, in the *Similar* items, the letter *c* indicates the number of clusters, *e* the total of similar elements in the cluster and *s* denotes the similarity factor calculated based on the section 4.3 for the 3 algorithms.

In the first test, E1, only groups were considered as a sample (Figure 3.a), the baseline CRF formed a cluster with 7 similar elements with an average similarity factor of 60.9%, while HAC grouped 146 with a similarity factor of 16.3% and K-means grouped 40 with a similarity factor of 39.7%. MITRE ATT&CK version 12 has 853 cataloged threats, between malicious groups and pieces of software, therefore, the 146 and 40 elements grouped by HAC and K-means respectively indicate a high number of grouped elements when compared to only 7 threats selected by the baseline CRF, which considers in its selection a similarity index superior to 50% in the sampling. Furthermore, these elements grouped by HAC and K-means with a low similarity factor raise questions about what caused such a result and what can be done to achieve more satisfactory indices.

As the CRF considers only the reference features, the E2 experiment, used as input base for the HAC and K-means, only the features found in the reference campaign. Therefore, the objective of this second experiment was to verify if decreasing the num-

ber of input features, the HAC would reduce the number of elements in the cluster and increase the similarity factor. As can be seen in Figure 3.b, due to the number of input features reduced in E2 to the same as in the reference campaign used by CRF, the result was better than E1 for HAC and worst for K-means. In HAC, the number of elements dropped to 3 and the similarity factor rose to 63.5%. In contrast, K-means elements rose to 46 and similarity index dropped to 37.8%.



(a) E1 (b) E2

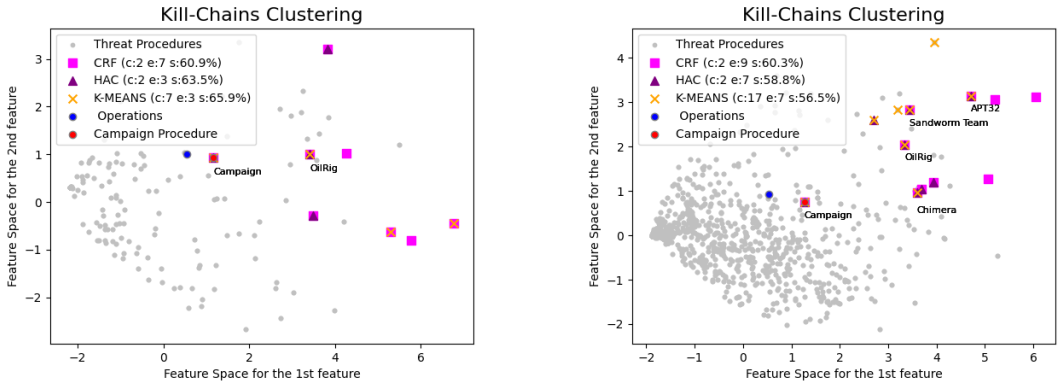
Figure 3. Clustering Experiments 1 and 2

In search of better results, the experiment E3 was carried out. E3 is similar to E2, changing only the number of clusters for K-means until find number of cluster that achieve the criteria: number of elements equal to elements in HAC and the higher similarity index.

As can be seen in Figure 4.a, the result for K-means was better than the HAC, where K-means found 3 elements with a similarity factor of 65.5%. Moreover, as HAC and K-means have an equal amount of elements, the higher similarity factor of K-means indicates that their elements are more similar than HAC, representing a better performance of K-means for selection of malicious groups in E3.

The last essay, E4 in Figure 4.b, was derived from E3, with the only difference been considered all threats, such as malware and tools besides only groups. Due to sample have been increased, CRF baseline become 9 elements and similarity index 60.3%. HAC and K-means clustered 7 elements and achieved 58.8% and 56.5% respectively for similarity index. Although K-means considers 17 the ideal number of elements, according to the criteria explained in E2, it still had an inferior performance in the selection of potential threats compared to HAC, which divided the sample into 2 groups, similar and non-similar.

