Investigating Threats to ICS and SCADA Systems Via Honeypot Data Analysis and SIEM

Tameem ud Din[1], Usman Zia[1], Mahnoor[1], Laiq Hasan[1], Syed M. Ali Uddin Hafee[2],
[1]University of Engineering and Technology Peshawar Pakistan,
[2]NED University of Engineering and Technology Karachi, Pakistan,
***Correspondence**: Tameem Ud Din - 20PWCSE1866@uetpeshawar.edu.pk

Citation | Din. T, Zia. U, Mahnoor, Hasan. L, Hafee. S. M. A. U, Investigating Threats to ICS and SCADA Systems Via Honeypot Data Analysis and SIEM, IJIST, Special Issue. pp 189-197, April 2024

Received| May 04, 2024, **Revised|** May 12, 2024, **Accepted|** May 18, 2024, **Published|** May 27, 2024.

Supervisory Control and Data Acquisition (SCADA) and Industrial Control Systems (ICS) are crucial for managing essential infrastructure, but their exposure to the internet has made them vulnerable to cyber threats, which can lead to significant consequences. This study presents an innovative approach to investigating cyber threats to SCADA and ICS systems by combining open-source honeypot deployment, log analysis, and integration with open-source SIEM solutions to enhance threat detection capabilities and incident response. A Conpot honeypot was deployed in a containerized environment on a cloud platform and exposed to the internet to collect real-world threat data, which was then analyzed by the Wazuh SIEM solution and integrated with TheHive for security orchestration and automated response. The analysis of the honeypot logs and SIEM alerts revealed various types of attacks, including brute force login attempts, reconnaissance and vulnerability scanning, and unauthorized access attempts, originating from multiple countries and targeting different industrial protocols. The integration with TheHive enabled the creation of playbooks for automating response actions, such as blocking malicious IP addresses or isolating infected systems. The study demonstrates the effectiveness of this combined approach using open-source tools in protecting critical infrastructure and enhancing cybersecurity posture for SCADA and ICS systems.

Keywords: ICS; SCADA; OT; Honeypot; Critical Infrastructure Protection.

Introduction:

Supervisory control and data acquisition (SCADA) systems are considered a type of industrial control system that allows users to monitor, acquire data from, and control industrial processes locally or remotely through sensors and actuators [1]. Power plants and water treatment facilities are examples of traditional industrial systems that were designed to operate in highly controlled and separated settings. However, the recent exposure of Industrial Control Systems (ICS) to the Internet has made access and technological adaptation easier, which has led to the exploitation of security holes by attackers to launch attacks against ICSs [2]. These attacks can significantly impact the economics and national security of countries. To identify possible threats and comprehend the terrain of these assaults, ICS honeypots are deployed [3]. Honeypots are an interesting security concept; instead of keeping attackers out, you want to invite them in [4]. They are typically divided into three categories: low-, medium-, and high-interaction. The key distinctions between these types are their capacities for data collection, maintenance, and deployment [5]. The higher the level of interaction, the more data the honeypot can capture. High interaction ones are most similar to real systems and can collect the most data. In industrial environments where attacks occur, such as ICS/SCADA, honeypots need to be very similar to real systems to reduce detection risk [6]. This paper investigates the threats to SCADA and ICS systems using open-source honeypot deployment, log analysis, and integration with open source SIEM solutions to enhance threat detection capabilities and incident response. Furthermore, it studies the viability of putting open-source honeypots in the cloud to detect attacks on these systems. In addition, the study investigates how to integrate these honeypots with an open-source SIEM solution. This combination strategy tries to improve threat detection by connecting honeypot data with other security logs. Finally, the research extends beyond simply recognizing attackers. It digs into a deep examination of the collected traffic to determine attack kinds, sources, and other useful information for creating defensive tactics. The core contributions of this paper are: 1) Evaluation of Open-Source Honeypot Deployment on Cloud 2) Integration of Honeypot with Open-Source SIEM for Threat Intelligence 3) In-Depth Analysis of Attack Patterns.

Background And Related Work:

This section gives an overview of honeypots, and SIEM Solutions, summarizes relevant efforts, and honeypot deployments.

