Alert Correlation and Attack Scenario Recognition

Subjects: Computer Science, Interdisciplinary Applications

Contributor: Faeiz M. Alserhani

Planned and targeted attacks, such as the advanced persistent threat (APT), are highly sophisticated forms of attack. They involve numerous steps and are intended to remain within a system for an extended length of period before progressing to the next stage of action. Anticipating the next behaviors of attackers is a challenging and crucial task due to the stealthy nature of advanced attack scenarios, in addition to the possible high volumes of false positive alerts generated by different security tools such as intrusion detection systems (IDSs).

Keywords: IDS ; APT ; alert correlation ; centralized learning

1. Introduction

An intrusion detection system (IDS) refers to a hardware or software entity that monitors a network to detect and identify unauthorized activities or violations of established policies. In the event that an intrusion is detected, IDS captures pertinent data regarding the event, disseminates alerts, and executes appropriate remedial or preventive measures as deemed required [1]. The pervasive utilization of the internet by individuals has led to a significant escalation in network attacks, resulting in substantial detrimental consequences for both institutions and individuals alike. As a consequence of the heightened occurrence and enhanced complexity of cyberattacks, contemporary networked enterprises are compelled to adopt rigorous security protocols in order to ensure the integrity of their data transmissions. Consequently, IDS has become an indispensable component of any security framework [2][3].

IDS is a form of passive monitoring system utilized to detect potential threats, supply a comprehensive account of an ongoing or attempted theft, and initiate alerts that may be reviewed by analysts stationed in a security operations center (SOC) or incident responders. The majority of attacks are executed through the examination of network communications, wherein packets traversing the network are intercepted and subjected to analysis. This process involves the identification of heuristics and patterns, commonly referred to as signatures, which serve as means to detect and classify common attacks. Upon detection, operators are promptly alerted, thereby enabling appropriate action to be taken [4]. Network attacks often involve a series of concerted attempts to attack various parts of the system. Defending against these multi-stage attacks requires knowledge of the current attack state and the expected future attack phase. The sub-components in a web network are interlinked, and analyzing the logs of one device may reveal events triggered in another [5]. Advanced attack scenarios are a contemporary form of attack employed by adversaries, characterized by the utilization of sophisticated and elusive exploits and payloads. The primary objective of APTs is to establish prolonged persistence within a targeted system and subsequently move laterally to accomplish various goals, including but not limited to extensive data collection, disruption of operations, or denial of services [6]. In addition, it should be noted that anomaly-based detection methods have been found to produce an extensive amount of false positive results [7]. The concept of alert correlation (AC), also referred to as IDS post-processing, has been suggested as a means to address these constraints. The system that is able to predict full or part of attack scenarios can provide early and preventive measures to reduce the severity and threat caused by network attacks, hence adopting a proactive approach [8]. In the context of IDS models based on learning, one additional challenge is the insufficiency of datasets that adequately describe various advanced attacks and APT patterns, which is crucial for effectively training a robust detection model [9]. So, to protect enterprise networks from cyberattacks, IDS is continuously developed to solve various limitations, one of which is the production of a substantial quantity of alerts that are of low quality. Furthermore, a significant proportion of the alerts generated by IDSs are classified as false positives. It is essential to have a robust level of security to ensure transparent and reliable communication between parties to find malicious trends and help administrators fine-tune, organize, and deploy efficient controls.

Machine learning (ML) is a form of artificial intelligence methodology that has the capability to autonomously extract valuable insights from extensive datasets [10]. ML-based IDS can attain reasonable levels of detection performance given the availability of ample training data. Additionally, these IDS employ machine learning models that possess enough generalization capabilities to identify attack variations and novel attacks [11]. Deep learning (DL) is a subfield within the broader domain of machine learning that has demonstrated remarkable capabilities in achieving exceptional levels of performance. DL approaches have demonstrated superior performance in handling large datasets as compared to traditional machine learning techniques. Furthermore, DL techniques possess the capability to autonomously acquire feature representations from unprocessed data, afterwards generating outcomes encompassing multiple hidden learning layers. DL trains several neural nodes instead of one like linear regression in statistical learning [12].

