Deep learning for predictive alerting and cyber-attack mitigation

***Abstract*—The successful security management of ICT systems and services is essential for an effective cyber security posture. The main objective is to minimize and control the damage caused by cyber-attacks and incidents, to provide effective response and recovery, and to invest efforts in preventing future cyber incidents. To achieve this objective, cyber threat intelligence (CTI) is widely applied, as it is considered a crucial mechanism to proactively defend against modern and dynamically evolving cyber threats and attacks. However, there are multiple challenges in the field of CTI, as there is an enormous amount of unstructured threats data in cyberspace that needs to be collected, classified, analyzed, and shared between states, organizations, or companies. Facing this challenge, data mining techniques and machine learning algorithms are essential for providing meaningful CTI information due to their ability to extract indistinct and hidden patterns in the data. Based on data mining techniques and machine learning algorithms' potential for successfully implementing cyber threat intelligence services, this paper develops an efficient predictive alerting model in a threat intelligence engine using the Deep Residual Network (DRN) model. Further, the main goal is to compare the performance of the DRN model with other machine learning models such as Sequential Rule Mining, IntruDTree, ScaleNet, etc. According to our experimental results, the DRN outperformed other tested machine learning models by achieving better results on parameters such as precision, recall, and F-measure.**

***Index Terms*—Cyber threat intelligence, Situational awareness system, Deep residual network, Fuzzy C-means clustering.**

I. Introduction

As we live in the 4th industrial revolution era, the focus and concern are not on accessing this hybrid world but rather securing the information systems and human interactions in cyberspace [1]. Alongside this innovation in digitalization, institutions, organizations, and companies face evolving cyber threats and complex attacks. In this context, cyber security has reached great attention because of the potential for huge economic losses, and the direct impact on the state's stability and security [2]. Security breaches have always existed - even before digitalization. However, the consequences of cyber security breaches prove to be more severe as information today spreads quickly, reaches a broader audience, may be very costly, and takes a longer recovery time. In the past, regular maintenance and updating of the antivirus and antispam software were sufficient to protect the sensitive and valuable information resources of organizations and companies. However, since APTs [3] are searching for new ways to exploit vulnerabilities, there has been a shift from standard antivirus protection to a defense based on risk assessment and cyber threat intelligence. To cope with this challenge, cyber threat intelligence is utilizing data mining techniques and machine learning algorithms to provide actionable information that can prevent cyber-attacks and can improve the cyber-security posture of an organization, company, or country [4]. In this context, data mining techniques focus on the extraction of meaningful and essential information from large datasets by analyzing and discovering invisible patterns and relationships of data for creating knowledge that can be used to predict, understand, and find anomalies and associations. In other words, it transforms the processed data into useful information, and knowledge [5], [6]. Today, cyber-security utilizes data mining to extract indistinct and hidden patterns in large data sets for several needs, like threat correlation and alert generation [7]. Moreover, the amount of data generated through security artifacts from divergent sensors has been growing exponentially in recent years, and this trend will probably continue in the future. Hence, cyber security has occupied the era of big data for managing huge quantities of data. Still, this huge volume of data must be processed to be used for effective cyber situational awareness [8].

Predictive alerting is used to predict future events by examining actual and historical data [15]. Predictive analytics models may be used to predict future events and behavior characteristics. Combining big data and predictive analytics for cyber defense enables the transformation of vast amounts of data into actionable intelligence. Predictive alerting is based on learning that creates data models and applies the models to detect threats. The main challenge is accuracy. Therefore, the research of algorithms for predictive alerting is an active domain. Furthermore, most predictive analytics methods provide a score where a higher score implies a greater possibility that an event will occur. A lower score suggests a decreased likelihood of the event occurring [15]. Utilizing such models and techniques, this paper evaluates the performance of a developed predictive alerting system employing the DRN method and comparing its performance with other artificial intelligence and machine learning methods, such as "Sequential Rule Mining," "IntruDTree," "Intelligent Intrusion Detection Model" and "ScaleNet."