Honeypots:

Honeypots are software applications designed to mimic real systems, deceiving attackers while serving as a decoy. Honeypots are classified into two types: production honeypots, which are installed within a network to deceive insiders, and research honeypots, which are deployed on the internet to attract attackers from the outside. Furthermore, honeypots can be classed according to how closely they engage with attackers. Low-interaction honeypots simulate basic services but do not completely function, whereas high-interaction honeypots provide a full operating system with real services [7]. To set up a honeypot, a specific honeypot image is deployed on the machine. Once configured, a port is assigned to the honeypot. Any attempt to access this port redirects the attacker to the honeypot, which appears to be a genuine system. Through this redirection, the attacker’s machine ID, IP address, and other critical parameters are logged in the honeypot’s log file, using these logs we can move to further analysis. We used conpost, an open-source, general-purpose ICS honeypot.

Conpot:

The ICS honeypot we have used is called Conpot. Conpot is a SCADA honeypot that serves as a valuable tool for emulating SCADA systems and detecting potential threats within SCADA device networks. Developed by The Honeynet Project, Conpot is designed to be easy to implement and provides simulation capabilities for protocols like HTTP, Modbus, and SNMP, as well as integration with programmable logic controllers (PLC). It features a logging system that monitors and records any unauthorized changes made by intruders, offering detailed event logs with millisecond accuracy. By mimicking the behavior of real SCADA systems, Conpot utilizes a logging system to monitor any changes that are made by intruders. The honeypot logs events of HTTP, SNMP, and Modbus services with millisecond accuracy and offers basic tracking information such as source address, request type, and resource requested in the case of HTTP [8].

Security Information and Event Management:

Security Information and Event Management (SIEM) is a vital security software/platform that analyzes security events. It works by examining log files sent to it and following the paths we set. SIEM analyzes log files using established rules and generates alerts depending on the results. Custom rules can be written for specific types of logs to generate alerts when certain circumstances are satisfied. To successfully analyze logs, we must create a decoder for any log types that do not follow the standard format. SIEM does, however, include preconfigured decoders for JSON structured logs.

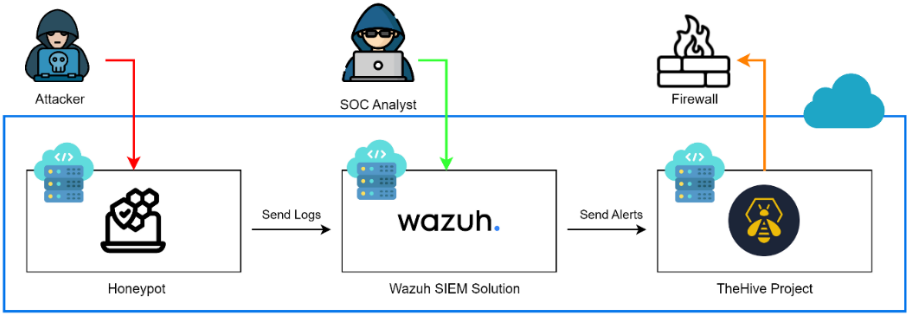


Figure 1: Holistic Approach to SCADA/ICS Threat Investigation with Honeypot, Wazuh, and the Hive.

Related Work:

The use of honeypots as a security precaution and research tool has been extensively studied, including industrial control systems (ICS) and supervisory control and data acquisition (SCADA). for protecting Operational Technology (OT) environments. The following literature review focuses on significant studies and research efforts on the subject of honeypots and their use to secure critical infrastructure. Nasution et al. concentrate on the low-interaction honeypot HONEYD's potential for enhancing security systems in their paper [9]. They go through how HONEYD is configured and deployed, as well as how it can identify and record assaults, giving security analysts important information. A thorough analysis of the application of honeypots and honeynets in Cyber-Physical Systems (CPS), Industrial IoT (IIoT), and the Internet of Things (IoT) is provided by Franco et al. [10]. Their research emphasizes the benefits and drawbacks of setting up honeypots in these settings and stresses the significance of acquiring information on new dangers. Jicha et al. [8] investigate the characteristics of SCADA honeypots, with a special emphasis on the Conpot variation. Their research provides a thorough examination of Conpot's capabilities, including its ability to mimic numerous industrial protocols and collect important information on prospective assaults on SCADA systems. Sharma and Kaul [11] conduct a detailed investigation on intrusion detection systems and honeypot-based proactive security measures in Vehicular Ad hoc Networks (VANETs) and VANET clouds. Their findings emphasize the necessity of honeypots for recognizing and mitigating risks in these dynamic and diverse contexts. Uitto et al. [12] describe the anti-honeypot strategies and procedures used by attackers to discover and avoid honeypots. Their research underlines the significance of establishing strong and resilient honeypot systems to fight such evasion strategies while maintaining effective threat monitoring and analysis. Cao et al. [13] describe DiPot, a distributed industrial honeypot system that monitors internet scanning and attack behaviors on industrial control systems. DiPot's strengths are its high-level modeling, broad data analysis capabilities, and accessible graphical frontend, which allow users to obtain insights into the present condition of ICS security. Lopez-Morales et al. [14] introduce HoneyPLC, a highly interactive, adaptable, and malware-collecting honeypot that supports a wide range of PLC models and manufacturers. Their trials show HoneyPLC's capacity to successfully disguise itself as a real device, enticing and misleading attackers while collecting vital data samples for further study. Radoglou-Grammatikis et al. [7] provide TRUSTY, a strategic custom honeypot deployment and analysis framework. It uses an adversarial game model between the attacker and the defense to optimize the number of honeypots deployed based on the attacker's behavior and the available computing resources.