In addition, the study examined the concepts of centralized and federated learning in order to assess the accuracy of detection while taking into account concerns regarding privacy, heterogeneity, and data availability. Federated learning (FL) is a collaborative ML learning approach that is implemented across various clients, ensuring that clients' personal data are not transferred to a central server provider and instead remain stored on the local client device. The conventional approach to neural network (NN) training involves the utilization of local datasets by all clients, which are then shared with a central server. On the other hand, distributed convolutional neural network training employs parallel training methods. The training of federated learning is operated under the assumption that local client datasets are separate and cannot be accessed or shared by others. The study is to provide a comparative analysis between an ensemble centralized model that combines three machine learning algorithms and a CNN_FL federated learning model. The evaluation process is to examine the detection rate in addition to the metrics typically used in such a design. In the FL approach, several local models using heterogeneous algorithms are trained in each device (clients), and the outcomes of these models are used to build and update a global model (server). The inputs of the global model are the results of the effective collaboration of data across multiple clients. This collection of various trained models contributes to the building of an intelligent IDS able to perform detection and prevention facilities with higher performance. Multistage attacks are modeled based on typical attack frameworks such as cyber kill chains [13].

2. Alert Correlation and Attack Scenario Recognition

The potential damage of multi-stage and advanced attacks necessitates the establishment of a comprehensive analysis to identify the steps conducted by intruders to perform such attacks. Detection of the links between different activities based on triggered alerts is crucial to create a global view of the attack in progress. This knowledge, whether it is complete or partial, can provide network administrators with valuable information in order to counter these attacks and to apply a mitigation procedure.

Rahman et al. [14] devised models for detecting advanced attack scenarios by employing a centralized approach. The NSL-KDD dataset is employed for evaluating the efficacy of federated architecture in IDS [7]. The study considers a range of practical scenarios and instances of intrusion attacks. Based on the empirical findings, a comprehensive evaluation is conducted to compare the FL, centralized, and self-learning methodologies. FL consistently demonstrated superior performance compared to the alternative methodologies in nearly all training iterations. The concept of “virtual reality” pertains to the procedure of developing programs that simulate a virtual reality experience. The intrusion detection capabilities of a federated network increase proportionally with its scale. In [15], an IDS in Wireless Edge Networks (WENs), combining GRU and SVM models under a custom FL algorithm, was proposed. They used Attention Mechanism to determine the significance of the uploaded model parameters. This is conducted with the goal of both measuring the global model's performance improvement and sorting the clients according to their importance. The authors [16] provided a strategy to improve the training effect by sharing a limited sample of data globally. This offers a scheme for the case of non-IID data in federated learning, which is an important consideration.

To detect distributed anomaly intrusion on an IoT-based industrial control system, a hybrid model that includes federated learning, autoencoder, transformer, and Fourier mixing sublayer was developed [17]. It delivered a high detection performance for time-series data while simultaneously solving the problem of anomaly detection on a minute-by-minute time scale with rapid learning. F. Wilkens et al. [18] used a kill chain state machine (KCSM) to detect complex attacks like advanced persistent threats (APTs) without having to spend more time analyzing large volumes of alerts. Their method generated scenario graphs from state machines by deriving potential attack stages from single and meta-alerts and modeling the resulting attack scenarios. The algorithm generates APT scenario graphs, which are graphical representations of the attack, with nodes representing involved hosts and edges representing infection activity.