The core contribution of this research paper is the design of a predictive alerting system using deep learning models. In this method, the input log file is normalized, and alert segregation is based on Fuzzy C-Means clustering (FCM) [18]. Then, the feature selection is performed from clustered sets for an effective predictive alert process. Further, a predictive alert is performed by the threat intelligence engine [19] using Deep Residual Network (DRN), and cyber-attack mitigation is done by blacklisting predicted results.

The paper is structured as follows. Next section provides an overview of related work in the field of intrusion detection, threat intelligenece gathering and alert prediction. The core contribution of this research paper, which is a designed DRN model for threat predictive alerting is presented in Section III and evaluation of experimental results by comparing them with other relevant machine learning methods is given in Section IV. In Section V, we discuss on the achieved results and possible improvements of the presented predictive alerting process based on the deep learning technique.

II. Related Work

The application of data mining to cyber security can increase the efficiency of detecting malicious activity due to the ability to extract information from previously stored unstructured data related to different types of incidents captured from monitoring systems and reporting mechanisms [22]. Moreover, sorting and correlating data can help cyber analysts take preventive measures and predict future attacks; therefore, the network can be proactively protected [23]. Data mining is generally classified into two main categories [24]: i) Descriptive data mining (information from data itself) includes cluster analysis and association rules. ii) Predictive data mining (information extracted from previous data) comprises classification and regression models.

Intrusion can be defined as a set of plans and actions taken to threaten or attack "Confidentiality," "Integrity," or "Availability" of a computer network or system. Therefore, intrusion detection aims to discover these threats and attacks on computer systems and networks by observing multiple activities or attributes. Furthermore, intrusion detection is one of the most critical components of network security. Traditionally, this job was done manually, but with the advancement of software systems based on data mining, this process has evolved from manual to automatic analysis systems. Generally, intrusion detection techniques are classified into two categories [25]:

- Misuse detection, also known as a rule-based approach, includes network traffic monitoring to capture matches of learned patterns of attacks and their signatures. If a pattern match is found, it triggers an event and raises the alarm, alerting the security analyst to take action. According to the size of the networks, these alarms can generate up to millions of alarms per day [26]. The main disadvantage of this approach is that it follows only predefined patterns. Therefore it cannot detect new or previously unknown threats and attacks. The main approaches for misuse detection include expert systems, signature analysis, state-transition analysis, and data mining [25].
- Anomaly Detection intends to fix the disadvantage of misuse detection by detecting attacks with undefined signatures. This approach includes building models of regular data and detecting deviations from these models of observed data. Furthermore, the advantage of this algorithm is the ability to detect novel and undefined threats or attacks by scanning the deviation from normal data. A disadvantage of this technique is generating a high percentage of false positives [26]. The main approaches for anomaly detection include statistical methods, expert systems, and data mining.

As elaborated above, data mining is crucial for a practical cyber threat intelligence and situational awareness system. Husák et al. [27] devised a Sequential Rule Mining approach for predictive cyber situational awareness and a personalized blacklisting process. Here, sequential rule mining was implemented for predicting security events, and it was utilized to generate a predictive blacklist. This approach improved the success rate but failed to train machine learning techniques for automatically choosing the good rules. Sarker et al. [28] presented IntruDTree for automatically selecting good rules for the cyber security intrusion detection process. In this model, the ordering of security features is done based on their significance as well as a tree-driven intrusion detection system was formulated through selected features. The computational cost of this technique is satisfactory, although this approach lacks generalization. Al-Omari et al. [11] devised an intelligent intrusion recognition scheme for the intrusion detection process in cyber security to improve generalization capability. A Decision Tree (DT) was considered for ordering security features. This method effectively reduces the computational effort needed, although it failed to include feature filtering and a wrapping scheme for better performance. Vinayakumar et al. [8] developed DeepDGANet for intrusion detection to increase system performance. In this method, a Domain Generation Algorithm (DGA) was implemented and tested using bots to generate domain names periodically. Even though it was not tested on an adversarial environment, this method provides detailed information about identified malware. The Sequential Rule Mining Approach was proposed to provide predictive cyber situational alerting as adaptive blacklisting [27]. However, this approach failed to support threat intelligence and an alert association by detecting subgroups of alerts suitable for specific investigations. The intrusion detection tree, IntruDTree [28], the machine learning security method, was developed. However, it failed to utilize large datasets in IoT infrastructures and evaluate their efficiency at the application level in cybersecurity. An intelligent intrusion detection model for cyber security was introduced in [11]. However, this technique failed to predict cyber threat types and was not evaluated with other security structures. The Consortium Blockchain-based "DefenseChain" platform was designed [8] for cyber threat intelligence allocation and defense process; however, this model failed to identify different real-life scenarios in which distributed trust values might be used to authorize the definition of threats to protect data access in contrast to targeted cyber-attacks.

