The Evolution of the CRUSOE Toolset: Enhancing Decision Support in Network Security Management

Martin Husák*, Lukáš Sadlek*†, Martin Hesko†, Vít Šebela†, Stanislav Špaček*
*Institute of Computer Science, Masaryk University, Brno, Czech Republic
†Faculty of Informatics, Masaryk University, Brno, Czech Republic
husakm@ics.muni.cz, sadlek@mail.muni.cz, 484964@mail.muni.cz, 531030@mail.muni.cz, spaceks@ics.muni.cz

***Abstract*—This demo paper presents the recent development of the CRUSOE toolset. CRUSOE enables cyber situational awareness and provides decision support for network security management. The first public version from 2021 used a combination of active and passive network monitoring to enumerate cyber assets and discover their vulnerabilities, visualize the collected data in a dashboard, conduct a risk assessment to recommend the most resilient infrastructure configuration, and facilitate attack mitigation. It also used novel approaches, such as a graph database for storing the data on cyber assets, which essentially became a knowledge graph for network security management. In the recent development, we managed to automate the deployment of CRUSOE via Ansible and Docker. Further, we implemented additional recommender systems and attack impact assessment capabilities and their visualizations. Finally, several sample datasets were created to facilitate the demonstration of the toolset and to enable testing it without one's data.**

I. Introduction

The growing size and complexity of today's computer network and infrastructures make it difficult to build and maintain the *cyber situational awareness* (CSA), i.e., the ability to perceive and comprehend the cyber environment and be able to project the situation in the near future. Namely, the personnel of cybersecurity incident response teams (CSIRT/CERT) and security or network operation centers (SOC/NOC) should be aware of the situation in the network to effectively prevent, discover, and mitigate cyber attacks and avoid mistakes while doing so. To achieve the CSA, we developed the CRUSOE toolset in 2021, published it as open source [1], and described it in a research paper [2]. Since then, we maintained the toolset and enhanced it with novel features based on our latest research. In this paper, we present the novel features of the toolset, namely the novel decision support capabilities, automated deployment, and sample datasets.

The CRUSOE toolset allows the users to achieve CSA in large and heterogeneous environments, such as campus networks, where it is difficult to keep track of all the devices and their security posture. The toolset user is a member of a security team or operations center and is guided through the OODA loop (Observe, Orient, Decide, Act) that facilitates decision-making in time-critical situations. Each phase is supported by a dedicated set of components referred to by the name of the phase. The Observe phase is supported by a plethora of orchestrated data collection tools, namely active network scanners (e.g., Nmap, dedicated vulnerability scanners, CMS scanners) and passive network traffic measurement based on IPFIX. These tools provide the list of cyber assets (hosts and services) and their fingerprints, which are updated at least once a day. This is accompanied by manually inserted information on network segmentation, identification of critical infrastructure, contacts on responsible users or administrators, and other entries. The Orient phase is supported by a dashboard that visualizes the collected data. The user can access contextual information on any known asset in the network or check how many devices in each subnet may be vulnerable to a particular CVE. The Decide phase originally contained only a tool that chose the most resilient configuration of the infrastructure (or its critical parts) out of possible configurations that fulfilled the operational needs with respect to current threats. This phase, together with corresponding visualizations, was enhanced the most in the recent development, as presented later. Finally, the Act phase is supported by a unified interface that facilitates interaction with various attack mitigation tools (routers, firewalls, mail filters, etc.).

In case of an incident, the user can assess the situation by checking the recent information on the assets in the dashboard, namely, check whether the incident affects critical infrastructures and look up the type of assets in question and responsible contact quickly. Decision support can be used to reconfigure critical infrastructures, while the Act tools allow faster application of mitigation actions. The novel version of CRUSOE enhances the decision support capabilities for both critical and non-critical assets, among other features.

II. Novel Capabilities

There are two novel capabilities of the CRUSOE toolset that complement the existing recommendation of the most resilient configuration implemented in the Decide phase. The novel capabilities belong to the Decide phase as well; they recommend the next possible targets or current attacks and model the spread of infection in the infrastructure.

First, the *Recommender System* is based on our previous proposal and follow-up implementation [3], [4]. The original motivation for this component was the question: "If a certain device is infected (e.g., with ransomware), which other devices can be infected or become infected in the near future?" This is paramount in incident response, namely in cases of ransomware infection – once an infection is reported at one or more devices, it is likely that there are other devices that are or will be infected, too, and it is crucial to quickly identify them and notify their

