

Comparative Analysis of AI-Based Methods for Enhancing Cybersecurity Monitoring Systems

Federica Uccello[1(✉)], Marek Pawlicki[2], Salvatore D'Antonio[1], Rafał Kozik[2], and Michał Choraś[2]

[1] University of Naples 'Parthenope' Centro Direzionale, Isola C4, 80133 Napoli, Italy
federica.uccello@assegnista.uniparthenope.it, salvatore.dantonio@uniparthenope.it

[2] Bydgoszcz University of Science and Technology,, PBS Bydgoszcz, Poland
chorasm@pbs.edu.pl

Abstract. Traditional security monitoring systems may struggle with detecting advanced and novel cyber threats, especially as cyberattacks are becoming more sophisticated. As a consequence, there is a growing need for advanced techniques to improve defense mechanisms. Among these, Artificial Intelligence (AI) has emerged as a promising solution. The paper explores the growing popularity of AI technologies in cybersecurity and investigates three distinct methodologies aimed at improving attack detection criteria within security monitoring systems. The first methodology is based on the extraction and analysis of correlation indexes between features and target variables, enabling the identification of recurring patterns indicative of anomalies. Secondly, Association Rule Mining (ARM) is utilized to identify hidden patterns and relationships between features, leading to more accurate detection criteria. Lastly, explainable AI (xAI) is leveraged to mine advanced rules, as well as increase the transparency of the model, enabling security analysts to understand the decision-making process behind threat detection. Through a comparative analysis, the efficacy of each method is evaluated in terms of detection accuracy, precision, recall, and F1-score. Promising experimental results demonstrate the potential of AI-driven approaches to enhance the capabilities of security monitoring systems, providing organizations with a new layer of protection against an evolving threat landscape.

Keywords: Explainable Artificial Intelligence · Association Rule Mining · Rule-Based SIEM · Attack Detection

1 Introduction

Cybersecurity has become one of the most important fields in today's digital landscape, as the recent technological advancement has pushed virtually any system to move towards an Information Technology (IT) paradigm, where data has become one of the most critical assets. Together with multiple benefits, this

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
O. Gervasi et al. (Eds.): ICCSA 2024 Workshops, LNCS 14816, pp. 100–112, 2024.
https://doi.org/10.1007/978-3-031-65223-3_7

revolution has brought increasingly sophisticated cyber-actors [1] and cyber-threats [2], which can lead to disastrous consequences, as underlined by the European Commission in the European Data Strategy [3].

Therefore, enabling rapid attack detection remains a priority, and tools like Security Information and Event Management (SIEM) and Intrusion Detection Systems (IDS) stand as well-established solutions within a comprehensive monitoring and threat protection strategy [4].

However, attackers continuously evolve tactics to exploit vulnerabilities in digital systems [5]. The ever growing evolution of attack techniques require constant effort in enhancing traditional security monitoring, which often struggle to keep up with the changing requirements [6–8]. To respond to these concerns, there has been an increasing interest in featuring Artificial Intelligence (AI) and Machine Learning (ML) techniques to enhance defense mechanisms [9].

The present work seeks to investigate and compare three distinct methodologies aimed at enhancing attack detection criteria within security monitoring systems. The major contribution of the paper include:

- Usage of Correlation Index Analysis, Association Rule Mining (ARM), and Explainable AI (xAI) to mine advanced correlation criteria for SIEM.
- Testing of the mined correlation rules to discriminate Malicious and Benign network traffic.
- Evaluation of the rules in terms of Precision, Recall, F1-score and Accuracy.
- Comparison between the three different approaches, based on the experimental results.

The remainder of the paper is organised as follows: Sect. 2 overviews the related work. Section 3 presents the methods employed, including the SIEM and the AI techniques involved. Section 4 details the proposed methodology for advanced rules extraction, while Sect. 5 showcases the experimental setup and the results, comparing the performance of the different methods. The results are discussed in higher detail in Sect. 6. Finally, Sect. 7 concludes the paper.

2 Related Work

Recent advancements in ML, ARM, xAI have shown promise in augmenting security measures. Various studies have explored the application of these techniques in cybersecurity, highlighting their potential to improve threat detection accuracy and efficiency.