Research Gap:

These studies show the growing importance of honeypots in critical infrastructure security, notably for SCADA and ICS systems. They provide useful insights on honeypot deployment, configuration, and analysis, as well as the problems and tactics involved in their effective use in detecting and mitigating cyber threats. While earlier research has examined the impact of honeypot deployment, there is still a gap in building a real-time threat analysis system. This system would use open-source and free tools to continually monitor honeypot data and extract threat intelligence to proactively mitigate threats. This is how the remainder of the paper is organized. Section 2 explains the approach and the experimental setting. The findings are examined and described in Section 3. Section 5 concludes the paper.

Material and Methods:

This study offers a comprehensive approach to investigating cyber hazards to SCADA and ICS systems. As shown in Fig. 1, it uses honeypot technology in conjunction with Wazuh, an open-source SIEM system, to analyze threats and attacks in real-time. For one week, the honeypot was deployed in a containerized environment on a cloud platform and exposed to the internet in order to collect logs. A persistent storage solution within the container ensured log retention for Wazuh's real-time analysis. Wazuh alerts were also linked to the Hive project, allowing for the building of playbooks. These playbooks managed the blocking of new malicious IP addresses detected by the honeypot using TheHive's SOAR platform, automating the response to possible threats.

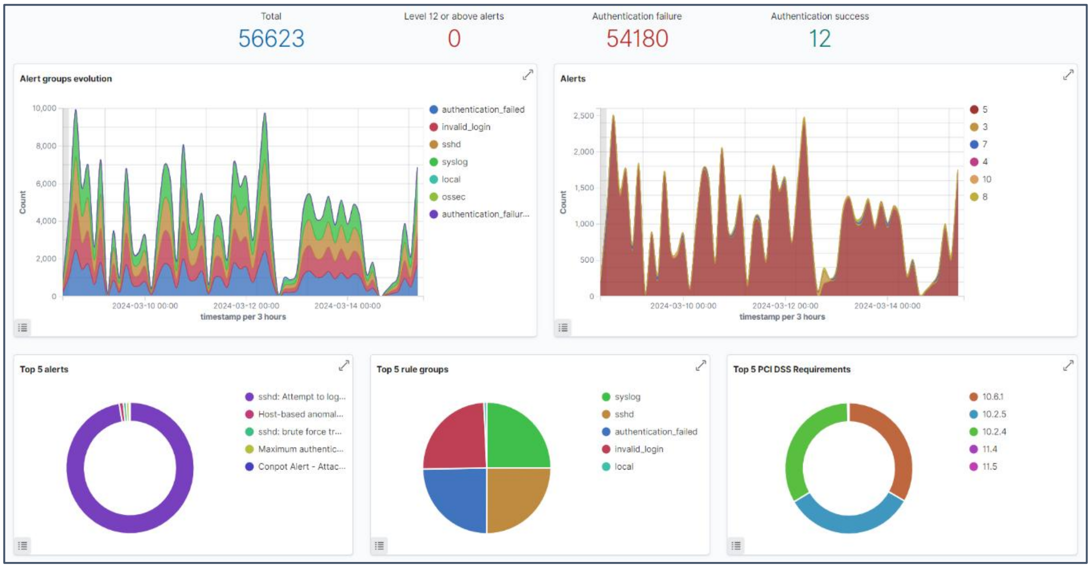
The experiment was set up by installing and configuring the Conpot honeypot, emulating a SCADA environment with various industrial protocols such as Modbus, SNMP, and HTTP. The honeypot was then exposed to the internet to attract potential attackers and gather real-world threat data. In parallel, the Wazuh SIEM agent was installed on the same system hosting the Conpot honeypot. The Wazuh agent was configured to monitor and process the log files generated by the honeypot, enabling the analysis of security events and potential threats in the dashboard of Wazuh. Further, the Wazuh was integrated with TheHive for security orchestration and automation purposes. The details of each stage are discussed in the subsections below.