Table 5 presents a comparative table of CRF, HAC and K-means (KMNS) in relation to the results obtained in E4. This table lists the name of the threats that each algorithm has grouped together. There are 4 threats that are common to CRF, HAC and K-means, and among them, the presence of OilRig is important to be present at this intersection, as the campaign carried out by MEAE simulated its behavio. Although the CRF has grouped more elements, its average similarity factor is higher than the HAC and K-means, this is due to the HAC and K-means having grouped threats such as the APT29, APT39 and Wizard Spider, which has a similarity factor lower than the others found.



(a) E3 (b) E4

Figure 4. Clustering Experiments 3 and 4

Table 5. E4 clustering results for MEAE OilRig Campaign.

Potential Threats	CRF	HAC	KMNS
OilRig	✓	✓	✓
APT32	✓	✓	✓
APT29	-	-	✓
APT39	-	-	✓
Sandworm Team	✓	✓	✓
Turla	✓	✓	-
Lazarus Group	✓	-	-
Wizard Spider	-	✓	✓
Chimera	✓	✓	✓
Kimsuky	✓	-	-
Cobalt Strike	✓	✓	-
QakBot	✓	-	-
Similarity Factor	60.3%	58.8%	56.5%
Elements in cluster	9	7	7
Nº Clusters	2	2	17

We conclude that the findings are consistent with the offensive campaign, demonstrating that the approach proposed in this work is capable of identifying which malicious campaigns present greater risks, given a certain kill chain.

Despite the existence of some works that apply machine learning algorithms to compare TTPs (Tactics, Techniques, and Procedures) from different APT (Advanced Persistent Threat) groups, the results presented here point out for a new Direction. We were capable to identify the most similar attacking group pattern to a offensive campaign carried out. This identification allows organizations to develop a set of countermeasures to defend against possible attacks.As far as we know we know, the results produced here are not available in the literature.

6. Conclusions

Like most research found in the state of the art, it aims to detect and recognize patterns in combating adversaries, although studies have been found to address this problem through proposals for a cyber threat attribution using unstructured reports in cyber threat intelligence and an Automated Reclassification for Threat Actors. The purpose of this research was to present a method for identifying the critical path of the threat and the cross reference features method for calculating the baseline similarity factor of adversaries most similar to the procedures used in the assessment campaign carried out.

This similarity factor was used as a baseline for comparison between 2 unsupervised learning algorithms in the threat clustering using the MEAE data-set. The 4 experiments detected OilRig as the main similarity, the behavior simulated by MEAE, demonstrating that the algorithms used are robust in the task of clustering large samples. The cross reference features (CRF) method proposed here served as a baseline for evaluating the level of similarity achieved by the clusters formed by HAC and K-means.

Although K-means is considered by several researchers as a superior performance method than the others, it was necessary to reduce the number of features and increase the number of clusters to obtain better results throughout the experiments, unlike CRF and HAC which maintained similar results with the initial parameters used. Considering the CRF as a baseline for the similarity factor, for the essays that consider only malicious groups, K-means outperformed the HAC by 2.4 percentage points, while in the essay with all threats, HAC surpassed K-means by 2.3%. The similarity factors achieved by both algorithms were determined by the average CRF in their respective clusters. Therefore, the similarity factor of 60.3% presented by CRF, indicates the average CRF of the 7 threats with CRF greater than 50%. Results above or below the average CRF need to be interpreted according to the number of elements grouped. Results with a similarity factor greater than 60.3% with less than 7 elements indicate that the cluster was formed with the most similar threats, while, regardless of the number of elements, similarity factors below 60.3% indicate a similarity lower than the baseline.

References

[Ahn et al. 2022] Ahn, G., Kim, K., Park, W., and Shin, D. (2022). Malicious file detection method using machine learning and interworking with mitre att&ck framework. *Applied Sciences*, 12(21):10761.

[Alloghani et al. 2020] Alloghani, M., Al-Jumeily, D., Mustafina, J., Hussain, A., and Aljaaf, A. J. (2020). A systematic review on supervised and unsupervised machine learning algorithms for data science. *Supervised and unsupervised learning for data science*, pages 3–21.