In particular, the enhancement of SIEM through ML and AI has been subject of various research [10–12], with the main goals to robust security measures and provide optimization. The concept of hybrid IDS has also been explored in various domains [13,14], including the Internet of Things (IoT) [15,16]. The potential of ML-enabled Network Intrusion Detection System (NIDS) is also explored in [17].

ARM is a data mining technique that allows to uncover hidden relationships within dataset's features. The literature shows several applications of ARM as

a toolkit for the enhancement of cyber incident detection. As an instance, the integration of associative rules in SIEM systems is investigated in [18]. The work shows an approach incorporating fuzzy sets theory and data mining techniques to construct associative rules based on an analysis of cyber incidents. Sequential rule mining is featured in [19] to analyse intrusion detection alerts, with the goal to reduce the amount of alerts received by security personnel. Similarly, [20] addresses the issue of false positive alerts in IDS using Modified Frequent Pattern (MFP-Tree) with the K-means algorithm to detect a range of attacks. In [21], a method for mining association rules from logs to identify intrusion behavior in cloud platforms is introduced.

Within AI, xAI methods aim to provide human-understandable explanations of ML models, to improve model transparency, robustness, and user trust. The usage of xAI in security has been discussed, among the others, in [22,23], where it is employed to enhance IDS with a focus on transparency. Trust management in IDS is the main concern of [24], where the need for transparency in ML models utilized in the cybersecurity domain is discussed. Therefore, the usage of xAI techniques such as decision tree algorithms, is proposed to enhance trust in IDS. Boosting detection accuracy and maintaining data integrity is the core focus of [25], where an evaluation within various intrusion scenarios is presented. The work aims at identifying key network features contributing to intrusion. The detection of attacks based on xAI has been also studied in [26]. The work leverages autoencoders and xAI to identify the most influential features for attack detection. Then, individual thresholds are set for such features to support attack detection.

With respect to the literature, the present work seeks to enhance traditional SIEM systems by the exploration and comparison of three AI techniques to mine advanced and highly accurate correlation criteria. The proposed methodology has the advantage of being versatile and suitable for various application domains, as it can be generalised and employed within wider scopes.

3 Innovative Methods for Improving SIEM

This section provides an overview of the SIEM system and AI techniques employed throughout the research.

3.1 Security Information and Event Management

Within the present work, SIEM systems serve as the base framework for the development and testing of the proposed hybrid SIEM, enhanced through the definition of AI-powered detection criteria. In general, SIEM systems consist of essential modules and key functionalities that can be schematized as follows:

– Data collection agent: this module is composed by one or more probes in charge of gathering data from heterogeneous sources. IDS can also be a used as a part of this set. If needed, this module can be connected to another one dedicated to parsing and normalization.

- Event correlation engine: this component acts as a kernel for the SIEM, as it correlates the data collected and raises alerts in case of anomalies.
- Historical Database/Log storage: an essential module for the storage of relevant data.
- User Interface: this module is usually connected to the storage one, and it can be used for monitoring and reporting. It can include dashboards to display information in user friendly way.

Event correlation is a crucial functionality and, traditionally, it follows a rule-based logic, consisting in identifying patterns or sequence that might be symptoms of potential threats. The correlation rules are typically crafted based on domain knowledge, known attack patterns, and heuristic analysis. However, this process of defining rules can lead to inaccurate criteria, especially when dealing with advanced and novel threats. This work extends the SIEM presented in the authors' previous research [27], with the goal to address such limitations. The SIEM was fed with simulated traffic, generated based on the considered datasets (Sect. 5), and the various correlation rules defined through AI techniques were tested and evaluated as shown in the following sections.

3.2 Correlation Index Analysis

The first approach for the generation of SIEM rules was the simplest one, based on Correlation Index Analysis. The goal was to identify the most relevant features associated with malicious behavior within the considered datasets and uncover patterns and relationships between various dataset features and the target label.

To perform the analysis, the correlation index between each dataset feature and the target label was calculated, focusing on malicious samples. The feature with the highest correlation index was selected for inclusion in the correlation rules, due to its strong association with Malicious behavior. Following the analysis, the correlation logic was defined by defining individual thresholds for the most correlated features.