Conpot Deployment:

Selection of Honeypot: The Conpot honeypot was used for this study because it is open source and can efficiently simulate SCADA systems. Conpot supports a variety of industrial protocols, including Modbus, SNMP, and HTTP, making it ideal for replicating a realistic SCADA system. Installation and Configuration: The Conpot honeypot was packed into a Docker container and placed on a dedicated system, with all necessary dependencies such as Python, Twisted, Scrapy, and PyYAML. The honeypot was then set up to imitate the needed SCADA system characteristics, such as supported protocols and services. The IEC 60870 5-104 protocol was disabled, and the default template for the configuration was utilized. Exposure to the Internet: To attract real-world cyber-attacks and collect vital threat intelligence, a honeypot with a small footprint (Conpot) was deliberately exposed to the Internet. This included opening critical ports and making the honeypot accessible via external networks. We mapped the typical ports used by common SCADA protocols to the Docker container running Conpot, leaving its default settings for a realistic attack surface.

SIEM Solution Installation and Configuration:

SIEM Solution Selection: The Wazuh SIEM Solution was chosen for this study because it is open-source, scalable, and compatible with a variety of log formats, including JSON, the major format utilized by the Conpot honeypot. Installation and Configuration: The Wazuh agent was installed on the server that ran the Conpot honeypot, and the Wazuh SIEM was configured to monitor and process not only the Conpot log files but also monitor the systems logs of the server running the Conpot. Furthermore, the Wazuh indexer server and dashboard were installed on a separate server. Rule Definition: Custom rules were built within the Wazuh SIEM to help with the examination of the Conpot log files. These rules were designed to detect specific patterns, abnormalities, or occurrences of interest associated with SCADA system attacks and vulnerabilities.



Total 56623

Level 12 or above alerts 0

Authentication failure 54180

Authentication success 12

Figure 2: Wazuh SIEM Dashboard: Real-time Monitoring and Analysis of Security Alerts and Events.

Data Collection and Analysis:

Logs Monitoring: The Conpot honeypot logs were constantly monitored, and any interactions or attacks were recorded with millisecond delay, including source IP addresses, requested resources, and any attempt to make unauthorized changes by possible attackers. Logs Analysis through Wazuh Dashboard: The Wazuh SIEM dashboard was used to examine the log data collected from the Conpot honeypot, making use of specified rules and decoders. Figure 2 illustrates the Wazuh SIEM dashboard, which enables real-time monitoring and analysis of security alerts and events. This investigation sought to uncover potential threats, attack patterns, and vulnerabilities associated with SCADA systems.

Threat Intelligence Gathering: The analysis of the Conpot honeypot logs and SIEM alerts provided valuable threat intelligence, including the tactics, techniques, and procedures (TTPs) employed by malicious actors targeting SCADA systems. Additionally, the public IP addresses of potential threat actors were collected for further investigation or potential blocking.

Integration with the Hive:

Installation and Configuration:

The Hive integration involved setting up a dedicated server for its deployment. Within The Hive, a new organization was created specifically for managing security incidents. A dedicated user account was also created and assigned to this organization. To establish communication between the Hive and Wazuh, an API-based connection was configured. This two-way communication channel allows Wazuh to automatically send security alerts to the Hive for further analysis and response.

Writing Playbooks:

The Hive's playbook functionality was leveraged to automate response actions based on incoming alerts. Playbooks were written to orchestrate specific actions, such as blocking malicious IP addresses identified by the honeypot or SIEM system directly on the firewall. This automated response streamlines the security workflow, allowing for faster threat mitigation and minimizing potential damage to SCADA/ICS systems. When the honeypot or SIEM detects a potential threat, such as a hostile IP address attempting to exploit a vulnerability, the Hive playbooks can be automatically activated. These playbooks would then execute a predetermined set of steps, such as blocking the malicious IP address at the network layer or isolating infected systems. This quicker response helped to reduce the impact of cyberattacks and prevent potential harm to SCADA/ICS systems.