In general, methods of predictive blacklisting demonstrated that it is also beneficial to predict the future behavior of previously identified malicious sources. These methods are especially effective when combined with other alerts sharing

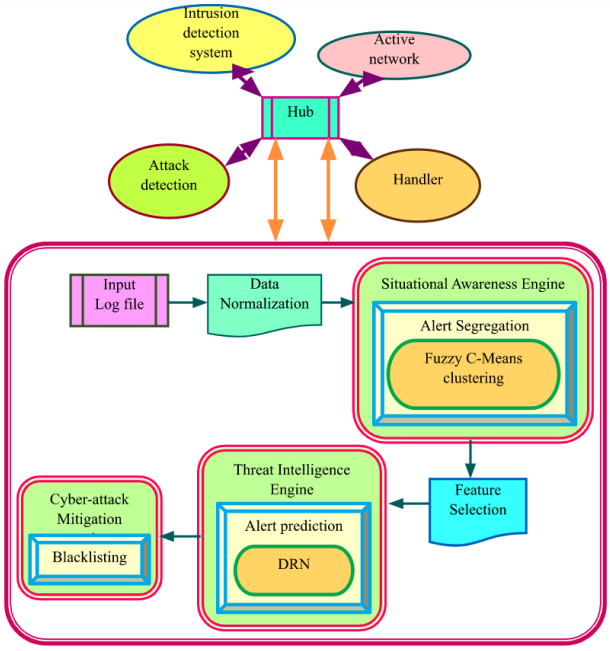


Fig. 1: Processing Pipeline of the Predictive Alerting System

mechanisms. These have grown in popularity in recent years because they allow comparing attacks on multiple targets to predict future ones. Furthermore, this is supported by the author's experience with multi-organization alerts sharing communities at nation-level networks. As a result, the proposed DRN model aims to support organizations by predicting alerts in threat intelligence engines and providing cyber-attack mitigations through blacklisting IP addresses.

III. Predictive Alerting

This section presents the paper's main contribution, which is a predictive alerting system. The purpose of the system is to provide predictions of alerts, which the ML-based model learned from the data processed. The predictions are actionable for taking measures against the threats expected. We propose the Deep Residual Network Predictive Alerting (DRNPA) plugins be integrated with the SABU platform [31]. The SABU platform was built for exchanging IDS alerts, honeypot logs, and third-party security data. Security teams process the alerts as a part of the incident-handling process. Often, this yields a generation of new rules for security devices. The DRNPA aims to aid in this process by providing automated analysis of the input data using the learned model to identify possible threats (alerts). The predicted alerts are further employed in the attack mitigation phase. Currently, the new items for blacklists are generated based on predicted alerts.