[Atiku et al. 2020] Atiku, S. B., Aaron, A. U., Job, G. K., Shittu, F., and Yakubu, I. Z. (2020). Survey on the applications of artificial intelligence in cyber security. *International Journal of Scientistic and Technology Research*, 9(10):165–170.

[Irshad and Siddiqui 2022] Irshad, E. and Siddiqui, A. B. (2022). Cyber threat attribution using unstructured reports in cyber threat intelligence. *Egyptian Informatics Journal*.

[Jaber and Fritsch 2023] Jaber, A. and Fritsch, L. (2023). Towards ai-powered cybersecurity attack modeling with simulation tools: Review of attack simulators. In *International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, pages 249–257. Springer.

[Ji et al. 2023] Ji, Y., Liu, H., Xiao, N.-C., and Zhan, H. (2023). An efficient method for time-dependent reliability problems with high-dimensional outputs based on adaptive dimension reduction strategy and surrogate model. *Engineering Structures*, 276:115393.

[Kim et al. 2019] Kim, H., Kwon, H., and Kim, K. K. (2019). Modified cyber kill chain model for multimedia service environments. *Multimedia Tools and Applications*, 78(3):3153–3170.

[Kinge et al. 2023] Kinge, A., Hrithik, P., Oswal, Y., and Kulkarni, N. (2023). Customer analytics research: Utilizing unsupervised machine learning techniques. In *Data Intelligence and Cognitive Informatics*, pages 501–515. Springer.

[Leszczyna 2021] Leszczyna, R. (2021). Review of cybersecurity assessment methods: Applicability perspective. *Computers & Security*, 108:102376.

[Lin et al. 2022] Lin, S.-X., Li, Z.-J., Chen, T.-Y., and Wu, D.-J. (2022). Attack tactic labeling for cyber threat hunting. In *2022 24th International Conference on Advanced Communication Technology (ICACT)*, pages 34–39. IEEE.

[Martin 2014] Martin, L. (2014). Gaining the advantage cyber kill chain. *https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/Gaining_the_Advantage_Cyber_Kill_Chain.pdf*. (Accessed on 01/03/2023).

[Mayukha and Vadivel 2023] Mayukha, S. and Vadivel, R. (2023). Reconnaissance for penetration testing using active scanning of mitre att&ck. In *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, pages 693–705. Springer.

[Moumouh et al. 2023] Moumouh, C., Chkouri, M. Y., and Fernández-Alemán, J. L. (2023). Cybersecurity awareness through serious games: A systematic literature review. In *International Conference on Networking, Intelligent Systems and Security*, pages 190–199. Springer.

[Neto et al. 2021] Neto, A. J. H., Dos Santos, A. F. P., and Dos Santos, M. (2021). Polymer: An adaptive kill chain expanding cyber threat hunting to multi-platform environments. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 2128–2135. IEEE.

[Noor et al. 2019] Noor, U., Anwar, Z., Amjad, T., and Choo, K.-K. R. (2019). A machine learning-based fintech cyber threat attribution framework using high-level indicators of compromise. *Future Generation Computer Systems*, 96:227–242.

[Park et al. 2023] Park, N.-E., Lee, Y.-R., Joo, S., Kim, S.-Y., Kim, S.-H., Park, J.-Y., Kim, S.-Y., and Lee, I.-G. (2023). Performance evaluation of a fast and efficient intrusion detection framework for advanced persistent threat-based cyberattacks. *Computers and Electrical Engineering*, 105:108548.

[Shin et al. 2021] Shin, Y., Kim, K., Lee, J. J., and Lee, K. (2021). Art: Automated reclassification for threat actors based on attck matrix similarity. In *2021 World Automation Congress (WAC)*, pages 15–20.

[Shin et al. 2022] Shin, Y., Kim, K., Lee, J. J., and Lee, K. (2022). Focusing on the weakest link: A similarity analysis on phishing campaigns based on the att&ck matrix. *Security and Communication Networks*, 2022.

[Van Den Berg 2017] Van Den Berg, J. (2017). The unified kill chain. *https://www.unifiedkillchain.com/assets/The-Unified-Kill-Chain-Thesis.pdf*. (Accessed on 01/03/2023).