The researchers in [17] proposed a model called MLAPT [19], which generates a correlation between alerts and the corresponding APT scenarios. Subsequently, ML models were employed to forecast APT events at their first stages, achieving a prediction accuracy of 84.8%. The authors [20] suggested a machine learning-based intrusion detection system that uses DT, RF, SVM, KNN, and DNN in both models of centralized and federated learning. A used Edge-IIoT set that had more than 10 different kinds of IoT devices was used to collect a dataset that categorized 15 attacks. These attacks were grouped into five threats, and some of their features were identified with high correlations. In centralized learning, the best DDN accuracy was found to be 94.67% for 15-class and 96.01% for 6-class. On the other hand, RF obtained 99.9% in binary classification. However, the federated learning results achieved were better; the accuracy results of the global model training with 10 clients and 10 iterations were 93.37%, 95.99%, and 100%, respectively, for 15-class, 6-class, and binary classification. M. Khosravi et al. [21] proposed a method for detecting (APTs) that relies on causal analysis and correlation between alerts. Multiple sensors' alarms are monitored over a lengthy period to determine which ones are most likely to be part of the APT attack's well-known IKC. Finally, it acceptably calculated the host infection score over all APT phases using a semi-real-world dataset and simulation.

A system for detecting APT attacks based on federated learning was proposed in [22] to differentiate various APT attack patterns. The global model is updated across multiple clients with various iterations. The malicious events collected as alerts are then fed to the correlation module to determine which alerts are most relevant to APT attack steps. Based on alert type-determined APT stages, an APT scenario indicated the probability of the change of the attack's step. With 400 iterations, their model applied to UNSW-NB15 datasets and synthetic datasets from five clients obtained 96.7% accuracy, which is higher than local models. The authors [23] proposed models for anomaly detection on two datasets, namely Contagio and CICIDS2017, using an unsupervised learning approach. Subsequently, the study will explore various known malware attacks targeting networks.

The authors [24] presented federated learning and CNN to detect abnormal IoT traffic without alert correlation. Mayfly optimization was used to minimize feature dimension, and the FL framework was then trained for each CNN local model for collaborative training without sharing private data. The Aposemat IoT-23 dataset detected anomalous IoT traffic with 97.73% accuracy using multi-class detection. The researchers [25] presented a federated learning framework to identify APT attacks in an SDN environment. They employed ML and DL techniques to categorize harmful indications. The researchers ran an experiment using the NF-UQ-NIDS dataset and models to showcase the viability of FL in addressing cyber threats while preserving privacy for data holders within the SDN environment. W. Giura et al. [26] proposed a model for APT detection that can be applied to general occurrences, expanding beyond the scope of IDS alerts. The attack stages are structured in a hierarchical pyramid, wherein the ultimate objective occupies the apex, while the preceding steps are organized into several strata. HTTP-based connections are considered more advantageous compared to alternative options for several reasons. Firstly, HTTP-based command and control (C&C) traffic is generally recognized as permissible within the majority of enterprise environments. Secondly, alternative C&C protocols like peer-to-peer (P2P) and Internet Relay Chat (IRC) exhibit distinctive network characteristics, such as specific ports and package content, which can be readily detected and obstructed [27]. The propagation of malware occurs through the utilization of custom encrypted partitions on removable media, as well as the exploitation of vulnerabilities within authentication protocols [28][29]. Kasongo et al. [30] suggested an ensemble model incorporating feature selection, specifically targeting the 19 most significant features out of the total 42 features available in the UNSW-NB15 dataset. The performance accuracy yielded a result of 75%. The study [31] presented an improved CNN architecture for the purpose of identifying malicious attack traffic, with a particular focus on zero-day attacks that have not been previously reported within the network. The binary classification findings of the study indicate that the model exhibited superior performance in detecting previously undetected instances of intrusion, as compared to the standard CNN model. The authors [32] conducted a study on intrusion anomaly detection, specifically examining different kernel functions within Support Vector Machines. They also utilized the Principal Component Analysis feature selection technique in their investigation. The datasets provided are the UNSW-NB15 datasets. The Gaussian kernel achieved the highest level of accuracy, measuring at 93.94% when applied to the UNSW-NB15 datasets.