We developed two new methods for i) alert segregation using Fuzzy C-means clustering and ii) alert prediction employing DRN as plugins for SABU. Technically, the SABU system as depicted in Fig. 1 includes the following processing phases:

1) Input Normalization: The input alert log files are preprocessed to unify their format as they can come from different sources, e.g., IDS, network telemetry, threat intelligence data, etc. The data are processed by the SABU and provided in CSV format. Each alert record r consists of k-fields, which include common information, e.g., the IP address, timestamp, source identification, and other alert specific data entries: $r = \langle x_1, \ldots, x_n \rangle$. In this paper, we use a dataset that consists of alerts defined by fields as shown in I. The SABU platform also performs data normalization. Normalization aims to turn characteristics into a comparable scale, aiming to boost the model's performance and training stability. Herein, the original data is converted using Z-Score Normalization [21].

2) Situational Awareness Engine: The alert segregation is provided as a plugin to situational awareness engine. Novel alert segregation method based on the Fuzzy C-Means clustering [18] was developed. It processes the stream of alerts, keeps the vectors of their main features in a predefined time window, and matches the alerts with similar features according to a set of rules. For grouping the alerts, the Fuzzy C-Means algorithm is used. The characteristic of Fuzzy clustering is that each data point can belong to more than one cluster, which is suitable for alert segregation because feature selection is performed based on the most significant features from every cluster. The output from alert segregation is thus a set of clusters. Each cluster represents a group of events sharing significant characteristics, e.g., events corresponding to the same attack technique.

3) Feature selection: After the completion of alert segregation, a feature selection process is performed to select the significant features for every cluster. The most important features are selected in each cluster using Hellinger distance, where features with values under the given threshold is selected and used for further processing in the Threat Intelligence Engine. The total number of input features is 23 as presented in Table I.

4) Threat Intelligence Engine: The intrusion predictive alerting is executed based on DRN [19] at the threat intelligence engine. After the accomplishment of feature selection, predictive alerting is performed in the threat intelligence engine. The predictive alerting is carried out using a Deep Residual Network – DRN, with the preprocessed selected features serving as the input to the DRN. The training efficiency and generalization ability are highly improved in DRN, thereby it is used for a predictive alerting process. The DRN method employs a variety of layers, including pooling, convolutional (Conv), linear classifiers, residual blocks, and linear classifiers. Under conditions with insufficient amounts of training data, DRN usually improves both learning and training performance. The DRN is used in this research to predict alerts successfully. The parameters used by the DRN are given in Table II.

5) Cyber-attack Mitigation Engine: the mitigation is done by blacklisting suspicious IP addresses based on the predictive alerting outcomes. Finally, the predicted output obtained from DRN is utilized to perform cyber-attack mitigation by blacklisting IP addresses based on the predicted results, and

ID	Name	Description
1	ip	The IP address of the event.
2	tor	Is the event related to Tor communication?
3	blocklist_de_ssh	Is the IP address in SSH blocklist?
4	uceprotect	Is the IP address in a DNS blacklist?
5	sorbs-dul	Is the IP address in a blacklist of dial-ip ranges?
6	sorbs-noserver	Is the IP address in DNS-based Block List (DNSBL) maintained by Sorbs?
7	sorbs-spam	Is the IP address recognized by Sorbs Spam Block List?
8	spamcop	Is the IP address in SpamCop Blocking List?
9	spamhaus-pbl	Is the IP address part of DNSBL database of end-user IP address ranges suspected of sending SPAM emails?
10	spamhaus-pbl-isp	Is the IP addres in the DNSBL list?
11	spamhaus-sbl-cbl	Is the IP address in the Spamhaus Exploits Block List?
12	hostname_exists	Does hostname exists?
13	dynamic_static	Static or dynamic IP address.
14	dsl	Is the host connected via DSL?
15	vpn	Does the host use a VPN to communicate?
16	nat	Does the host use NAT for Internet access?
17	ip_in_hostname	Is the IP address part of the hostname?
18	censys_protocols	List of protocol of the connection identified by Protocol scanning service (PSS).
19	censys_tags	List of Tags generated by PSS.
20	censys_device_type	The identified type of the device.
21	censys_product	The identified product of the host.
22	centsys_os	The identified OS running on the host device.
23	censys_os_ver	the identified version of operating system.

