

Cyber security: performance analysis and challenges for cyber attacks detection

Azar Abid Salih¹, Maiwan Bahjat Abdulrazzaq²

¹Department of Information Technology Management, Technical College of Administration, Duhok Polytechnic University, Duhok, Iraq

²Department of Computer Science, Faculty of Science, University of Zakho, Duhok, Iraq

Article Info

Article history:

Received Nov 13, 2022

Revised Jun 1, 2023

Accepted Jun 3, 2023

Keywords:

Attack detection

Cyber-attack

Cybersecurity

Deep learning

Machine learning

ABSTRACT

Nowadays, with the occurrence of new attacks and raised challenges have been facing the security of computer systems. Cyber security techniques have become essential for information technology services to detect and react against cyber-attacks. The strategy of cyber security enables visibility of various types of attacks and vulnerabilities throughout computer networks, whilst also provides detecting cyber-attacks and effective ways of identifying and preventing them. This study mainly focuses on the performance analysis and challenges faced by cyber security using the latest techniques. It also provides a review of the attack detection process including the robust effectiveness of intelligent techniques. Finally, summarize and discuss some methods to increase attack detection performance utilizing deep learning (DL) architectures.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Azar Abid Salih

Department of Information Technology Management, Technical College of Administration

Duhok Polytechnic University

Duhok, Kurdistan Region, Iraq

Email: azar.abid@dpu.edu.krd

1. INTRODUCTION

As information and communication technologies advance, novel challenges emerge that need to be addressed in system security. Because of advancements in cyber security technology, attackers are considering new and advanced methods of attacking networks [1]. The aim of network security is to defend and prevent unauthorized access to computers, networks, programs, and information [2]. Modern society is becoming increasingly interconnected via the internet. Therefore, cybersecurity is an essential consideration for information technology as well as internet services [3]. The term “cyber security” refers to high-tech information security that includes strategies for protecting networks, devices, software, and information from unauthorized access [4]. The integrity security system of cyber security refers to a set of safety procedures that can be used to defend cyberspace against attacks [5]. The primary goal of a cyber-defense system is to keep data confidential, integral, and available [6]. With the widespread use of networks and the emergence of new types of cyber threats, network security is becoming increasingly important [7]. The main challenge in the area of network security is identifying different types of network attacks, especially those that have never been observed before, and updating behavior on a regular basis [8]. Threats and attacks come in a variety of forms, including active and passive attacks, external and internal attacks, known and unknown attacks, and spam [9]. In general, security attacks are divided into two main kinds: first, active attacks that try to alter the regular functionality of a network, such as denial-of-service (DOS), which is the riskiest attack that involves numerous attackers and overwhelms a system with a large amount of incoming requests. The second, which is a passive attack, happens when the attackers try to decode and analyze the information exchanges to get significant information about the legitimate transmitter-receiver nodes [10]. Intelligent techniques are contributing for detecting and analyzing performance various types of attacks in the of cyber space [11]. Network security is

based on machine learning (ML) algorithms capable of processing big datasets, feature learning, feature extracting, classifying, predicting, and identifying attacks [12]. However, traditional ML algorithms are not capable of providing distinctive feature descriptors to address the problem of attack detection, due to model complexity limits [13]. ML has made a significant breakthrough with the structure of neural networks, which are referred to as DL because of their architecture of deep layers for solving big amounts of data [14].

The defense process of cyberspace against cyber-attacks faces numerous challenges that affect the performance of detection attacks [15]. The main obstacles point which are complexity of attacks, new features of attacks, the frequent updates of attacks by covering themselves in regular time [16]. In each detection attack strategy, several tools and techniques are grouped together based on their functionalities, such as system-level monitors, and network traffic analyzers. In data collection, the amount of raw data captured can be used to convert any packet details into understandable ones by using software defined networks (SDNs) [17]. The modern programming environment for research on cyber-attacks represents a useful tool for training and testing to detect vulnerabilities and attacks [18]. The python programming language with most important packages of libraries is used to offer a cybersecurity solution for network traffic issues [19], [20]. Furthermore, the different statistical tools used in the data preparation stage for analyzing data, such as the orange and R reduce the complexity of the data set and find relations between features [21], [22].

This paper aims to highlight the challenges faced in the field of cybersecurity in detecting and identifying cyberattacks using various intelligent techniques, DL and ML. The rest of this paper is structured as follows: In section 2, overviews of cybersecurity, section 3 relevant concepts, cyber-attack detection, and section 4 cyber security challenges are illustrated. The literature review in detail is shown in section 5, and the discussion and comparison are presented in section 6. The section 7 is about assessment and recommendations. Finally, section 8 presents the conclusion.