3.3 Association Rule Mining

Subsequently to the previous approach, ARM was used to mine more advanced discrimination criteria. The purpose of ARM is to discover hidden patterns and relationships within datasets, and its application to the present research focused on the APRIORI algorithm [28].

APRIORI is a well-known algorithm for association rule discovery: it was applied to the preprocessed datasets to mine relationships between multiple features. In comparison to the Correlation Index Analysis approach, the APRIORI derived rules provide more detailed insights into the behaviors and patterns associated with network traffic, in particular with anomalous one.

3.4 Explainable Artificial Intelligence

The third method for the generation of discrimination criteria was xAI. In the literature, the most well-known xAI approaches include feature importance, rule-based explanations, and counterfactual explanation. These methods can help professionals understand how ML models are making decisions, as well as identifying potential biases in the models and supporting debugging problems that could arise. In addition, xAI can be employed to identify the most important features for the model's predictions.

Within the present experiment, multiple methods and criteria to build rules were considered. Ultimately, enhanced rules were mined through SHapley Additive exPlanations (SHAP) [29] and ANCHORS [30]. In the next section, all the aforementioned methods for rule extraction are formalized and detailed.

4 Proposed Approach for SIEM Rules Definition

The entire selection of AI techniques for rule definition were preceded by common initial steps.

At first, appropriate network traffic datasets D_1 and D_2 were selected (Sect. 5). To ensure quality and suitability of the datasets, and consistency among them, essential data preprocessing steps were conducted: this included handling missing values, removing duplicates, and addressing class imbalance issues. To formalize the following approach, let $X = x_1 , ... , x_n$ represent the set of features in the generic dataset D_i, and let $y = \{Malicious, Benign\}$ be the corresponding labels. The aim is to identify a set of x_{rule} and use them to build discrimination criteria between Malicious and Benign samples. Table 1 and Table 2 provide a short definition of the features involved in the definition of the set of rules for D_1 and D_2, respectively.

Table 1. D_1 Features selected for the definition of detection rules.

Feature	Description
Bwd Packet Length Mean	It stands for the mean packet length observed in the backward direction (from destination to source) in the flow. It provides direction-specific insights.
Fwd Packet Length Mean	It represents the average length of packets in the forward direction (from source to destination) in the flow.
Destination Port	This feature indicates the port number used by the destination in the network flow. It can help identifying specific services or applications involved in the communication.
Total Backward Packets	This feature represent the total number of packets transmitted in the backward direction during the flow.
Flow Duration	It refers to the time duration of a network flow

Table 2. D_2 Features selected for the definition of detection rules.

Feature	Description
Source Rate	It denotes the rate at which data packets are transmitted from the source.
Protocol	This feature specifies the communication protocol utilized in the network. It is useful for identifying different types of network traffic.
Source Loss	It indicates the percentage of data packets lost during transmission from the source. Monitoring source loss helps in assessing network reliability and identifying potential issues.
Destination Loss	The percentage of data packets lost during transmission to the destination.
TCP RTT	The TCP connection setup round-trip time

4.1 Rules Derived Through Correlation Index Analysis

As explained previously, linear correlation analysis has been performed between the dataset features and the target label, referred to as malicious samples. Let C be the correlation index and x_i be the ith feature in the dataset. The linear correlation analysis selects the feature x_i with the highest correlation index, as shown in Eq. 1:

$$j = argmax_i(C(f_i, target_{label})) \tag{1}$$

After the analysis, the features with the highest correlation index were selected and used for the creation of the correlation rules. The present method led to the creation of the rules formalized in Eq. 2 and Eq. 3, which were the best one defined using this method. In the equations, $BPLengthMean$ denotes the mean packet length observed in the backward direction, while $SrcRate$ represents the source packets per second. The criteria for the definition of the thresholds are explained in Sect. 5.