978-3-903176-66-9

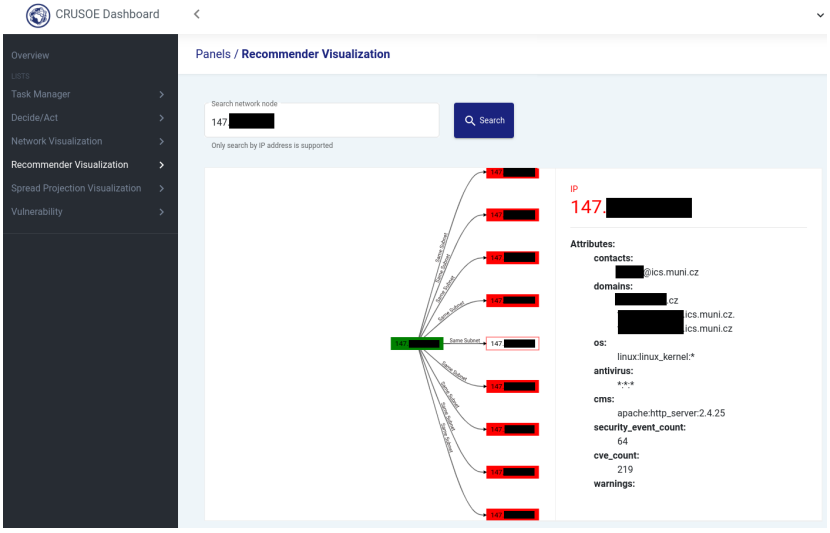

Fig. 1: Recommender system visualization: For the requested hosts, the ten most similar hosts in close proximity are displayed, the edge caption highlights their most similar feature, and full details on the host are displayed in the panel on the right.

users or take preemptive measures. Since CRUSOE collects data that can be used for the recommendation, there is no need to collect them on demand, and the user can simply run the recommendation. This is done via a graphical interface, which can be seen in Figure 1. The user types in an IP address or domain name and receives a list of devices identified by their IP addresses that are most similar to the device on the input and are close to it. The user can then, for example, quickly collect the contacts of users or administrators to send them warnings.

We assume that the malware will spread among devices that are similar. The similarity between the two devices can be viewed as them having the same OS and other SW, serving a similar purpose as a workstation or server, or having a similar history of cybersecurity incidents. Also, we assume malware spreads among the devices that are ”close” in any meaning of proximity or distance. We can assume the physical distance or the numerical difference between the devices' IP addresses. However, we mostly rely on distances between nodes in the CRUSOE graph database, assuming the graph traversal over certain types of nodes and edges. Thus, we can calculate the distance between the hosts in the same or different subnet, hosts used by the same users, or hosts located in the same or different organization's departments. The proposed recommendation algorithm combines the various similarity and distance features into a single *risk score*, which is then used to prioritize most similar hosts in close proximity and put their list on the output. Readers interested in details are kindly referred to our previous work [3], [4].

Second, the *Spread Projection* goes further than the recommender system while using its fundamental principles. While the recommender system primarily serves for early warning, it can be used in many other cases as well, namely in attack modeling and impact assessment, either in real-time or during a static security assessment. The spread projection uses the risk score from the recommender system to estimate the probability of an attack spreading from one device to another (e.g., via automated malware propagation or attacker's lateral movement). This is used to estimate the attack spread in the whole network or its part (e.g., subnet or certain infrastructure). The graphical interface of this component can be seen in Figure 2. The user types in the IP address of a device where the initial compromise happened or can happen. Further, the user selects a subnet or a subsystem from a list of known ones. In return, the user receives a graphical representation of an attack spread based on a Bayesian network. Each host in the subnet of infrastructure is modeled as a node in the Bayesian network. Starting from the initial host, new nodes are connected by edges representing a potential attack spread if the risk score (based on the similarity and distance of the two hosts) is bigger than a given threshold. Connecting new nodes is iterated with all newly added nodes until no new nodes can be added. The user can see how far the attack can propagate and with what probability. This is advantageous, namely in situations where the attacker can move laterally (this is referred to as pivoting or island hopping in literature) to exploit the target not directly but via a seemingly unrelated host. Readers interested in the detailed description of the Bayesian network construction algorithm and its implications are kindly referred to the research paper proposing the idea [5].

III. Deployment Automation and Datasets

Apart from providing novel capabilities, we also aim to facilitate deployment and allow potential users to familiarize themselves with the toolset before deploying it. Therefore, we first automated the deployment using Ansible and Docker. The original Ansible scripts, often suited only to individual components, were unified to allow for the whole toolset to be built using only one command. Various configurations were centralized into a single file for easier setup before the