2. CYBER SECURITY

Cyber security, or information technology security, is the protection and measures adopted to defend computer systems, networks, software, hardware, and information from illegal access [23]. The field of cyber security is becoming increasingly significant due to the increased reliance on computer systems, smart devices, internet service, wireless networks, intelligent applications, and numerous devices that constitute the internet of things (IoT), including having a defensive strategy against malicious attacks [24]. Computer systems around the world need systems based on the detection and predictive analysis of cyber threats performance depending on building intelligent model. The model design of cyber security is constructed on integral, confidential, and available information [25]. The purpose of cyber security is to protect cyber space such as personal information, e-government data, e-marketing websites, e-business reports, and various applications from unauthorized penetration [26]. In addition, it covers the security of software and equipment that ensure and guarantee the privacy and integrity of the data under protection from a variety of threats and attacks [27]. The possible dangers of computer security that are strenuous or hard to detect and it is a security violation are referred to as threats, and any attempt to commit any violation is referred to as an attack [28].

2.1. Cyber security applications

In order to defend against various kinds of cyberattacks, numerous organizations and projects have been established. Cyber security applications are employed in many different fields including industry, financial institutions, enterprise, government, and businesses, to collect and store confidential information on computers and transmit it over networks to other computers [29]. The key domains of cybersecurity vary depending on the digital infrastructure that has to be protected. Such as: information security, network security, IoT security, information communication security, database security, cloud security, application security, server security and mobile security are all terms that can be used to describe different types of security [30].

2.2. Intelligent techniques for cyber security

Currently, the trend in cyber security popularity relates to learning algorithms, which have been increasing [31]. DL and ML are advantageous for developing security models because they are more accurate, particularly in the feature learning process when dealing with enormous of data [32]. The DL networks need a massive amount of data to feed the network for the building model. The datasets are based on data gathered from various network traffic resources. It represents a set of data records containing a variety of features and pertinent data, from which the cybersecurity model originates [33]. Many datasets exist in the field of cybersecurity, including malware analysis, intrusion detection, DDoS attack detection and prediction, IoT attack prediction, and spam analysis, all of which are different cases in cyber security used for different purposes by utilizing learning algorithms [34]. In general, attack detection can be achieved using ML,

evolutionary algorithms, intelligent algorithms, DL, as well as data mining techniques. The majority of attack prevention strategies use traffic analysis to identify and prevent malicious activity [35].

2.3. Cyber attacks

A cyber-attack is any attempt to control computer systems or networks without permission with the intent to destroy, disable, or steal data held within these systems. These attacks may target any internet-connected device or person. Users individually, crucial public services, big businesses, governments, or even whole countries. In order to defend against cyber-attack, it's essential to recognize and analyze the variety of cyber-attack scenarios and methods that are used nowadays. The key to efficient cyber security is tracking updating and growing cyberattacks in cyberspace [36]. The section that follows goes through some of the most recent and common cyber-attacks seen today.

2.3.1. Malware

It is a type of malicious software, known as malcode, a type of damaging computer code or web script with the purpose of gaining access or causing harm to a computer, client, server or network [37]. Malwares are any codes that can be generated by an attacker or team in order to steal data or circumvent access controls in order to compromise a system [38]. The different kinds of malware include viruses, worms, trojan horses, ransomware, time bombs, polymorphic, rootkits, adware, spyware, and keyloggers, but the main categories are viruses, worms, and trojans, each of which has its unique characteristics and traits [39].

2.3.2. Phishing

The phishing attack is regarded as one of the most harmful attacks. Phishing, also known as brand spoofing, is a process of gaining access to personal data in order to disrupt or misuse it by showing itself as a legitimate user [40]. Usually, sending them via phishing emails. These emails can collect fake webpages and capture the details of personal informations. Phishing attacks comes in variety of forms, such as malicious websites, phishing over email, and phishing over the phone [41].

2.3.3. Distributed denial of service

Refers to an attacker who makes several requests to shut down a system or network on a single target [42]. The DDOS build-up DoS attack, in which numerous attackers are compromised, will target a single victim and produce a denial of service for the user of the targeted resource. DDoS attacks are commonly divided into three categories: application layer threats, protocol threats, and volume-based threats [43]. DDOS attacks are the most dangerous attacks in this decade because they do not necessitate a penetration of the whole network, just finding a gap to disturb the computer system [44].

2.3.4. SQL injection

The process of passing SQL code into interactive web applications cyber-attack involves using code to enter databases and steal information [45]. An attempt to issue SQL commands to a database via a website interface is known as a SQL injection attack. Due to the availability of user-input fields, all of these website features are vulnerable to SQL injection attacks. The SQL injection attacks take advantages of the power of code for damaging purposes, typically by intrusion into the back end of an application [46].

2.3.5. Zero-day

It is identified as a novel attack. Zero-day attacks refer to the harm to or theft of data from a system affected by a vulnerability problem. These attacks exploit either new gaps or vulnerabilities in software or hardware in novel ways and cannot be matched against known signatures. When a vulnerability has already been exploited, the attacker uses a zero-day exploit to conduct a cyber-attack, which can lead to issues such as identity theft or data loss [47].

3. CYBER ATTACKS DETECTION

Cybersecurity is a continuously evolving on network, constantly active process with cyber-attacks development more frequent and sophisticated. Due to enormous amounts of traffic data, traditional detection systems cannot keep up with the requirements of security administrators and analysts to defend and protect against various new types of attacks [48].