$$y_{corr} = \begin{cases} Malicious\ if\ BPLengthMean \geq T_{BPLengthMean} \\ Benign \qquad\qquad\qquad otherwise \end{cases} \tag{2}$$

$$y_{corr} = \begin{cases} Malicious\ if\ SrcRate \geq T_{SrcRate} \\ Benign \qquad\qquad\qquad otherwise \end{cases} \tag{3}$$

4.2 Rules Derived Through Association Rule Mining

To mine more advanced correlation rule, the APRIORI algorithm was applied to the datasets. As a result, the criteria shown in Eq. 4 and Eq. 5 were the best

one defined using this method. In the equations, $FPLengthMean$ denotes the mean packet length observed in the forward direction, while $DstRate$ represents the destination packets per second.

$$y_{arm} = \begin{cases} Malicious\ if\ BPLengthMean \geq TM_1 \times FPLengthMean \\ Benign \qquad\qquad otherwise \end{cases} \tag{4}$$

$$y_{arm} = \begin{cases} Malicious\ if\ SrcRate > TM_2 * DstRate \\ Benign \qquad\qquad otherwise \end{cases} \tag{5}$$

4.3 Rules Derived Through Explainable Artificial Intelligence

The third and last method involved the usage of xAI techniques to define more complex detection rules. The criteria shown in Eq. 6 and Eq. 7 were the best ones defined using this method. In Eq. 6, $DestPort$ denotes the destination port, $TotBWDP$ the total number of backward packets, and $FlowDuration$ the time duration of network flow. mean packet length observed in the forward direction, while $DstRate$ represents the destination packets per second. In Eq. 7, $Protot$ denotes the network protocol, $SrcLoss$ and $DstLoss$ the Source/Destination packets respectively that are re transmitted and/or dropped, while $TcpRtt$ indicates the TCP connection setup round-trip time.

$$y_{xAI} = \begin{cases} Benign \quad if \quad DestPort \in Whitelist \\ Malicious\ if \quad TotBWDP \leq T_{TPL} \\ \qquad\qquad\qquad or \\ \qquad\qquad FlowDuration > T_{FlowDuration} \\ Benign \qquad\qquad otherwise \end{cases} \tag{6}$$

$$y_{xAI} = \begin{cases} Benign \quad if \quad Proto \in Whitelist \\ Malicious\ if\ SrcLoss > DstLoss \\ \qquad\qquad\qquad or \\ \qquad\qquad TcpRtt < \epsilon \\ Benign \qquad\qquad otherwise \end{cases} \tag{7}$$

In the next section, the experiment setup and the performance evaluations for the three methods are detailed.

5 Experimental Setup and Results

5.1 Datasets

Within the present experiments, two datasets have been employed. The fist one is the CICIDS2017 [31], a well-recognised benchmark dataset designed for evaluating IDS against sophisticated and evolving network attacks. The choice of this dataset was motivated by its careful addressing of the flaws of other datasets,

which might suffer from outdated information, lack of diversity and anonymization of critical data. In addition, the methodology was tested using a more recent and specialised dataset, the WUSTL-IIOT-2021 [32], meant to emulate real-world industrial systems.

Essential data preprocessing steps were conducted to ensure the quality and suitability of the datasets. This included handling missing values, removing duplicates, and addressing class imbalance issues by selecting an equal number of Benign and Malicious samples for each dataset. For dimensionality reduction, the SelectKBest feature selection method from the scikit-learn library was employed. SelectKBest utilizes univariate statistical tests, such as ANOVA F-value, to identify the most relevant features for the analysis. Afterwards, the datasets were employed to simulate real-time network traffic to feed to the SIEM, once the last column representing the label was removed. Then, the SIEM rules were employed to label the simulated traffic as Benign or Malicious, and the performance of the detection was evaluated by comparison with the ground truth.

5.2 Parameter Settings

During the process to define correlation rules, the threshold values and parameters were derived from the statistical properties of the selected features and the discovered criteria.

In Eq. 2 and Eq. 3, the thresholds $T_{BPLengthMean}$ and $T_{SrcRate}$ have been set equal to the mean value of the selected feature for all the samples of the original dataset.