TABLE I: SABU generated fields for log records

the results are interpreted into a CSV file.

The execution of the proposed analysis of the intrusion detection alerts framework on SABU is performed using Python in the SABU alert sharing platform. Moreover, the implementation of the proposed intrusion detection alert structure is performed using a Dataset of intrusion detection alerts from a sharing platform [30].

IV. Experiments

The experiments of the proposed DRN-based predictive alerting are presented in this section. The proposed predictive alerting system was implemented in the PYTHON programming language and consumed data from the SABU alert-sharing platform.

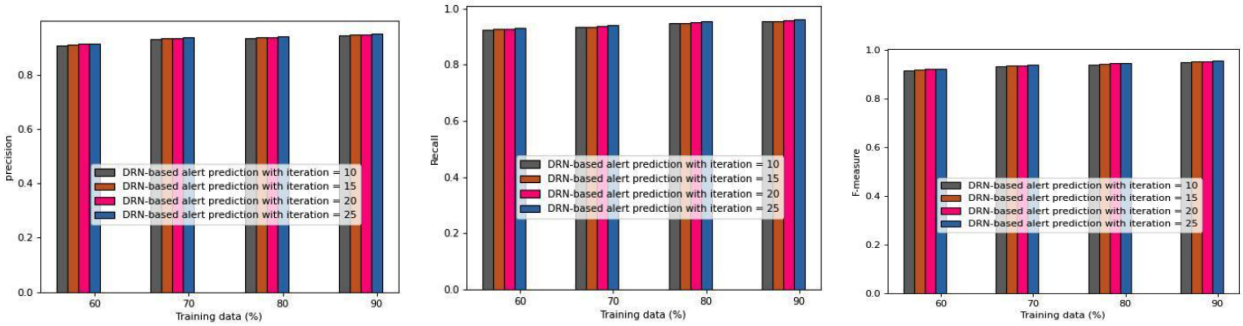
The dataset of intrusion detection alerts [30] from the sharing platform is used to run the DRN-based predictive alerting algorithm. This database includes the main file, intrusion detection alerts, and four auxiliary records with enhanced data. Alerts were collected from the SABU alert-sharing platform for one week and accumulated in the IDEA setup. Nearly 12 million alerts were generated by 34 intrusion detection systems, other data sources, and honeypots distributed across three organizations. The hostname, Uniform Resource Locators (URLs), IP addresses, and other identifiers in alerts are anonymized, but the information in auxiliary files allows for malicious actors to be reported.

The performance metrics, namely precision, recall, and F-measure, are used to evaluate the DRN-driven predictive alerting system. Figure 2 depicts DRN-driven predictive alerting performance metrics for training data in different iterations. Figure 2a demonstrates performance analysis of DRN-based predictive alerting with precision metric. The precision for DRN driven predictive alerting model with iterations 10, 15, 20, and 25 is 0.9446, 0.9465, 0.9480, and 0.9500 for 90% of data used for training. The performance analysis of devised DRN-based predictive alerting for recall metric is shown in Figure 2b. The recall of DRN-based predictive alerting with iteration 10 is 0.9546, 15 is 0.9560, 20 is 0.9586, and 25 is 0.9607. Figure 2c outlines the analysis of DRN-based predictive alerting for the F-measure metric. The F-measure of DRN-based predictive alerting is 0.9401, 0.9425, 0.9445, and 0.9467, while iteration is 10, 15, 20, and 25.