Cyber attacks detection by ML has made progress in cyber-attack detection and is efficient for analyzing data, which is used to build an appropriate model for making the right decisions. In cyber security datasets, various algorithms are used, such as support vector machines (SVM) and principal component analysis (PCA), which are applied to reduce the dimensionality [49]. One of the big challenges facing the cyber security domain is the traditional ML approaches described in previous works cannot effectively deal with

sophisticated security problems and massive data in cyberspace. As a result, developing DL architecture-based cyber-attack detection methods has become a primary solution due to their ability to explore and exploit the complicated relationships between the collected datasets or collected information by monitoring network, system status, behavior, features, and the usage of the system, which could detect unauthorized users [50].

DL is used to improve cyberspace security, and various challenges in incorporating DL into cyber security are analyzed [51]. In the last decade, DL innovation has been employed in a variety of cases related to cyber security, including DDoS detection and prediction, intrusion detection, spam detection, phishing and malware detection on computer networks analysis and for real-world data [52]. In the DL model, the progress of a cyber security model is to automatically select features and provide feature learning in network layers [53]. The feature selection of anomaly detection techniques is a critical point for cyber network [54]. It can help to select general features such as unusual behavior of user to detect, predict, and prevent various types of attacks [55]. The intelligence techniques have the main role of selecting the number of features and analyzing the traffic network to identify the type of attack in a data set and in real time data [56].

4. CHALLENGES OF USING INTELLIGENT TECHNIQUE FOR CYBER SECURITY

Securing information systems has become one of the biggest challenges of the present day. With the increase of cyber-attacks that can cause more damage to computer systems, it is becoming difficult to identify cyber-attacks in big cyber space [57]. The cyber security model in organizations needs a security analyst to protect the systems. The security analyst faces many challenges related to cyber security attack detection, prediction, and prevention [58]. As a result, after reviewing state-of-the-art algorithms and techniques of ML, DL and artificial intelligence (AI) algorithms, those were used to face challenges for the attack detection process. Many other problems and issues are exposed as well. Some of these main issues are discussed below.

4.1. Dataset

The data set is a collection of data gathered from network traffic with various features and different network resources [59]. The large size of cyber data is one of the main challenges because the data sets include massive amount of redundant, outliers, and useless data, such as the in KDD-Cup 1999 data set [60]. In this case of dataset, the data preprocessing stage needs data mining and statistics techniques to extract benefits data from big data [61]. The most famous datasets uncovered some types of cyber-attack features because the datasets were outdated and cyber-attacks were covered and could be updated with new behaviors. As a result, dealing with the most recent update to the dataset is one of the best ways to detect new attacks [62].

DL techniques perform better when dealing with large amounts of data for both training and testing stages [63]. Data is collected from heterogeneous sources in order to build a strong security system [64]. Real-time and novel attacks are rarely seen in widely available databases. Due to these limitations, the most recent and exemplary benchmark dataset has yet to be identified. The main problems with datasets are the poor quality of the data and the availability of less data. Furthermore, the data contains imbalanced attack categories [65].

4.2. Real time data

Several traditional detection methods fail due to their limitations in real-time data, because of the complexity, universality, or heterogeneity of sources. Real-time data collection needs the procedure of monitoring network traffic for feature extraction. There are many tools used to analyze and monitor network traffic, such as Netflow, Kali Linux, and Wireshark. The real-time environment issues for attack detection are limited to dealing with machine learning algorithms [66]. In the real world, feature extraction makes detection attacks more reliable and useful in sampling. On the other hand, they note that the real datasets could be lacking with respect to the features identified because they are collected from different sources and regular growing attacks [67]. Cybercriminals generate new attacks and update them on a daily basis in order to disclose the network's weaknesses. If there are more false alarms, security administrators will lack confidence in the system [68]. The DL deals with real-time data, which is the optimal solution, particularly for feature learning purposes [69]. The model of DL can adapt with real-time data even updated data and added new complex features [70].

4.3. Time complexity

In the case of a real-time environment, it is more complex to detect an attack in less time than expect it to occur because of network issues and equipment [71]. On the other hand, the offline data set to build a cyber security model is split into two main stages: training and testing data. The time consumed in both stages is different from the intelligent algorithms that run and detect attacks. The performance and time required for detecting attacks depend on the algorithm that is used to build the security model. It also depends on the environment of the model working on it, including software and system properties. Most of the attack detection models suffer from longer training time, which affects the performance of the model. The cyber security gap

exists between the times an attacker successfully escapes prevention security systems. The complexity of time for ML and DL models for attack detection and building models has been improving speed and computational cost by using advanced hardware to reduce training and testing model time [72]. To overcome the latency in attack detection, ML and DL approaches necessitate amount of high-performance resources and data while training models. Using multiple GPUs is one approach, but it is neither power-efficient nor cost-effective [73].