To define a suitable value for the TM_1 and TM_2 parameters shown in Eqs. 4 and 5, multiple approaches have been considered. Ultimately, the optimal value for TM_1 has been obtained by calculating the average $BPLengthMean$ to $FPLengthMean$ ratio for Malicious instances. Similarly, TM_2 has been set equal to the average $SrcRate$ to $DstRate$ ratio for Malicious instances.

In Eq. 6), the thresholds T_{TPL} ans $T_{FlowDuration}$ were calculated using the same logic employed for $T_{BPLengthMean}$ and $T_{SrcRate}$. The $Whitelist$ and $Blacklist$ have been derived checking the exclusive values of $DestinationPort$ for Benign and Malicious labels, respectively. The $Whitelist$ parameter from Eq. 7 has been derived using a similar criteria, but considering the $Prot$ feature instead. Finally, the ϵ parameter is defined as shown below (Eq. 8):

$$\epsilon \rightarrow 0 \tag{8}$$

This value has been selected as threshold for the $TcpRtt$ feature, as it has been observed that this feature was close to zero, on average, for Malicious samples only.

5.3 Results

The following tables show a summary of the classification report based on Precision (9), Recall (10), F1-Score (11), and Accuracy (12). In the following equations, for the sake of brevity, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) are denoted using the acronyms.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{9}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \tag{12}$$

Table 3. Performance report of SIEM rules derived through Correlation Index Analysis.

Correlation Rules	Precision		Recall		F1-score		Accuracy
	Benign	Malicious	Benign	Malicious	Benign	Malicious	
Equation 2	0.73	0.98	0.98	0.63	**0.84**	0.76	**0.77**
Equation 3	0.62	**0.99**	**0.99**	0.49	0.77	0.57	0.70

Table 4. Performance report for SIEM rules derived through ARM.

Correlation Rules	Precision		Recall		F1-score		Accuracy
	Benign	Malicious	Benign	Malicious	Benign	Malicious	
Equation 4	**1**	0.99	0.99	**1**	**0.99**	**0.99**	**0.99**
Equation 5	0.90	0.98	0.98	0.90	0.94	0.94	0.94

Table 5. Performance report for SIEM rules derived through xAI.

Correlation Rules	Precision		Recall		F1-score		Accuracy
	Benign	Malicious	Benign	Malicious	Benign	Malicious	
Equation 6	**1**	0.96	0.96	**1**	0.98	0.98	0.98
Equation 7	**1**	0.99	0.99	**1**	**0.99**	**0.99**	**0.99**

6 Discussion

As shown in the Tables, the different rules implemented present varying degrees of performance. Table 3 suggest a potential trade-off between precision and recall in the SIEM rules derived though simple Correlation Index Analysis. The performance metrics of the both rules, while providing a baseline for correlation, reveal the need for further refinement, as denoted by the Accuracy values.

The ARM approach, as depicted in Table 4, demonstrates a significant advancement in the system's discriminatory power. ARM allows to discover association rules that significantly contribute to the identification of Malicious traffic. In particular, Eq. 4 exhibits high Precision, Recall, F1-Score, and Accuracy, indicating a near-perfect performance in distinguishing between Benign and Malicious instances The effectiveness of ARM rules can be attributed to their ability to capture intricate relationships and dependencies within the network traffic data. The discovered patterns empower the intrusion detection system to make more informed decisions, leading to a substantial improvement in its overall performance.

The xAI approach, represented in Table 5, offers a unique perspective on correlation rules. The interpretability of the rules derived through xAI techniques enhances the system's transparency and facilitates a deeper understanding of the decision-making process. Both Eq. 6 and Eq. 7 showcase near-perfect Precision, Recall, F1-Score, and Accuracy. Additionally. Both xAI rules show balance between Precision and Recall, highlighting the superior potential of xAI techniques in comparison to the previous approaches. the xAI rules provide human-understandable insights into the features and patterns contributing to the decision-making process.

The comparison of the three approaches reveals a progressive refinement in the system's performance. The SIEM rules derived through correlation analysis provide a baseline, while the ARM and xAI approaches contribute advanced correlation rules that significantly enhance the system's ability to discern between Benign and Malicious traffic.