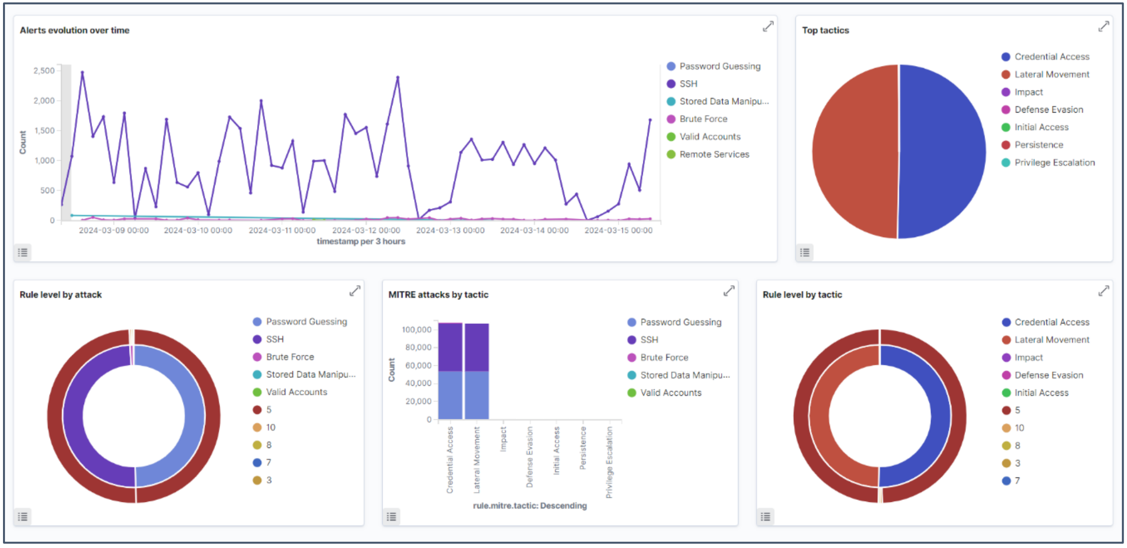


Figure 3: Wazuh SIEM Dashboard: Top Tactics and Techniques Observed in Attacks Based on the MITRE ATT&CK Framework

Result and Discussion:

Observed Attacks and Threats:

The analysis of the Conpot honeypot logs and Wazuh SIEM alerts provided valuable insights into the tactics, techniques, and procedures (TTPs) employed by malicious actors targeting SCADA systems. Figure 3 showcases the top tactics and techniques observed in attacks, categorized according to the MITRE ATT&CK framework. The following subsections provide a detailed overview of the findings.

Types of Attacks:

The study identified several types of attacks attempted against the simulated SCADA environment, including Brute Force Login Attempts: A significant number of brute force attacks (7,548) were observed, where attackers attempted to gain unauthorized access to the system by trying multiple combinations of non-existent usernames and passwords. Reconnaissance and Vulnerability Scanning: Numerous attempts were made to probe the honeypot for open ports, services, and potential vulnerabilities. These reconnaissance efforts are often precursors to more advanced attacks. Exceeding Authentication Limits: In some instances, attackers exceeded the maximum allowed authentication attempts, indicating their persistence in trying to gain unauthorized access.

Attacker's Region:

The analysis of the source IP addresses revealed that the attacks originated from various regions across the globe. Table 1 presents the distribution of attack sources by country or region.

As shown in Table 1,

Table 1: Distribution of Attack Sources by Country/Region.

Country/Region	Number of Unique IP Addresses
United States (US)	201
China (CN)	12
United Kingdom (GB)	10
Germany (DE)	10
Pakistan (PK)	8
Netherlands (NL)	8
Russia (RU)	8
Others	52

The majority of attacks originated from the United States, with 201 unique IP addresses involved. Other notable sources include China, the United Kingdom, Germany, Pakistan, the Netherlands, and Russia.