4.4. Growing new attacks

Through the growth of cyber-attacks, there are two main challenges in using ML to handle these new attacks. Firstly, ML and DL models are used to locate previously unseen activity. Secondly, newer attacks are frequently more technically advanced than earlier ones [74]. Cyber security models are frequently trained with more past features, which means previous actions recorded in the dataset. The latest attacks may escape from classifiers and generate a false alarm or reduce the detection rate. In case of malware each year, there are over 6 million new strains of malware attacks discovered. It is also easy for an attacker to circumvent signature-based approaches. Malware can be reformed without altering its behavior but changing its signature [75].

Previous detection solutions have many limitations because they depend on information about known threats but are often powerless to identify the unknown. The behavior-based detection of the unknown is the main challenge in cybersecurity because of unseen attacks, in this case the detection rate becomes very low and false negatives increase. Exploration allows analysts to discover new anomalies that are not yet covered by rules and select features useful for detection [68]. Researchers are currently focused automated ML and DL models to detect new and previously undetected cyber attack is one of the greatest practices to adopt [76].

4.5. Predictive attacks

The network will be more protected if analysts are able to predict or anticipate cyber-attacks. The current prediction process indicates that there is limited success. The prediction based on network vulnerabilities can be effective, but requires up-to-date knowledge of the network, which is challenging due to errors in attacks detection, incomplete network information and attack obfuscation [77]. Using ML and data mining techniques to predict cyber-attacks can be helpful to reduce the number of features cyber-attacks [78]. A building prediction model should extract historical data and previously seen features of attacks action stored in order to predict next action of cyber-attacks [50]. This model needs recurring attack patterns and knowledge of different types of attacks previously attack known using feed-forward DL recurrent neural networks (RNN) and long short-term memory (LSTM) is a sequence learning [79].

4.6. Machine learning algorithms

The selection of an effective and suitable ML technique for a certain problem remains a challenging matter of critical concern. It depends on the types of data. ML provides solutions for most cyber security issues, such as detection, prevention, identification, and perdition attacks [80]. The main problem with ML-detection software is the impossible number of alerts it generates with low accuracy detection to develop a system that handles false positives and reduces false alerts [81]. The unavailability of benchmark and updated datasets for ML model training is a significant challenge. The traditional and single cyber security ML algorithm may fail to cope with various problems of attack detection as a result, using the hybrid DL and ML method to overcome issue. Building robust ML models and combining the advantages of both incremental and DL can yield incredible results. The main challenge of learning algorithms is insufficient quantities of training data in the case of overfitting or underfitting. The overfitting refers to a model that the training data too well. on other hand, underfitting refers to a model that can neither model the training data nor generalize to new features [82].

4.7. Complex features

Every time a new device connects online, in turn, both the quantity of data flows and the variety of attack aspects are growing. The features contain different categories of data types. The exponential growth of the IoT, 5G technology, cloud computing, and big data is accompanied by the generation of complex types of data [83]. In cyberspace, there is a wide range of attack types with complex features and behaviors their sophistication is increasing. The security model needs some important features to identify the type of attack. The selection of features helps reduce the time and complexity of the model. The feature selection processes are used to select the most general behaviors of attacks in the process of detection and classification [84].

4.8. Evaluation metric

The performance of models in detecting attacks, particularly in detection tasks, is described through evaluation metrics. ML algorithms are evaluated using many performance indicators [85]. Classification and clustering tasks both have different metrics for evaluating the accuracy of a model [86]. Selecting an appropriate metric is a main challenge because different metrics are proposed to evaluate different problems and the application of cyber security cases will result in different outcomes based on the metrics of accuracy,

confusion matrix, time, cohen Kappa, error rate, precision, sensitivity, recall, and specificity. It is necessary to consider an agreed-upon set of measures for comparing results with similar scenarios in order to make additional improvements [68]. The process of detecting attacks in a real-time network is evaluated and analyzed using different metrics such as average response time, average errors, and software response [87].

5. LITERATURE REVIEW

Recently, much research has been presented in the literature to improve the issues of cybersecurity, which involves preventing, detecting, predicting, defending, and recovering from cyber-attacks. This section discusses some methods used for attack types, data collection, challenges facing researchers, and problem domains and methods used for problem solving. Vinayakumar *et al.* [88], presented advanced ML algorithms such as hybrid DL framework for real-time malware visual detection. Also, for dimensionality reduction technique used feature engineering and feature learning to reduce volume of big dataset. Furthermore, a new upgraded approach for detecting effective new attacks known as zero-day malware detection based on static and dynamic analysis has been developed. On both public and private datasets from various sources, the proposed approach was used. The use of DL for robust intelligent zero-day malware detection significantly improved attack detection accuracy.

Wang and Jones [89], this paper focused on the large volume of data produced by network flows and system tools. The different characteristics of data in network traffic are analyzed using the R language and its functions. They performed the data preprocessing stage, including removing redundant and missing values, to get the best quality of data. In addition, a statistical analysis of variables and a clustering analysis based on the k-means cluster are carried out. The proposed method efficiently solves and analyzes different data types and attacks detection. Obitade [90], presented one of the challenges when detecting cyber-attacks in the big space of data with complex features. Analyze and organize big data with different features to improve cyber space protection. As a methodological framework, the resource-based view (RBV) provided advanced analytic functions of big data analytics as a way to extract value from the space of massive amounts of data. The survey in this paper uses data from 479 business and information security executives and draws from the RBV. Arivudainambi *et al.* [91], proposed a novel method of malware attack detection which is a hybrid of PCA and ANN for classifying malicious traffic rather than using the traditional single model because the single method fails to handle various types of attacks. The experimental results show that the proposed method is efficient for classifying attack with 99% accuracy and requiring less time to detect attacks.