Parameter	Value
Batch size	128
No. Filters	16
Kernel size	3
Activation Function	ReLU

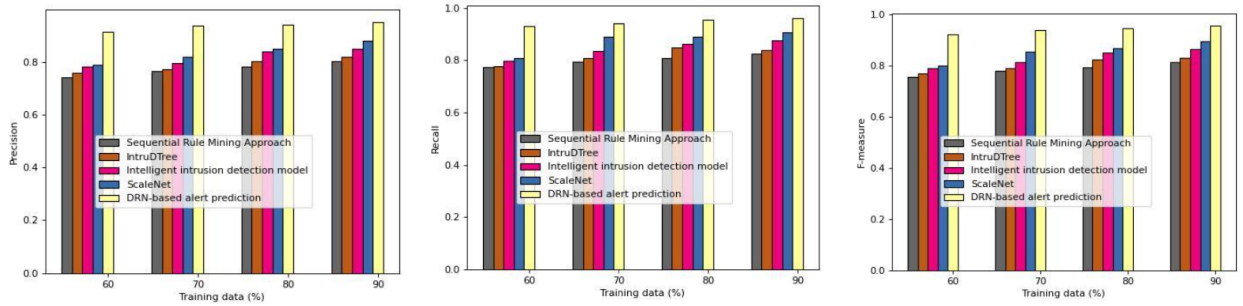
TABLE II: DRN Parameters

The existing predictive alerting approaches, such as sequential rule mining scheme [27], IntruDTree [28], intelligent intrusion detection method [11], and ScaleNet [8] were considered for comparing the performance of proposed DRN-based predictive alerting techniques. Figure 3 compares the DRN-based predictive alerting model in terms of various performance metrics through altering training data. The comparative analysis of the DRN-based predictive alerting model for the precision metric is in Figure 3a. The DRN model achieved better precision of 0.9405, whereas existing methods reached 0.7811, 0.8018, 0.8376, and 0.8479 when 80% of data were used for training. The performance enhancement of the proposed approach is 16.94%, 14.74%, 10.94%, and 9.84%, better than the existing methods. Figure 3b compares the DRN-driven predictive alerting approach regarding recall metric. The recall of sequential rule mining scheme is 0.8086, IntruDTree is 0.8474, intelligent intrusion detection technique is 0.8628, ScaleNet is 0.8911, and developed DRN-based predictive alerting is 0.9529. At the same time, training data is 80%, and performance improvement achieved by the designed



(a) Precision Metric (b) Recall Metric (c) F-Measure

Fig. 2: Predictive alerting performance metrics



(a) Precision Metric (b) Recall Metric (c) F-Measure

Fig. 3: Predictive alerting performance metrics

Comparative analysis	Sequential Rule mining model	IntruDTree	Intelligent intrusion detection model	ScaleNet	DRN
Precision	0.8032	0.8187	0.8509	0.8801	**0.9500**
Recall	0.8248	0.8401	0.8747	0.9076	**0.9607**
F-measure	0.8139	0.8292	0.8626	0.8937	**0.9553**

TABLE III: Performance comparison to related methods

model is 15.14%, 11.07%, 9.45%, and 6.48%. The comparative analysis of the designed DRN-based predictive alerting approach for the F-measure metric is outlined in Figure 3c. The F-measure of existing and developed predictive alerting methods are 0.7946, 0.8240, 0.8500, 0.8690, and 0.9467 in 80% of training data, whereas performance enhancement of the proposed approach is 16.06%, 12.96%, 10.20%, and 8.20%.

We compared the results obtained by the proposed model with the conventional techniques by varying the training data percentage from 60% to 90%. The best-achieved results are in Table III. The precision of the sequential rule mining scheme is 0.8032, IntruDTree is 0.8187, intelligent intrusion detection technique is 0.8509, ScaleNet is 0.8801, and developed DRN-based predictive alerting is 0.9500 in case the training data forms 90% of all data. The precision metric of the predictive alerting technique is increased because of the normalization process. The recall metric of existing and developed predictive alerting methods are 0.8248, 0.8401, 0.8747, 0.9076, and 0.9607 in 90% of training data. Due to the utilization of the DRN model, the recall of the designed predictive alerting approach is highly improved. The DRN model also obtained a better F-measure metric of 0.9553, whereas existing methods achieved 0.8139, 0.8292, 0.8626, and 0.8937 for 90% of training data. The clustering of normalized data based on FCM schemes efficiently enhances the F-measure of a developed DRN-based predictive alerting system.

