An ATT&CK-KG for Linking Cybersecurity Attacks to Adversary Tactics and Techniques*

Kabul Kurniawan[1], Andreas Ekelhart[1,2], and Elmar Kiesling[1]

[1] WU Wien - Vienna University of Economics and Business, Welthandelsplatz 1, Vienna, Austria {first.last}@wu.ac.at
[2] SBA Research, Floragasse 7, Vienna, Austria aekelhart@sba-research.org

Abstract. Leveraging knowledge graph techniques to detect and analyze cyber attacks is a promising research direction at the interface between the semantic web and security research communities. In this paper, we build on prior work and develop a vocabulary to extend a cybersecurity knowledge graph with adversary tactics and techniques. Using this vocabulary, we represent rich threat intelligence instance data from MITRE Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) in a knowledge graph. This knowledge can be used to contextualize indicators of compromise from log messages, identify potential attack steps, and link them to cybersecurity knowledge. To demonstrate the benefits of the approach, we link low-level threat alerts produced by community rules to the cybersecurity knowledge graph.

Keywords: Cybersecurity · Knowledge Graph · Attack Pattern · ATT&CK.

1 Introduction

Cybersecurity Threat Intelligence (CTI) can help defenders to identify vulnerabilities, understand malware behavior, and contextualize attack patterns. ATT&CK[3], published by MITRE, is a well-established source of such intelligence; it provides a taxonomy and instance knowledge for adversary tactics and techniques curated from real-world observations. ATT&CK supports various security use cases from the adversary and defender perspective, such as adversary emulation, red teaming, defensive gap assessment, identification and modeling of adversary behavior [5].

Although such CTI resources can be very useful for security experts, they are also typically difficult to relate to other cybersecurity information and operational data and not suitable for querying and automated machine interpre-

 This work was sponsored by the Austrian Science Fund (FWF) and netidee SCIENCE under grant P30437-N31. The competence center SBA Research (SBA-K1) is funded within the framework of COMET — Competence Centers for Excellent Technologies by BMVIT, BMDW, and the federal state of Vienna, managed by the FFG. The authors thank the funders for their generous support.

[3] https://attack.mitre.org/matrices/enterprise/

tation [3]. Several researchers addressed this limitation by introducing knowledge graphs for related cybersecurity information. Unified Cybersecurity Ontology (UCO) [6], for instance, integrates a number of cybersecurity standards (STIX[4], CyBox[5], CVE[6], CAPEC[7], CEE[8], and CVSS[9]) into an openly available ontology. The SEPSES Cybersecurity Knowledge Graph (CSKG)[3] builds on a similar set of standards, but provides a continuously updated, integrated cybersecurity knowledge graph with rich instance data accessible via various access mechanism. Neither UCO nor the SEPSES CSKG, however, include the CTI embodied in ATT&CK. Xiong et al. [7] propose a threat modelling language called *enterpriseLang* based on the MITRE ATT&CK Matrix. It uses a domain-specific language (DSL) based on the *Meta Attack Language* (MAL) framework to describe system assets, attack steps, and defenses. Hemberg et al. [2] propose *BRON*, which links MITRE ATT&CK, CVE, CWE[10], and CAPEC to identify tactics and techniques via a graph database. However, both *enterpriseLang* and *BRON* rely on custom, hard-coded data models and do not provide a standard model for ATT&CK. Furthermore, such custom data models lack semantic interoperability and make it difficult to, e.g., store and exchange the model and query the data via SPARQL.

We fill this gap by developing a standard model for ATT&CK based on RDF-S[11] and OWL[12], and integrate its instance data into the SEPSES CSKG. To identify high-level attack steps while abstracting from low-level indicators, we introduce a method to (i) translate community-based threat detection rules from sources such as *Sigma*[13] into SPARQL queries and (ii) to link the alerts they produce to adversarial tactics and techniques defined in ATT&CK. Figure 1 illustrates the resulting overall process to link cybersecurity attacks to adversary tactics and techniques.

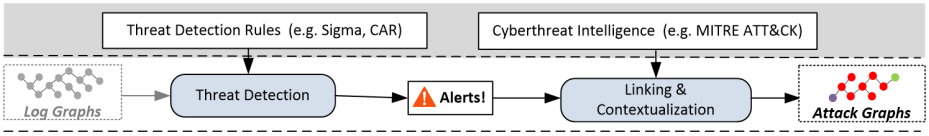

Fig. 1: Threat detection and adversarial tactic and technique linking

[4] Structured Threat Information Expression, `https://stixproject.github.io/`
[5] Cyber Observable eXpression, `https://cyboxproject.github.io/`
[6] Common Vulnerability Exposure, `https://cve.mitre.org/`
[7] Common Attack Pattern Enumeration and Classification, `https://capec.mitre.org/`
[8] Common Event Expression, `https://cee.mitre.org/language/syntax.html`
[9] Common Vulnerability Scoring System, `https://www.first.org/cvss/`
[10] Common Weakness Enumeration, `https://cwe.mitre.org/`
[11] Resource Description Framework Schema, `https://www.w3.org/TR/rdf-schema/`
[12] Web Ontology Language, `https://www.w3.org/TR/owl-features/`
[13] Sigma rule repository, `https://github.com/SigmaHQ/sigma`