References

1. Bhattacharya, S.; Maddikunta, P.K.; Kaluri, R.; Singh, S.; Gadekallu, T.R.; Alazab, M.; Tariq, U. A Novel PCA-Firefly Based XGBoost Classification Model for Intrusion Detection in Networks Using GPU. Electronics 2020, 9, 219.
2. Preuveneers, D.; Rimmer, V.; Tsingenopoulos, I.; Spooren, J.; Joosen, W.; Ilie-Zudor, E. Chained Anomaly Detection Models for Federated Learning: An Intrusion Detection Case Study. Appl. Sci. 2018, 8, 2663.
3. Bhatti, D.G.; Virparia, P.V. Soft Computing-Based Intrusion Detection System With Reduced False Positive Rate. In Design and Analysis of Security Protocol for Communication; Wiley: Hoboken, NJ, USA, 2020; pp. 109–139.
4. Anwar, S.; Mohamad Zain, J.; Zolkipli, M.F.; Inayat, Z.; Khan, S.; Anthony, B.; Chang, V. From Intrusion Detection to an Intrusion Response System: Fundamentals, Requirements, and Future Directions. Algorithms 2017, 10, 39.

5. Jadidi, Z.; Hagemann, J.; Quevedo, D. Multi-step attack detection in industrial control systems using causal analysis. Comput. Ind. 2022, 142, 103741.

6. Sharma, A.; Gupta, B.B.; Singh, A.K.; Saraswat, V.K. A novel approach for detection of APT malware using multi-dimensional hybrid Bayesian belief network. Int. J. Inf. Secur. 2023, 22, 119–135.

7. Manzoor, E.; Milajerdi, S.M.; Akoglu, L. Fast Memory-efficient Anomaly Detection in Streaming Heterogeneous Graphs. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1035–1044.

8. Ansari, M.S.; Bartos, V.; Lee, B. Shallow and Deep Learning Approaches for Network Intrusion Alert Prediction. Procedia Comput. Sci. 2020, 171, 644–653.

9. Zhang, J.; Zhao, Y.; Wang, J.; Chen, B. FedMEC: Improving Efficiency of Differentially Private Federated Learning via Mobile Edge Computing. Mob. Netw. Appl. 2020, 25, 2421–2433.

10. Michie, D.; Spiegelhalter, D.J.; Taylor, C.C. Machine Learning, Neurall and Statistical Classification; Ellis Horwood Series in Artificial Intelligence: New York, NY, USA, 1994; Volume 13.

11. Liu, H.; Lang, B. Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey. Appl. Sci. 2019, 9, 4396.

12. Dong, S.; Wang, P.; Abbas, K. A survey on deep learning and its applications. Comput. Sci. Rev. 2021, 40, 100379.

13. Martin, L. Cyber Kill Chain. 2014. Available online: http://cyber.lockheedmartin.com/ (accessed on 27 July 2023).

14. Rahman, S.A.; Tout, H.; Talhi, C.; Mourad, A. Internet of Things Intrusion Detection: Centralized, On-Device, or Federated Learning? IEEE Netw. 2020, 34, 310–317.

15. Chen, Z.; Lv, N.; Liu, P.; Fang, Y.; Chen, K.; Pan, W. Intrusion Detection for Wireless Edge Networks Based on Federated Learning. IEEE Access 2020, 8, 217463–217472.

16. Zhao, Y.; Li, M.; Lai, L.; Suda, N.; Civin, D.; Chandra, V. Federated learning with non-IID data. arXiv 2018, arXiv:1806.00582.

17. Truong, H.T.; Ta, B.P.; Le, Q.A.; Nguyen, D.M.; Le, C.T.; Nguyen, H.X.; Do, H.T.; Nguyen, H.T.; Tran, K.P. Light-weight federated learning-based anomaly detection for time-series data in industrial control systems. Comput. Ind. 2022, 140, 103692.