Protocols:

The honeypot was configured to emulate various industrial protocols commonly used in SCADA systems. Table 2 presents the distribution of attacks based on the targeted protocols and the corresponding port numbers. As shown in Table 2, the majority of attacks (1,168) targeted the HTTP protocol on port 8800, while 150 attacks aimed at the SNMP protocol on port 16100. It is important to note that the honeypot was intentionally exposed to the internet to attract potential attackers and collect real-world threat data. The observed attacks and their distribution provide valuable insights into the tactics, techniques, and procedures (TTPs) employed by malicious actors targeting SCADA and ICS systems.

Table 2. Distribution of Attack Sources by Protocol and Port Number.

Protocol	Port Number	Number of Attacks
HTTP	80	1168
SNMP	161	150

Discussion:

The study revealed a significantly higher number of attacks targeting the HTTP protocol on port 8800 compared to other protocols like SNMP. This observation can be attributed to several factors. Firstly, HTTP is a widely adopted protocol, and many attackers attempt to exploit web-based vulnerabilities or misconfigurations, making it an attractive target for reconnaissance and potential exploitation.

Another potential reason for the high HTTP traffic could be the detectability of the default Conpot web interface. While Conpot aims to emulate a realistic SCADA environment, the default configuration and web interface may have been recognized by attackers as a honeypot, prompting them to focus their efforts on the HTTP service. To enhance the authenticity of the simulated environment and attract a more diverse range of attack patterns, customizing the Conpot configuration to better mimic specific SCADA systems or industrial environments could be beneficial.

Furthermore, the study did not explicitly mention attacks targeting the SSH protocol, which is often used for remote administration. Leaving the SSH port open, although not directly related to SCADA protocols, could potentially attract unwanted brute-force attacks or unauthorized access attempts, which may not be relevant to the study's scope. It is recommended to avoid exposing unnecessary ports to the internet when deploying honeypots or simulating SCADA environments to maintain a focused and relevant attack surface.

Lastly, the relatively lower traffic observed for protocols like SNMP compared to HTTP could be due to the difficulty in accurately emulating the behavior and nuances of industrial protocols within a low-interaction honeypot environment with a default configuration template.

To address this challenge, future studies could explore incorporating more realistic industrial protocol implementations, potentially through the use of virtualized or emulated hardware components and integrating additional data sources or logs from real SCADA systems (with appropriate anonymization and security measures) to enrich the honeypot's behavior and log patterns.

Conclusion:

This study successfully investigated threats to SCADA/ICS systems by combining honeypots, SIEM, and security orchestration. A honeypot like Conpot acted as a decoy, collecting real-world attack data when exposed on a cloud platform. This data was then analyzed by Wazuh, a SIEM solution, to identify attack patterns and sources. Finally, Hive, a security orchestration platform, automated responses based on predefined protocols, such as blocking malicious IPs and isolating infected systems. This study demonstrates the effectiveness of this combined approach, and importantly, the value of open-source tools (Conpot, Wazuh, the Hive) in protecting critical infrastructure.

Future work could focus on enhancing the realism of the honeypot deployment, exploring the detection and analysis of more advanced attack techniques, investigating the scalability and performance of the proposed solution, integrating with external threat intelligence platforms, and fostering collaboration and information sharing among organizations and security communities.

Acknowledgement:

We would like to express our sincere gratitude to the University of Engineering and Technology, Peshawar, and NED University of Engineering and Technology, Karachi, for providing the necessary resources and support for this research project. We are thankful to our supervisors, Dr. Laiq Hasan and Syed M. Ali Uddin Hafee, for their invaluable guidance and mentorship throughout the project. Additionally, we acknowledge the open-source communities behind the tools and technologies used in this study, particularly Conpot, Wazuh, and TheHive, for their dedication to advancing cybersecurity research and development.

Author's Contribution:

Tameem Ud Din, Usman Zia, and Mahnoor were responsible for the conceptualization, implementation, and execution of the project. They conducted extensive research, deployed the honeypot and SIEM solution, performed data analysis, and integrated the automated response capabilities. Dr. Laiq Hasan provided overall supervision, guidance, and valuable insights throughout the project, ensuring its successful completion. Syed M. Ali Uddin Hafee contributed his expertise in cybersecurity and industrial control systems, providing critical feedback and recommendations to enhance the project's effectiveness.

Conflict of interest:

The authors declare that there is no conflict of interest in publishing this manuscript in the International Journal of Information Security and Threat Modeling (IJIST). The research was conducted solely for academic and scientific purposes, and the authors have no financial or other interests that could influence the objectivity or integrity of the work presented.