7 Conclusion

The paper explores the application of AI techniques in enhancing attack detection within security monitoring systems. Three distinct methodologies were investigated: Correlation Index Analysis, ARM, and xAI. Each method was employed to derive correlation rules aimed at improving the discrimination between Benign and Malicious network traffic within a rule-based SIEM system.

Through a comparative analysis of the experimental results, it was observed that each approach brings a different perspective in enhancing the capabilities of security monitoring systems. The Correlation Index Analysis method provides a straightforward approach to identifying correlation between features and target labels. While it offers a baseline for correlation rules, its performance may be limited in capturing complex relationships within the data.

On the other hand, ARM enables the discovery of intricate patterns and dependencies within the dataset, leading to more accurate and robust correlation rules. The ARM-derived rules exhibited significant improvement in detection

accuracy, precision, recall, and F1-score compared to Correlation Index Analysis. However, its application might not be as straightforward as the previous one, and it might be more challenging to apply to complex scenarios. Finally, xAI techniques offer human-understandable insights into the decision-making process of the model. The xAI-derived rules demonstrated near-perfect performance metrics, highlighting the superior potential of xAI techniques in enhancing attack detection criteria. Similarly to the ARM approach, this method might also bring challenges with the perspective of more intricate scenarios. Future research directions may involve further exploration of advanced AI techniques and their integration into security monitoring systems, as well as the development of hybrid approaches combining multiple methodologies to enhance detection accuracy and efficiency even further. Additionally, the scalability and real-time implementation of these techniques in large-scale environments might be worth of further investigation for practical deployment within actual systems.

Acknowledgment. This research has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 101020560 CyberSEAS, and from the Industrial Cyber Shield (ICS) No C83C22001460001 project funded by INAIL. The content of this publication reflects the opinion of its authors and does not, in any way, represent opinions of the funders. The European Commission and INAIL are not responsible for any use that may be made of the information that this publication contains.

References

1. Pawlicka, A., Choraś, M., Pawlicki, M.: The stray sheep of cyberspace aka the actors who claim they break the law for the greater good. Pers. Ubiquit. Comput. **25**(5), 843–852 (2021)
2. Kozik, R., Choraś, M.: Current cyber security threats and challenges in critical infrastructures protection. In: 2013 Second International Conference on Informatics & Applications (ICIA), pp. 93–97 (2013)
3. E. Commission, A european strategy for data.' https://digital-strategy.ec.europa.eu/en/policies/strategy-data
4. Younus, Z., Alanezi, M.: A survey on network security monitoring: tools and functionalities. Mustansiriyah J. Pure Appli. Sci. **1**(2), 55–86 (2023)
5. Pawlicki, M., Pawlicka, A., Kozik, R., Choraś, M.: The survey and meta-analysis of the attacks, transgressions, countermeasures and security aspects common to the cloud, edge and iot.Neurocomputing, 126533 (2023)
6. Duo, W., Zhou, M., Abusorrah, A.: A survey of cyber attacks on cyber physical systems: recent advances and challenges. IEEE/CAA J. Automatica Sinica **9**(5), 784–800 (2022)
7. Li, Y., Liu, Q.: A comprehensive review study of cyber-attacks and cyber security; emerging trends and recent developments. Energy Rep. **7**, 8176–8186 (2021)
8. Guembe, B., Azeta, A., Misra, S., Osamor, V.C., Fernandez-Sanz, L., Pospelova, V.: The emerging threat of ai-driven cyber attacks: a review. Appl. Artif. Intell. **36**(1), 2037254 (2022)
9. Zhang, Z., et al.: Artificial intelligence in cyber security: research advances, challenges, and opportunities. Artifi. Intell. Rev., 1–25 (2022)