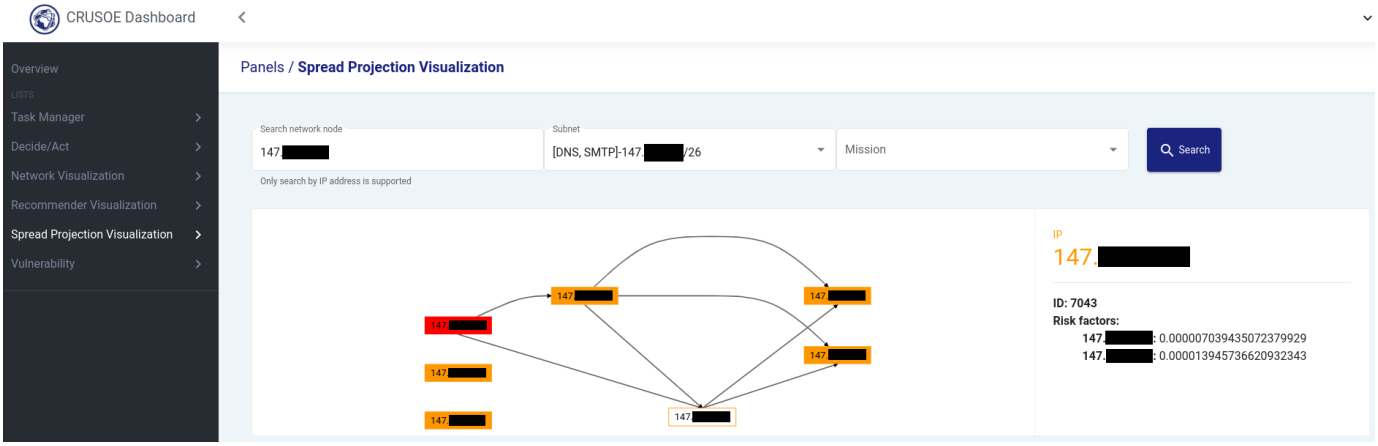

Fig. 2: Attack spread projection visualization: Starting from the node of initial compromise, the hosts in the selected subnet or infrastructure are connected to the Bayesian network of possible attack propagation. Details are provided in the panel on the right.

installation. Docker was used in a similar manner as Ansible, and our experiments show faster deployment with Docker, which we therefore recommend to users.

Second, we provide sample datasets. Thus, the potential users may first deploy the toolset locally, fill it with data from the dataset, and familiarize themselves or experiment with the toolset before deploying it in their live environment and running data collection themselves. Three datasets have been created so far. The first dataset, which is already publicly available, is based on the collection of network traffic and system logs from an exercise in cyber range [6]. We took the environment description from the cyber range exercise, converted it into the CRUSOE data model, and enhanced it with the data extracted from network traffic and systems logs. Since the source data were already public, no privacy issues were involved. For our internal needs, we prepared a dataset based on data collected in our own campus network, which we cannot share due to privacy and security concerns and difficult anonymization of the data. We also prepared another dataset containing a combination of IT and OT assets (industrial robotics environment), which will be published soon. All the datasets are saved as a dump of the Neo4j database and can be loaded using a single command as described in the documentation.

IV. Conclusion and Future Work

In this demo paper, we presented the recent additions to the CRUSOE toolset [1], [3] that extend the capabilities for achieving cyber situational awareness via continuous data collection and visualization, provide decision support, and facilitate attack mitigation. Namely, we integrated a novel recommender system into the toolset to recommend similar devices in close proximity to a given asset. The recommender system is used for attack impact assessment and is accompanied by visualizations. Further, we automated the deployment of CRUSOE via Ansible and Docker and provided sample datasets so that anyone can try running CRUSOE locally for testing purposes without the need to collect their own data.

In our future work, we aim to revise the Observe phase, focusing on improving the data quality and (semi-)automating data collection, where human input is still required. For example, the identification of critical hosts and services is typically done manually for the most critical assets but can be supported by automated identification of additional critical assets and their dependencies [7]. We will also reflect on changes in cybersecurity tooling, such as adapting CRUSOE to work with the novel CVSS v4.0. CRUSOE will remain an open-source project that is accessible on GitHub and developed by CSIRT-MU. Ultimately, we are also going to evaluate and quantify the benefits of using the dashboard, e.g., by measuring the time spent on handling an incident and reducing the human error rate caused by the lack of cyber situational awareness.

Acknowledgment

This research was supported by OP JAK "MSCAfellow5_MUNI" (No. CZ.02.01.01/00/22_010/0003229).

References

[1] CSIRT-MU, "CRUSOE: A Toolset for Cyber Situational Awareness and Decision Support in Incident Handling Inspired by the OODA Loop," https://github.com/CSIRT-MU/CRUSOE, 2021.

[2] M. Husák, L. Sadlek, S. Špaček, M. Laštovička, M. Javorník, and J. Komárková, "CRUSOE: A toolset for cyber situational awareness and decision support in incident handling," *Computers & Security*, vol. 115, p. 102609, 2022.

[3] M. Husák, "Towards a Data-Driven Recommender System for Handling Ransomware and Similar Incidents," in *2021 IEEE International Conference on Intelligence and Security Informatics (ISI)*, 2021.

[4] V. Bouček and M. Husák, "Recommending Similar Devices in Close Proximity for Network Security Management," in *2023 19th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2023, pp. 481–484.

[5] M. Husák and M. Javorník, "Lightweight Impact Assessment and Projection of Lateral Movement and Malware Infection," in *2023 IEEE Conference on Communications and Network Security (CNS)*, 2023.

[6] D. Tovarňák, S. Špaček, and J. Vykopal, "Traffic and log data captured during a cyber defense exercise," *Data in Brief*, vol. 31, p. 105784, 2020.

[7] L. Sadlek and P. Čeleda, "Cyber Key Terrain Identification Using Adjusted PageRank Centrality," in *IFIP International Conference on ICT Systems Security and Privacy Protection*. Springer, 2023, pp. 293–306.