Novo *et al.* [92], focused on common AI and ML algorithms applied to IDS cases. The first important step is to handle the problem of big data in cybersecurity. The UNSW-NB15 dataset is divided into groups. It is widely used in cybersecurity for the detection of various types of attacks based on features. They evaluated which neural network model (multilayer or recurrent) and ML algorithm get higher accuracy and decrease the computational load of the system depending on the group of data. According to the performance of the proposed model, the RNN is the best architecture for training data stage. Zhang *et al.* [93], proposed a novel method based on DL, which is the parallel cross convolutional neural network (PCCN) for attack detection. This method is used for the feature learning process to increase the detection performance of imbalanced abnormal flows. Also, enhance the feature extraction for multi-classes to reduce the amount of useless data. There are many challenges handled in this proposed model, such as detecting attacks in different classes, reducing detection time, and using DL architecture rather than traditional ML. The experiment was applied to the CICIDS2017 data set using the CICFlowMeter-V3 tool to extract 84-dimensional flow features. Also, the real-time dataset is used in the training and testing stages. Overall, PCCN gets better performance in accuracy detection.

Chen *et al.* [94], proposed a CNN algorithm for feature extraction because of the big volume and complexity of feature malware attacks. The challenges of malware attack detection are in the diversity and complexity of the types and structures of features that include source code, binary files, and other behaviors. The purpose of this research is to identify and detect different types of malware from different data collection samples of datasets from honeypots, Github, and the NCHC malware knowledge base, GNU. The accuracy of false positive alerts was higher than 90% in all experiments. Fang *et al.* [95], in this work, developed a method for predicting different cyber-attacks using DL architecture by utilizing Bi-directional recurrent neural networks with long short-term memory (BRNN-LSTM). The proposed work uses five real-world data sets which handle the problem of dataset long-range dependence and high nonlinearity. The method BRNN-LSTM achieved higher prediction accuracy compared to traditional statistics approaches with sufficient results.

Ibor *et al.* [96], proposed new model for predicting cyberattacks using DL architecture was conveniently into a new model using (ReLU) as the activation function in the hidden layers of a deep feed-forward neural network. This work achieved best represents the features beneficial for predicting cyberattacks in the use of a dataset of CICIDS2017 and UNSW_NB15. There are different processes performed in this

model, such as feature selection, ranked features, dimensionality reduction, and the k mean clustering algorithm at the initial stage, to generate a set of input vectors called hyper-features. The model is applied to a python environment. The experimental result gets high prediction accuracy of 99.99%.

Elmrabit *et al.* [97], evaluated 12 ML algorithms to detect attacks applied to three data sets, which are CICIDS-2017, UNSW-NB15, and the industrial control system (ICS). The challenges of this work are that attackers' behaviors change continuously over time, and it is difficult to point to one algorithm for detecting different cyber-attacks. The evaluation results of the proposed work find that the random forest (RF) algorithm achieves the highest accuracy in attack detection. Shaukat *et al.* [98], the objective of this work is to evaluate which of three algorithms for ML techniques, including DBN, decision tree, and SVM achieves high accuracy for cyber-attack detection. The challenge of this research was detecting a sophisticated and new attack which is zero-day attacks. The proposed model is applied to different datasets which are Spambase, Twitter Dataset, Enron, Spambase, NSL-KDD, malware dataset, KDD CUP99 with different types of attacks. The evaluation of this work illustrated that different ML algorithms are being used for different cyber attacks.

Singh *et al.* [99], in this research presented the detection of phishing attacks based on CNN as applied to e-commerce web pages. The proposed work faces challenging issues due to the semantic structure of web pages it attempts to identify a phishing attack. The data preprocessing is applied and feature extraction from the URLs is automatically done through its hidden layers. The result showed that the model achieved accuracy of 98.00% with reducing detection time. Zhang *et al.* [100], introduced the challenge of unknown attack detection, but they mentioned in research that it is still an unsolved problem completely because of the difficulties of collecting unknown attack samples and the inability to detect them in a timely manner. The proposed method uses learning the mapping relations between feature space and semantic space to recognize unknown attacks and deals with zero-shot learning. The test of this work has been applied to the NSL_KDD dataset. The experimental results showed that the accuracy of the unknown attack detection reached 88.3%.

Sedjelmaci [101], presented the distributed detection systems executed on the different critical 5G wireless systems to protect them from dangerous network attacks such as jamming and DDoS attacks. In this study, proposed a new cooperative detection system based on hierarchical reinforcement learning (RL) algorithms to classify network attacks. The experimental results showed that the RL detection system enhances the detection of new misbehavior attacks with high accuracy.