2 Knowledge Graph Construction

We first defined an ontology based on the existing schema by MITRE to represent and publish attack data [5]. Figure 2 provides an overview of the re-

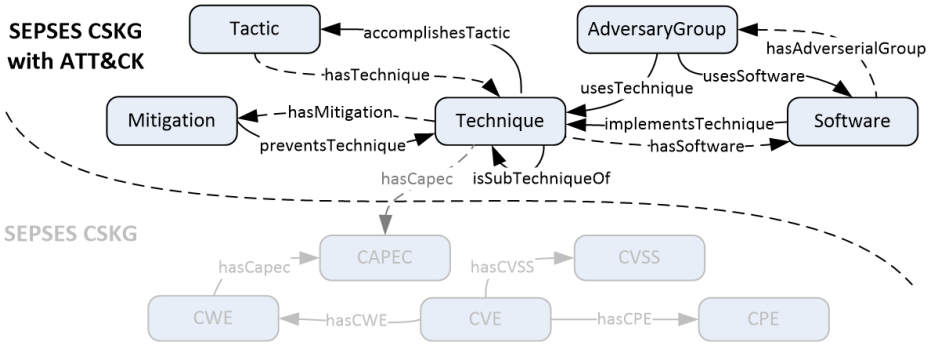

Fig. 2: An extended SEPSES CSKG Ontology with ATT&CK

sulting ATT&CK ontology[14]. The ontology consists of five main classes (i.e., Tactic, Technique, Mitigation, AdversaryGroup, and Software) and a set of data and object properties. The Tactic class represents tactical adversary goals during the attack, whereas a Technique represents *how* an adversary achieves the tactical objective; the `att:accomplishesTactic` property links Techniques and Tactics. Techniques may have more specific sub-techniques; to this end, the `att:isSubTechnique` property can be used to self-link the Technique class. Mitigation represents measures that can be used to prevent Techniques from being executed. We defined the `att:preventsTechnique` property to link between mitigations and Techniques. The AdversaryGroup represents a threat/actor group that typically represents persistent threat activity. The `att:usesTechnique` property links the group to the Technique class. Software represent software/tools that are used to implement a Technique, which is expressed via the property `att:implementsTechnique`. We also defined several inverse properties such as `at:hasMitigation`, `hasTechnique`, `hasSoftware`, etc. Furthermore, Technique links to the existing CAPEC attack pattern knowledge in the SEPSES CSKG via the `att:hasCAPEC` property.

We used RML[16], a declarative RDF mapping language to map and transform MITRE ATT&CK resources[17] into RDF based on the developed ATT&CK ontology. The constructed RDF graphs are automatically stored in a triplestore together with the existing SEPSES CSKG and integrated with information from

[14] `https://w3id.org/sepses/vocab/ref/attack`
[15] As per August 2, 2021 (cf. `https://w3id.org/sepses/dumps/attack`).
[16] RDF Mapping Language, `https://rml.io/`
[17] MITRE publishes the ATT&CK resource in a JSON format

#Axiom	#Class	#ObjectProperty	#DataProperty	#Individual
61	5	10	7	4054

Table 1: ATT&CK Knowledge Graph Statistics[15]

other standards, such as CAPEC, CVE, CWE, CPE[18] and CVSS (see Figure 3 steps (1) and (2)). Table 1 provides summary statistics on the elements currently in the ATT&CK knowledge graph.

3 Prototype Implementation

We implemented the threat detection concept introduced in Section 1 using *Sigma*, an open, community-driven, and generic rule format for threat detection in logs and included them in our threat detection pipeline. *Sigma* provides more than 950[19] rules/signatures (written in a YAML) for different log sources (e.g., application, network, web logs) and platforms (e.g Linux & Windows).

Figure 3 shows an example; based on the Sigma rule specification[20], we translate the existing *Sigma* rules into SPARQL queries and also transform other rule metadata (e.g. *title, description, log source, and tags*[21]) into RDF (3).

The translated rules can then be used to detect Indicators of Compromise (IoCs) in RDF log graphs[22] (4). Once detected, the respected alert will automatically be linked to the corresponding attack technique in the ATT&CK knowledge graph (5).