References:

[1] M. Mesbah, M. S. Elsayed, A. D. Jurcut, and M. Azer, "Analysis of ICS and SCADA Systems Attacks Using Honeypots," Futur. Internet 2023, Vol. 15, Page 241, vol. 15, no. 7, p. 241, Jul. 2023, doi: 10.3390/FI15070241.

[2] A. Nechibvute and H. D. Mafukidze, "Integration of SCADA and Industrial IoT: Opportunities and Challenges," IETE Tech. Rev., May 2024, doi: 10.1080/02564602.2023.2246426.

[3] A. Ara, "Security in Supervisory Control and Data Acquisition (SCADA) based Industrial Control Systems: Challenges and Solutions," IOP Conf. Ser. Earth Environ. Sci., vol. 1026, no. 1, p. 012030, May 2022, doi: 10.1088/1755-1315/1026/1/012030.

[4] "Know Your Enemy: Revealing the Security Tools, Tactics, and Motives of the Blackhat Community: Honeynet Project: 9780201746136: Amazon.com: Books." Accessed: May 06, 2024. [Online]. Available: https://www.amazon.com/Know-Your-Enemy-Revealing-Community/dp/0201746131

[5] "World Wide ICS Honeypots: A Study into the Deployment of Conpot Honeypots." Accessed: May 06, 2024. [Online]. Available: https://www.researchgate.net/publication/358166067_World_Wide_ICS_Honeypots_A_Study_into_the_Deployment_of_Conpot_Honeypots

[6] S. Chamotra, J. S. Bhatia, R. Kamal, and A. K. Ramani, "Deployment of a low interaction honeypot in an organizational private network," Proc. 2011 Int. Conf. Emerg. Trends Networks Comput. Commun. ETNCC2011, pp. 130–135, 2011, doi: 10.1109/ETNCC.2011.5958501.

[7] P. Radoglou-Grammatikis et al., "TRUSTY: A solution for threat hunting using data analysis in critical infrastructures," Proc. 2021 IEEE Int. Conf. Cyber Secur. Resilience, CSR 2021, pp. 485–490, Jul. 2021, doi: 10.1109/CSR51186.2021.9527936.

[8] A. Jicha, M. Patton, and H. Chen, "SCADA honeypots: An in-depth analysis of Conpot," IEEE Int. Conf. Intell. Secur. Informatics Cybersecurity Big Data, ISI 2016, pp. 196–198, Nov. 2016, doi: 10.1109/ISI.2016.7745468.

[9] A. M. Nasution, M. Zarlis, and S. Suherman, "Analysis and Implementation of Honeyd as a Low-Interaction Honeypot in Enhancing Security Systems," Randwick Int. Soc. Sci. J., vol. 2, no. 1, pp. 124–135, Jan. 2021, doi: 10.47175/RISSJ.V2I1.209.

[10] J. Franco, A. Aris, B. Canberk, and A. S. Uluagac, "A Survey of Honeypots and Honeynets for Internet of Things, Industrial Internet of Things, and Cyber-Physical Systems," IEEE Commun. Surv. Tutorials, vol. 23, no. 4, pp. 2351–2383, 2021, doi: 10.1109/COMST.2021.3106669.

[11] S. Sharma and A. Kaul, "A survey on Intrusion Detection Systems and Honeypot based proactive security mechanisms in VANETs and VANET Cloud," Veh. Commun., vol. 12, pp. 138–164, Apr. 2018, doi: 10.1016/J.VEHCOM.2018.04.005.

[12] J. Uitto, S. Rauti, S. Laurén, and V. Leppänen, "A Survey on Anti-honeypot and Anti-introspection Methods," Adv. Intell. Syst. Comput., vol. 570, pp. 125–134, 2017, doi: 10.1007/978-3-319-56538-5_13.

[13] J. Cao, W. Li, J. Li, and B. Li, "DiPot: A Distributed Industrial Honeypot System," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 10699 LNCS, pp. 300–309, 2018, doi: 10.1007/978-3-319-73830-7_30.

[14] E. López-Morales et al., "HoneyPLC: A Next-Generation Honeypot for Industrial Control Systems," Proc. ACM Conf. Comput. Commun. Secur., pp. 279–291, Oct. 2020, doi: 10.1145/3372297.3423356.