Mane *et al.* [102], presented genetic programming model used to defend such vulnerabilities in systems. A modern DDoS dataset is applied to detect novel attacks. This dataset contains new attacks collected from various environments and produced using network simulator (NS2). The results showed that the proposed model detects DDoS attacks with the accuracy of 98.67% applied by python.

Zuech *et al.* [103], proposed detecting web attacks using random under sampling and ensemble learners. Mostly, the challenge of this work is to treat class imbalance problems, which is an important consideration for cybersecurity and ML. Eight different ML classifiers are employed with different undersampling (RUS) ratios. The classification performance in detecting web attacks is applied to the CSE-CIC-IDS2018 dataset. Are the researchers' requirements to get answers to three questions about using statistical approaches, in various classifiers, and undersampling ratios, significant techniques for web attacks, after conducting research, it became evident that the answer to all questions is yes.

In 2021 Kumar and Sinha [104], suggested a novel robust intelligent cyber-attack detection method. The ML and DNN approach are used to defend against anomaly attacks, which are zero-day attacks that target unknown vulnerabilities in systems. The proposed work involves two phases: first, signature generation, and second, the evaluation phase. The result of the analysis of the proposed zero-day attack detection showed higher performance with an accuracy of 91.33% for the binary classification and an accuracy of 90.35% for multi-class classification on real-time data. The data is captured by creating a simulated environment that consists of 10 genuine. The performance applied to CICIDS18 showed the best results for detecting zero-day attacks.

Ugwu *et al.* [105], built a learning model which is LSTM for DDoS attack detection with SVD algorithm. This study was evaluated based on the UNSW-NB15 and NSL-KDD datasets. The processes of dimensionality reduction, feature extraction, and preprocessing stage are applied to reduce redundant features employed in building model. The results showed that the proposed model represents a significant improvement when compared with the traditional ML algorithm. Also, the LSTM model performed better on UNSWNB15 with an accuracy of 94.28% and NSL-KDD dataset with an accuracy of 90.59%. Yang [106], presented an intrusion detection system based on DL that applies BLSTM architecture to the system. The proposed model was tested and trained on the UNSW-NB15 data set. The results of this work illustrated that the intrusion detection system effectively improve at detecting known or unknown attacks on features of the network under the current network environment with high accuracy. Aslan and Yilmaz [107], proposed novel hybrid DL algorithms for the classification of malware rather than using traditional AI and ML algorithms, which are no longer effective in detecting all new and complex malware updates. This work involves four main stages, which are: data collection, design of a deep neural network architecture, training of the proposed deep neural network, and evaluation of the trained deep neural network. The proposed novel hybrid model is applied to the Maling,

Microsoft BIG 2015, and Malevis datasets. In general, the experimental results showed that the suggested method can effectively classify malware with high accuracy 97.78% attacks detection.

6. DISCUSSION AND COMPARISON

This paper provides a comprehensive review of the studies that have been conducted most recently on cybersecurity. There are many challenges faced by the analyst when using intelligent techniques to address cybersecurity issues pointed out with solutions in the papers reviewed. Also, it describes the different cyber-attack types and defense methods, and the suitable algorithms used in intelligent techniques (machine learning and deep learning) schemes for cybersecurity attack detection in terms of feature reduction, big data set analysis, real-time data, new types of attack, and time complexity. Finally, the issues and challenges of cybersecurity are highlighted, and provided solutions are discussed with assessments and recommendations.

The employment of DL algorithms in solving cyber security challenges is progressing. The main challenge is dealing with complex, big data sets and collecting real-time data with different types of features. The process of collecting online data uses tools and statistical techniques that are not reliable. In this study, we presented some benchmark datasets with descriptions and evaluated their performance in representing ways to demonstrate the current working state of attack detection methods with intelligent technique structures. DL algorithms have increased the efficiency of the decision model in detecting vulnerable risks in cyberspace. However, many gaps remain unaddressed due to the limitations of the DL method. A robust learning architecture is utilized to train models for feature analysis and pattern detection. Most studies prove that DL techniques achieve the best solution as compared with traditional ML algorithms in the process of detection, identification, prediction, prevention, recognition, and early detection of different types of attacks. As a result of the review, it is now possible to classify and identify the most common types of attacks, vulnerabilities, and defense mechanisms depending on specific techniques, which is hybrid DL architecture, because the process of building a security model needs many stages to overcome the different challenges it faces. To make the work more focused, the research objectives are presented in Table 1 (see in Appendix).

7. ASSESSMENT AND RECOMMENDATION

The recommendation and solution in order to overcome challenges should be to build a robust intelligent system that works with high accuracy detection attacks. The important point is that the different learning models are being used for specific different cyber-attacks. The use of hybride DL architecture is one of the best methods for issues of cyber security in case detection, prediction, and prevention of cyber-attacks. The DL couldn't produce sufficient results with the less volume of data, poor quality of data available and imbalanced data in classes. The large volume of data collection informs security decisions and detects the right type of attack because it consists of various types of features and feeds the network sufficient data.