V. Conclusion

This paper presents an efficient predictive alerting system employing the DRN method. This devised intrusion detection alert model is executed in a dataset of intrusion detection alerts from the SABU sharing platform. The proposed DRN model successfully enhances the system performance and generalization ability with less processing time by using the ReLU activation function, which is much faster than tanh or sigmoid. In addition, the utilized FCM for the clustering process efficiently improves the predictive alerting process. Furthermore, the performance of established DRN-based predictive alerting models is evaluated with other existing predictive alerting approaches through altering training data. In addition, the performance of the DRN-based predictive alerting method is computed with three performance metrics, namely precision, recall, and F-measure. Thus, the DRN-driven predictive alerting system

achieved better performance than other existing techniques with a precision of 0.9500, a recall of 0.9607, and an F-measure of 0.9553. However, this approach can be further improved by including more datasets for training to enhance system performance.

REFERENCES

[1] Xu, Min, Jeanne M. David and Suk-Hi Kim. "The Fourth Industrial Revolution: Opportunities and Challenges." International Journal of Financial Research 9 (2018): 90-95.

[2] Schwab, Klaus. "The Fourth Industrial Revolution". London, England: Portfolio Penguin. 2017

[3] Choi, J., Choi, C., Lynn, H.M. and Kim, P., "Ontology based APT attack behavior analysis in cloud computing", In proceedings of 2015 10th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA), pp.375-379, November 2015.

[4] Dharamkar, B. and Singh, R.R., "A review of cyber attack classification technique based on data mining and neural network approach", Int. J. Comput. Trends Technol, vol.7, no.2, pp.100-105, January 2014.

[5] Katoua, H. S. (September 2013). "Exploiting the Data Mining Methodology for Cyber Security". Egyptian Computer Science Journal Vol. 37 No. 6, 44-52.

[6] Jiawei Han, Micheline Kamber, Jian Pei. (June 2011). "Data Mining: Concepts and Techniques". San Francisco, CA, USA: Morgan Kaufmann Inc.

[7] Buczak, A.L. and Guven, E., "A survey of data mining and machine learning methods for cyber security intrusion detection", IEEE Communications surveys & tutorials, vol.18, no.2, pp.1153-1176, October 2014.

[8] Lee, W. and Stolfo, S.J., "A framework for constructing features and models for intrusion detection systems", ACM transactions on Information and system security (TiSSEC), vol.3, no.4, pp.227-261, November 2000.

[9] Singh, S. and Silakari, S., "An ensemble approach for cyber attack detection system: a generic framework", In proceedings of 2013 14th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, pp.79-84, July 2013.

[10] Vinayakumar, R., Soman, K.P., Poornachandran, P., Mohan, V.S. and Kumar, A.D., "ScaleNet: scalable and hybrid framework for cyber threat situational awareness based on DNS, URL, and email data analysis", Journal of Cyber Security and Mobility, vol.8, no.2, pp.189-240, April 2019.

[11] Al-Omari, M., Rawashdeh, M., Qutaishat, F., Mohammad, A.H. and Ababneh, N., "An Intelligent Tree-Based Intrusion Detection Model for Cyber Security", Journal of Network and Systems Management, vol.29, no.2, pp.1-18, April 2021.

[12] Mohammadi, S., Mirvaziri, H., Ghazizadeh-Ahsaee, M. and Karimipour, H., "Cyber intrusion detection by combined feature selection algorithm", Journal of information security and applications, vol.44, pp.80-88, February 2019.

[13] Tsai, C.F., Hsu, Y.F., Lin, C.Y. and Lin, W.Y., "Intrusion detection by machine learning: A review", expert systems with applications, vol.36, no.10, pp.11994-12000, December 2009.

[14] Milenkoski, A., Vieira, M., Kounev, S., Avritzer, A. and Payne, B.D., "Evaluating computer intrusion detection systems: A survey of common practices", ACM Computing Surveys (CSUR), vol.48, no.1, pp.1-41, September 2015.