10. Ban, T., Takahashi, T., Ndichu, S., Inoue, D.: Breaking alert fatigue: Ai-assisted siem framework for effective incident response. Appl. Sci. **13**(11), 6610 (2023)
11. Radoglou-Grammatikis, P.: Securecyber: an sdn-enabled siem for enhanced cybersecurity in the industrial internet of things. IEEE COMSOC MMTC Commun. - Front. **18**(2), 2023 (2023)
12. Alturkistani, H., El-Affendi, M.A.: Optimizing cybersecurity incident response decisions using deep reinforcement learning. Inter. J. Electr. Comput. Eng. **12**(6), 6768 (2022)
13. Al, S., Dener, M.: Stl-hdl: a new hybrid network intrusion detection system for imbalanced dataset on big data environment. Comput. Sec. **110**, 102435 (2021)
14. Kim, T., Pak, W.: Real-time network intrusion detection using deferred decision and hybrid classifier. Futur. Gener. Comput. Syst. **132**, 51–66 (2022)
15. Smys, S., Basar, A., Wang, H., et al.: Hybrid intrusion detection system for internet of things (iot). J. ISMAC **2**(04), 190–199 (2020)
16. Narayana Rao, K., Venkata Rao, K., Prasad Reddy, P.V.G.D.: A hybrid intrusion detection system based on sparse autoencoder and deep neural network. Computer Commun. **180**, 77–88 (2021)
17. Kim, T., Pak, W.: Robust network intrusion detection system based on machine-learning with early classification. IEEE Access **10**, 10754–10767 (2022)
18. Subach, I., Mykytiuk, A.: Methodology of formation of fuzzy associative rules with weighted attributes from siem database for detection of cyber incidents in special information and communication systems. Inform. Technol. Sec. **11**(1), 20 (2023)
19. Husák, M., Bajtoš, T., Kašpar, J., Bou-Harb, E., Čeleda, P.: Predictive cyber situational awareness and personalized blacklisting: a sequential rule mining approach. ACM Trans. Manag. Inform. Syst. (TMIS) **11**(4), 1–16 (2020)
20. Sivanantham, S., Mohanraj, V., Suresh, Y., Senthilkumar, J.: Association rule mining frequent-pattern-based intrusion detection in network. Comput. Syst. Sci. Eng. **44**(2) (2023)
21. Lou, P., Lu, G., Jiang, X., Xiao, Z., Hu, J., Yan, J.: Cyber intrusion detection through association rule mining on multi-source logs. Appl. Intell. **51**, 4043–4057 (2021)
22. Patil, S., et al.: Explainable artificial intelligence for intrusion detection system. Electronics **11**(19), 3079 (2022)
23. Zhou, Q., Li, R., Xu, L., Nallanathan, A., Yang, J., Fu, A.: Towards explainable meta-learning for ddos detection,' arXiv preprint arXiv:2204.02255(2022)
24. Mahbooba, B., Timilsina, M., Sahal, R., Serrano, M.: Explainable artificial intelligence (xai) to enhance trust management in intrusion detection systems using decision tree model. Complexity **2021**, 1–11 (2021)
25. Karna, S.K., Paudel, P., Saud, R., Bhandari, M.: Explainable prediction of features contributing to intrusion detection using ml algorithms and lime
26. Kalutharage, C.S., Liu, X., Chrysoulas, C., Pitropakis, N., Papadopoulos, P.: Explainable ai-based ddos attack identification method for iot networks. Computers **12**(2), 32 (2023)
27. Coppolino, L., et al.: Detection of radio frequency interference in satellite ground segments. In: 2023 IEEE International Conference on Cyber Security and Resilience (CSR), pp. 648–653 (2023)
28. Hegland, M.: The apriori algorithm–a tutorial. Math. Comput. Imaging Sci. Inform. Process. 209–262 (2007)
29. Bowen, D., Ungar, L.: Generalized shap: Generating multiple types of explanations in machine learning, arXiv preprint arXiv:2006.07155 (2020)

30. Ribeiro, M.T., Singh, S., Guestrin, C.: Anchors: high-precision model-agnostic explanations. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32 (2018)
31. Sharafaldin, I., Lashkari, A.H., Ghorbani, A.A.: Toward generating a new intrusion detection dataset and intrusion traffic characterization. ICISSp **1**, 108–116 (2018)
32. Zolanvari, M., Teixeira, M.A., Gupta, L., Khan, K.M., Jain, R.: WUSTL-IIOT-2021 Dataset for IIoT Cybersecurity Research (October 2021). http://www.cse.wustl.edu/~jain/iiot2/index.html