The issue with collecting real data by using security tools is that they lack coverage of all features to anticipate cyber-attacks. Also, the analyst is unable to compare results achieved by a specific case of real-time data cyber-attack with previous work because the data may be collected with different network resources in different domains. On the other hand, the offline dataset needs updating to cover the behavior of new types of attacks. Collection of more data to avoid imbalanced data in data sets, which refers to the number of features per type of attack that is not equally distributed. This leads to model bias for the type of attack that is more available in the data training set. The imbalanced data in classification attacks may give incorrect accuracy to the model. As a result, the first step in data preprocessing, which includes data cleaning, handling imbalanced data, feature scaling, removing outliers, data transformation and dimension reduction to prepare data before training model. There is a possibility that conventional attack detection methods using ML, data mining, and statistical techniques miss many new cyber-attacks. The main challenges of learning algorithms are insufficient quantities and non-reprehensive training data. Furthermore, both the poor quality of the data and irrelevant features are negatively affected by the performance of training model. Additionally, it is important to address the problem of overfitting and underfitting of the training data stage, because both cases increase model complexity and the duration time of training. To reduce the time of the training process, it is suitable using feature scaling technique. This technique limits the features to a certain range, which is often between zero and one, to speed up the training process because the disparity in the size of the features necessitates more processing and takes longer time. There are various metrics for evaluating that can be identified by the purpose of the model, such as detection rate by accuracy, time of detection, precision, recall, and f1 score, and so on, to analyze the performance of a cyber security model for attack detection.

The main point of this research is that there is no specific intelligent algorithm to detect and predict a specific type of attack with high accuracy. Hence the best way to recommend solutions for detecting different types of attacks is to build a robust model architecture based on the most-suited algorithms for a particular

problem. As a final point, recently, according to the Cisco annual internet report and the most recent of studies identified that the DDOS which is the most dangerous and fast-growing attack to disrupt the normal traffic of a targeted server, network and computer system.

8. CONCLUSION

The trend of cyber security technology has achieved improvements in computer system and network security measures through detecting and reacting against cyber-attacks. This paper presents the main important challenges and limitations of using intelligent techniques in cyber security with solution methods and analysis performance, as well as a comprehensive bibliography in this area. Furthermore, it provides description of the types of cyber-attacks in cyber space. The DDoS an active attack type has been dangerous threats to the cyber world because of their potential to bring down victims and their capability to expand and change. There are various types of attack features with different case domains. Detecting cyber-attacks has undoubtedly become a big data problem. A dataset is a crucial point for the training and testing of building models because there is the unavailability of representative datasets for each attack domain. The using of hybrid DL model to secure cyber-space dedicated of the recent research works to solve computer security challenges has made significant progress by ensuring the robustness of a network as well as maintaining the integrity of the data.

APPENDIX

Table 1. Cybersecurity challenges for cyber attacks detection

Ref	Attack type	Data	Challenges	Problem domain	Method used for problem solution
[88]	Zeroday malwares	Real time public and private datasets	In real time identifying large number unknown malwares.	Network traffic	Robust intelligent detection using Hybrid DL.
[89]	Various attacks	Masquerading user data and KDD CUP	Big data different characteristics in different network traffic	Network traffic	Correlation analysis of features clustering analysis based k-means,
[90]	Different cyber attacks	Collected from IS and business executives	Analyze big data volumes improve cyber protection	479 business organization and IS	Advanced data analytics using RBV methodological framework
[91]	Novel Malware	Network traffic dataset	Traditional ML and single model fail to cope with new attack	network traffic classification	Extreme surveillance. Hybrid PCA and ANN.
[92]	8 types of attacks	Divide data set UNSW-NB15 to groups	Evaluate of ML algorithm to computational load of system	IDS network traffic	Neural Network algorithms using RNN get best results
[93]	14 types of attacks	CICIDS2017 dataset and real time data	Detecting multi-class imbalanced traffic data to get best accuracy	Network traffic	Deep learning techniques, parallel cross CNN.
[94]	Different malware	honeypot Dataset, NCHCmalware, GNU	Growth data volume, variety and complexity in malware attack	Network traffic	Using CNN algorithm find malicious or benign code
[95]	Different cyber attacks	Using five real-world datasets	Data set long-range dependence and high nonlinearity	Real-world net cyber attack	Develop a DL by utilizing the BRNN with LSTM algorithms
[96]	14 types of attacks	CICIDS2017 UNSW_NB15	-Big data through device sensors. -Cyber attack predictions	Network traffic	EM algorithm then K clustered dataset and Developing DNN.
[97]	More than 14 attacks	CICIDS-2017, UNSW-NB15, industrial control system datasets	-Attackers behaviors change -Pointing one algorithm for detecting attacks	ALICEperformance at University of Leicester	Testing 6 ML and 6 DL algorithms with big 3 types of data sets. RF algorithm has best performance.
[98]	Different types of Attacks	Spambase, Twitter, Enron, NSL-KDD, Malware, KDD CUP99	-Detecting zero-day attacks. -Evaluate which ML algorithm suitable for which type of attacks	Network traffic	-Used DBN, DT and SVM -Various learning DL models are being used
[99]	Phishing on URLs	taken from e-commerce website	Identify phishing or legitimate web page	URLs, deceive website, email	using DLtechniques (CNN) with high accuracy.
[100]	10 types of attacks	NSL_KDD	-Detect unknown attacks. -Extract different features	Network traffic	Novel method of zero-shot learning based on sparse AE
[101]	Jamming, DDoS	a real telecom dataset	-Distributed detection, on 5G -detect unknown attacks	wireless communicatios	New detection system based on hierarchical RL algorithms.
[102]	DDoS	Modern DDoS Dataset	Detect novel attack of DDoS	IDS on network	PCA and genetic programming
[103]	Web attack	CSE-CICIDS2018.	-Big data imbalance classes -Evaluate ML algorithms	Website on net	DT, RF, CB, LGB, XGB, NB, LR
[104]	Zero-day, real time attack	CICIDS18 and real data	-Detect unknown attack, vulnerabilities of a software	vulnerabilities of a software or system opens	Built robust detection model in two phase signature generation and evaluation phase HVA, LVA
[105]	DDoS	UNSW-NB15, NSL-KDD	Expanding the scope of DDoS	Network traffic	DL (LSTM) and SVD algorithms.
[106]	Unknown	UNSWNB15	time detection of unknown attack	Network traffic	DNN-BLSTM sequence learning
[107]	New malware	Maling, Microsoft BIG2015, Malevis	Traditional AI, ML algorithm not effective in detecting new malware	cybersecurity domain on nnetwork	Novel hybrid DL architecture