[15] Vaibhav Kumar and M L Garg. "Predictive Analytics: A Review of Trends and Techniques". International Journal of Computer Applications 182(1):31-37, July 2018.

[16] Wu, Peilun, Hui Guo, and Nour Moustafa. "Pelican: A deep residual network for network intrusion detection." 50th annual IEEE/IFIP international conference on dependable systems and networks workshops (DSN-W), pp. 55-62. IEEE, 2020.

[17] Zaeemzadeh, Alireza Rahnavard, Nazanin & Shah, Mubarak. "Norm-Preservation: Why Residual Networks Can Become Extremely Deep?". IEEE Transactions on Pattern Analysis and Machine Intelligence. PP. 1-1. 10.1109/TPAMI.2020.2990339, May 2018.

[18] Xie, L., Wang, Y., Chen, L. and Yue, G., "An anomaly detection method based on fuzzy c-means clustering algorithm", In Second International Symposium on Networking and Network Security (ISNNS, 10) Jinggangshan, PR China, pp.89-92, April 2010.

[19] Chen, Z., Chen, Y., Wu, L., Cheng, S. and Lin, P., "Deep residual network-based fault detection and diagnosis of photovoltaic arrays using current-voltage curves and ambient conditions", Energy Conversion and Management, vol.198, pp.111793, October 2019.

[20] Purohit, S., Calyam, P., Wang, S., Yempalla, R. and Varghese, J., "DefenseChain: Consortium Blockchain for Cyber Threat Intelligence Sharing and Defense", In proceedings of 2nd Conference on Blockchain Research Applications for Innovative Networks and Services (BRAINS), pp.112-119, September 2020.

[21] Fei, N., Gao, Y., Lu, Z. and Xiang, T., "Z-score normalization, hubness, and few-shot learning", In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp.142-151, 2021.

[22] Tianfield, H., "Data mining based cyber-attack detection", System simulation technology, vol.13, no.2, pp.90-104, April 2017.

[23] Mehrnoosh Monshizadeh, Zheng Yan. (2014). "Security Related Data Mining". IEEE International Conference on Computer and Information Technology (pp. 775-782). Xi'an, China: IEEE.

[24] Daniel Barbara, Sushil Jajodia. (2001). "Applications of Data Mining in Computer Security". London: Kluwer Academic Publishers.

[25] Jiawei Han, Micheline Kamber, Jian Pei. (June 2011). "Data Mining: Concepts and Techniques". San Francisco, CA, USA: Morgan Kaufmann Inc.

[26] Sun, N., Zhang, J., Rimba, P., Gao, S., Zhang, L.Y. and Xiang, Y., "Data-driven cybersecurity incident prediction: A survey", IEEE communications surveys & tutorials, vol.21, no.2, pp.1744-1772, December 2018.

[27] Husák, M., Bajtoš, T., Kašpar, J., Bou-Harb, E. and Čeleda, P., "Predictive cyber situational awareness and personalized blacklisting: a sequential rule mining approach", ACM Transactions on Management Information Systems (TMIS), vol.11, no.4, pp.1-16, September 2020.

[28] Sarker, I.H., Abushark, Y.B., Alsolami, F. and Khan, A.I., "Intrudtree: a machine learning based cyber security intrusion detection model", Symmetry, vol.12, no.5, pp.754, may 2020.

[29] Oosterhoff, J., van Zwet, W.R. (2012). A Note on Contiguity and Hellinger Distance. In: van de Geer, S., Wegkamp, M. (eds) Selected Works of Willem van Zwet. Selected Works in Probability and Statistics. Springer, New York, NY.

[30] Martin Husak, Martin Zadnik, Václav Bartos, and Pavol Sokol. 2019. Dataset of intrusion detection alerts from a sharing platform. Retrieved November 15, 2019.

[31] CESNET and Masaryk University. 2016. SABU. Retrieved November 15, 2019 from https://sabu.cesnet.cz/en/start