18. Wilkens, F.; Ortmann, F.; Haas, S.; Vallentin, M.; Fischer, M. Multi-Stage Attack Detection via Kill Chain State Machines. In Proceedings of the 3rd Workshop on Cyber-Security Arms Race, Virtual, 15 November 2021; pp. 13–24.

19. Ghafir, I.; Hammoudeh, M.; Prenosil, V.; Han, L.; Hegarty, R.; Rabie, K.; Aparicio-Navarro, F.J. Detection of advanced persistent threat using machine-learning correlation analysis. Future Gener. Comput. Syst. 2018, 89, 349–359.

20. Ferrag, M.A.; Friha, O.; Hamouda, D.; Maglaras, L.; Janicke, H. Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning. IEEE Access 2022, 10, 40281–40306.

21. Khosravi, M.; Ladani, B.T. Alerts Correlation and Causal Analysis for APT Based Cyber Attack Detection. IEEE Access 2020, 8, 162642–162656.

22. Li, Z.; Chen, J.; Zhang, J.; Cheng, X.; Chen, B. Detecting Advanced Persistent Threat in Edge Computing via Federated Learning. In Proceedings of the Security and Privacy in Digital Economy: First International Conference, SPDE 2020, Quzhou, China, 30 October–1 November 2020; Springer: Singapore, 2020; pp. 518–532.

23. Neuschmied, H.; Winter, M.; Stojanović, B.; Hofer-Schmitz, K.; Božić, J.; Kleb, U. APT-Attack Detection Based on Multi-Stage Autoencoders. Appl. Sci. 2022, 12, 6816.

24. Xia, Q.; Dong, S.; Peng, T. An Abnormal Traffic Detection Method for IoT Devices Based on Federated Learning and Depthwise Separable Convolutional Neural Networks. In Proceedings of the 2022 IEEE International Performance, Computing, and Communications Conference (IPCCC), Austin, TX, USA, 11–13 November 2022; pp. 352–359.

25. Thi, H.T.; Son, N.D.H.; Duy, P.T.; Pham, V.-H. Federated Learning-Based Cyber Threat Hunting for APT Attack Detection in SDN-Enabled Networks. In Proceedings of the 2022 21st International Symposium on Communications and Information Technologies (ISCIT), Xi’an, China, 27–30 September 2022; pp. 1–6.

26. Giura, P.; Wang, W. Using large scale distributed computing to unveil advanced persistent threats. Sci. J. 2012, 1, 93–105.

27. Wang, X.; Zheng, K.; Niu, X.; Wu, B.; Wu, C. Detection of command and control in advanced persistent threat based on independent access. In Proceedings of the 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 22–27 May 2016; pp. 1–6.

28. Lajevardi, A.M.; Amini, M. A semantic-based correlation approach for detecting hybrid and low-level APTs. Future Gener. Comput. Syst. 2019, 96, 64–88.

29. Yin, Y.; Jang-Jaccard, J.; Xu, W.; Singh, A.; Zhu, J.; Sabrina, F.; Kwak, J. IGRF-RFE: A hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset. J. Big Data 2023, 10, 15.

30. Kasongo, S.M.; Sun, Y. Performance Analysis of Intrusion Detection Systems Using a Feature Selection Method on the UNSW-NB15 Dataset. J. Big Data 2020, 7, 105.

31. Hairab, B.I.; Elsayed, M.S.; Jurcut, A.D.; Azer, M.A. Anomaly Detection Based on CNN and Regularization Techniques Against Zero-Day Attacks in IoT Networks. IEEE Access 2022, 10, 98427–98440.

32. Almaiah, M.A.; Almomani, O.; Alsaaidah, A.; Al-Otaibi, S.; Bani-Hani, N.; Hwaitat, A.K.; Al-Zahrani, A.; Lutfi, A.; Awad, A.B.; Aldhyani, T.H. Performance Investigation of Principal Component Analysis for Intrusion Detection System Using Different Support Vector Machine Kernels. Electronics 2022, 11, 3571.