Learning APT Chains from Cyber Threat Intelligence

Ghaith Husari, Ehab Al-Shaer, Bill Chu, and Ruhani Faiheem Rahman
University of North Carolina at Charlotte
Charlotte, NC, USA
{ghusari,ealshaer,billchu,rrahman}@uncc.edu

ABSTRACT

With the rapidly evolving cyber attacks, cybersecurity specialists are actively using cyber threat intelligence to identify and respond to cyber attacks in a timely manner. However, this information will be highly useful for attack detection and mitigation if we can construct structured cyber threat intelligence and accurately generate TTP Chains to understand the steps of cyber attacks. In this poster, we present our preliminary Natural Language Processing (NLP) analysis to characterize the temporal relationship of attack actions of an APT attack to extract and construct the reported TTP chains using the popular standard, MITRE ATT&CK [1], and the Structured Sharing Language, STIX 2 [2], a machine-readable language that will help automate the process of understanding and responding to the cyber attacks shared in unstructured text via blogs, emails, and social media.

1 BACKGROUND

With the increasing need for cyber threat intelligence, various standards have emerged to share cyber threat information with the public community. Most of these standards are based on the observation that cyber threats follow attack patterns. Identifying attack patterns is essential to define models to represent the attack information and for their mitigation planning. STIX [2] standard has become very popular for sharing CTI data. In STIX schema, `campaign` construct is defined as a set of Indicators of Compromise (IoCs) that belong to a given attack pattern. These IoCs consist of 88 types of malicious artifacts (e.g., website, IP address, registry key, etc.). To specify the sequential nature of attack pattern in a threat report, STIX has a `kill_chain_phase` attribute for the IoCs. This attribute is based on the definition of the cyber kill chain, which defines 7 phases for cyber attacks: reconnaissance, weaponization, delivery, exploitation, installation, command control and actions on objectives.

2 PROBLEM DEFINITION AND MOTIVATION

MITRE ATT&CK techniques provide threat high-level behaviors in terms of threat actions (e.g., Spearphishing Link, Exfiltration over C2, etc.). Although these threat behaviors are used to report cyber attack using the ATT&CK standard, yet there is no end-to-end approach that captures the sequential relationship between these actions and artifacts from threat reports. Although current solutions, such as TTPDrill [3], extract these techniques from unstructured texts of cyber threat intelligence, it does not capture the temporal relations report by the threat report. Recognizing this sequential (or temporal) relationship between actions and their artifacts is essential for detecting attacks by analyzing their artifacts gathered from the network. These artifacts can be used as signatures for detecting and preventing cyber attacks in a timely manner.

3 WORK OBJECTIVE AND TECHNICAL APPROACH

In this poster, we will present our preliminary work to capture the temporal relationship between threat artifacts that are reported by cyber threat intelligence. First, we design a set if Regular Expressions (regex) to capture IoCs (artifacts) from the text, then we utilize Natural Language Processing and design our NLP typed dependency ruleset to extract the temporal relationship between these artifacts in threat reports. Based on our careful analysis of threat IoCs and artifacts, We adapt five stages of kill chain that are associated with the threat artifacts that we capture using our regexes.

- `Deliver`: This is the first stage of delivery of the malware to provide initial access for the threat. This starts with a malicious email that includes a link or an attachment, if clicked downloads and executes a script that will start the malicious campaign.
- `Exploit`: This phase includes exploiting a software vulnerability in an installed application on the victim's machine.
- `Command and control`: Once installed, the malware will communicate with the remote c2 server via IP address or website to receive further commands on which malicious actions to execute.
- `Stage`: After collecting sensitive data from the user's machine, the malware usually stores this data by creating a file on the victim's system in preparation of exfiltrating it.
- `Exfiltrate`: The last phase of the attack is to send the collected information to a remote location using email, IP address, or web service.

Threat actions, in CTI reports, are often associated with temporal expressions. These expressions can be used to determine the temporal order between threat actions. The output and accuracy of this process are important for ordering threat actions of a given cyber-attack in one timeline. This ordering is essential for merging multiple reports of the same cyber threat into one aggregated report. Also, it enables IDS specialists to train models to capture malicious malware behaviors (i.e., sequence of ordered threat actions).

4 PRELIMINARY EXPERIMENTATION AND FINDINGS

In this preliminary work, We used the dependency rules shown in Table 1 to process 5,000 threat reports scraped from Symantec and Fireeye CTI blogs [4, 5] to capture the temporal relationship

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
HotSoS, April 1–3, 2019, Nashville, TN, USA
© 2019 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-7147-6/19/04.
https://doi.org/10.1145/3314058.3317728

Table 1: An example of temporal relations extraction among threat actions

threat description	The actors behind the previous FormBook campaign used CVE-2017-0199 ... Equation Editor downloaded a file called *"xyz[1].123"* and **then** created the *scvhost.exe* process... The malware comes ... with rootkit functions such as key-logging, clipboard stealing, screenshots and webcam access. Passwords are stolen from the following applications, among others ... it creates the POST request string, as you can see below... **Then**, it encrypts it with 3DES **before** sending it.
threat actions with temporal order	exploit vulnerability download file **[Before]** create process keylogging clipboard stealing screenshots access webcam access steal passwords from applications create POST request string **[Before]** encrypt it with 3DES encrypt it with 3DES **[Before]** sending it.