As we can see, the execution of */tmp/vUgefal* located in */tmp/* has been detected based on a *Sigma* rule. Consequently, a *Program Execution in Suspicious Folder* alert has been raised. The detected alert is automatically linked to the *T1204.002*[23] (*User Execution: Malicious File*) technique in the ATT&CK knowledge graph, as it is identified in the Sigma rule (see *tags* value). Finally, an analyst can further explore and integrate additional information from the ATT&CK knowledge graph (e.g. via SPARQL Query federation) as defined in the ontology (e.g. tactics, mitigations, adversary group, and attack patterns (CAPEC)).

4 Conclusions and Future Work

In this paper, we extended the SEPSES CSKG with threat intelligence from MITRE ATT&CK and used it to link indicators of compromise to adversarial

[18] Common Platform Enumeration, `https://nvd.nist.gov/products/cpe`

[19] As per August 2 2021.

[20] `https://github.com/SigmaHQ/sigma/wiki/Specification`

[21] *tags* associates Sigma rules with ATT&CK techniques.

[22] Note: tools such as SLOGERT [1] or [4] can be used to transform log data into RDF.

[23] `https://w3id.org/sepses/resource/attack/technique/T1204.002`

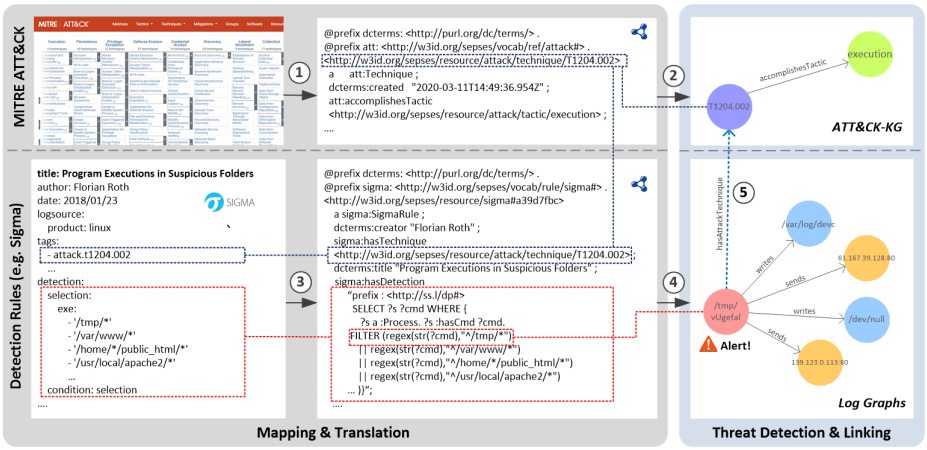

Fig. 3: KG-based threat detection with automated Sigma rule translation

tactics and techniques. The result is of high practical relevance, as demonstrated in the application that automatically identifies, contextualizes and links alerts from log messages to rich knowledge on techniques, tactics, attack patterns, vulnerabilities, and potential mitigations in the CSKG. For future work, we plan to evaluate the approach in a real-world setting, incorporate various alternative detection mechanisms (e.g., provenance-based detection and graph queries), and develop mechanisms to link individual steps in an attack campaign. Ultimately, this will provide a foundation for a new generation of tooling to support semantic security analytics.

References

1. Ekelhart, A., Ekaputra, F.J., Kiesling, E.: The SLOGERT Framework for Automated Log Knowledge Graph Construction. In: ESWC (2021)
2. Hemberg, E., Kelly, J., Shlapentokh-Rothman, M., Reinstadler, B., Xu, K., Rutar, N., O'Reilly, U.M.: Linking Threat Tactics, Techniques, and Patterns with Defensive Weaknesses, Vulnerabilities and Affected Platform Configurations for Cyber Hunting. arXiv:2010.00533 [cs] (2021)
3. Kiesling, E., Ekelhart, A., Kurniawan, K., Ekaputra, F.: The SEPSES Knowledge Graph: An Integrated Resource for Cybersecurity. In: ISWC (2019)
4. Kurniawan, K., Ekelhart, A., Kiesling, E., Winkler, D., Quirchmayr, G., Tjoa, A.M.: Virtual Knowledge Graphs for Federated Log Analysis. In: ARES (2021)
5. Strom, B.E., Applebaum, A., Miller, D.P., Nickels, K.C., Pennington, A.G., Thomas, C.B.: Mitre ATT&CK: Design and Philosophy. Technical report (2018)
6. Syed, Z., Padia, A., Finin, T., Mathews, L., Joshi, A.: UCO: A Unified Cybersecurity Ontology. In: Proceedings of the AAAI Workshop on Artificial Intelligence for Cyber Security (2016)
7. Xiong, W., Legrand, E., Åberg, O., Lagerström, R.: Cyber security threat modeling based on the MITRE Enterprise ATT&CK Matrix. Software and Systems Modeling (2021)