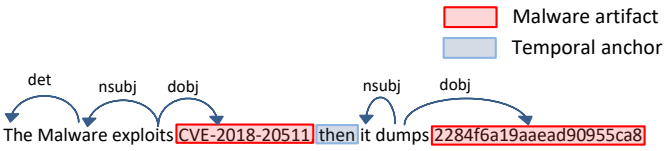


Figure 1: Example of Dependency Parsing, where the edge represents typed dependency type

Table 2: Excerpt of typed dependencies used in this work

typed dependency	Threat Action
nsubj	parent <verb>, child.<subj>
dobj	parent<verb>, child.<obj>
nsubjpass	parent<verb>, child.<obj>
nmod:agent	parent<verb>, child.<subj>
nmod:from	parent<verb>_from, child.<obj>
nmod:to	parent<verb>_to, child.<obj>
nmod:with	parent<verb>_with, child.<obj>
nmod:via	parent<verb>_via, child.<obj>
nmod:over	parent<verb>_over, child.<obj>
nmod:for	parent<verb>_for, child.<obj>
nmod:through	parent<verb>_through, child.<obj>
nmod:using	parent<verb>_using, child.<obj>
nmod:into	parent<verb>_into, child.<obj>


```
{
 "type": "malware",
 "id": "malware--0c7b-4a4d-aa9d-feb398cd0061",
 "created": "2019-03-8T08:17:27.000Z",
 "name": "Janero",
 "description": "Exploits vulnerability, dumps file",
 "labels": ["Trojan"]
 "kill_chain_phases": ["Exploit","Stage"]
 }
```


Figure 2: STIX 2 malware construct generated by our tool

between threat actions and artifacts to generate STIX constructs of these artifacts and associate the temporal relationship between them using STIX 2 attribute `kill_chain_phase`. These constructs are machine-readable and therefore, they can be processed automatically by countermeasures such as firewalls, Web Application Firewalls (WAF), and Intrusion Detection Systems (IDS).

[The types of artifacts exhibit high predictability] : Based on our analysis, we found that the different types of artifacts that are extracted from the dataset exhibit high temporal order correlation. For instance, as shown in Figure 1, vulnerability exploitation artifact (script) is often followed by creating a dump file on the Victim's host. This is after we extracted the temporal relations between them and calculated the correlation only based on temporal relations between every two types of artifacts. This finding is important for cybersecurity tools to detect malicious behaviors based on a sequence of temporal events and their artifacts to avoid the high false positive rate of IDS.

5 FUTURE WORK

We learned in this preliminary work that the types of malicious artifacts exhibit a high temporal dependency which indicates that some types of artifacts are preconditions to other types. We can capture this temporal relationship and use it to detect behaviors of cyber attacks based on the artifacts which they generate. These behaviors should reduce the number of false positives laid by IDS that do not consider the temporal relationship between artifacts.

However, using English temporal anchors (e.g., before, after, etc.) is a limited approach in the cyber domain and results in missing to capture a non-negligible amount of temporal relations.

To Address this limitation, we plan to expand our set of temporal anchors to include more temporal expressions used in CTI domain such as "upon successful execution of the previous step, the malware ...". To collect these terms, we intend to do the following: (1) Manually label artifacts reported by cyber threat information with an arc that represents the temporal relations between artifacts based on human intelligence. Then, we will use the text around these artifacts to learn the words that indicate the temporal relation amongst artifacts, and (2) we plan to use machine learning techniques to detect cyber malicious activity based on the types of artifacts and the temporal order of generating these artifacts by any given process to inform the security analyst of the suspicious processes running on the network machines.

REFERENCES

[1] MITRE Corp. (1958). MITRE ATT&CK. https://attack.mitre.org/ . (Last accessed: November 12, 2018)

[2] STIX. The Structured Threat Information eXpression.https://stixproject.github.io/

[3] Ghaith Husari, Ehab Al-Shaer, Mohiuddin Ahmed, Bill Chu, and Xi Niu. (2017). TTPDrill: Automatic and Accurate Extraction of Threat Actions from Unstructured Text of CTI Sources. In Proceedings of the 33rd Annual Computer Security Applications Conference (ACSAC 2017). ACM, New York, NY, USA, 103-115.

[4] Symantec Corp. Symantec Security Center. https://www.symantec. com/security response/.

[5] FireEye, Inc. FireEye cyber threat intelligence. https://www.fireeye.com/solutions/cyber-threat-intelligence.html.