REFERENCES

[1] B. Alhayani, S. T. Abbas, D. Z. Khutar, and H. J. Mohammed, "WITHDRAWN: best ways computation intelligent of face cyber attacks," *Materials Today: Proceedings*, 2021, doi: 10.1016/j.matpr.2021.02.557.

[2] Y. He, "Research on the key technology of network security based on machine learning," in *2021 IEEE 6th International Conference on Intelligent Computing and Signal Processing, ICSP 2021*, 2021, pp. 972–975, doi: 10.1109/ICSP51882.2021.9408756.

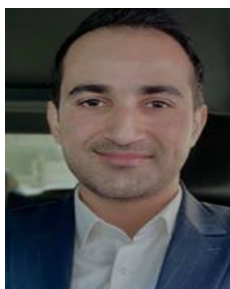
- [3] R. Derbyshire, B. Green, D. Prince, A. Mauthe, and D. Hutchison, "An analysis of cyber security attack taxonomies," in *Proceedings - 3rd IEEE European Symposium on Security and Privacy Workshops, EURO S and PW 2018*, 2018, pp. 153–161, doi: 10.1109/EuroSPW.2018.00028.
- [4] A. Verma, R. Surendra, B. Srikanth Reddy, P. Chawla, and K. Soni, "Cyber security in digital sector," in *Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021*, 2021, pp. 703–710, doi: 10.1109/ICAIS50930.2021.9395933.
- [5] S. Kumar, H. Kumar, and G. R. Gunnam, "Security integrity of data collection from smart electric meter under a cyber attack," in *Proceedings - 2019 2nd International Conference on Data Intelligence and Security, ICDIS 2019*, 2019, pp. 9–13, doi: 10.1109/ICDIS.2019.00009.
- [6] A. A. Mishra, K. Surve, U. Patidar, and R. K. Rambola, "Effectiveness of confidentiality, integrity and availability in the security of cloud computing: A review," in *2018 4th International Conference on Computing Communication and Automation, ICCCA 2018*, 2018, doi: 10.1109/CCAA.2018.8777537.
- [7] M. Lehto and J. Limnell, "Strategic leadership in cyber security, case Finland," *Information Security Journal*, vol. 30, no. 3, pp. 139–148, 2021, doi: 10.1080/19393555.2020.1813851.
- [8] P. Maheshwaran and S. Rajagopal, "A scheme for detecting the types of misbehaviour and identifying the attacks using reputation mechanism in a mobile ad-hoc network," in *Proceedings of the International Conference on Communication and Electronics Systems, ICCES 2016*, 2016, doi: 10.1109/CESYS.2016.7889961.
- [9] H. H. Addeen, Y. Xiao, J. Li, and M. Guizani, "A survey of cyber-physical attacks and detection methods in smart water distribution systems," *IEEE Access*, vol. 9, pp. 99905–99921, 2021, doi: 10.1109/ACCESS.2021.3095713.
- [10] A. El-Shafie, H. Chihouai, R. Hamila, N. Al-Dhahir, A. Gastli, and L. Ben-Brahim, "Impact of passive and active security attacks on mimo smart grid communications," *IEEE Systems Journal*, vol. 13, no. 3, pp. 2873–2876, 2019, doi: 10.1109/JSYST.2018.2868291.
- [11] A. Salih, S. T. Zeebaree, S. Ameen, A. Alkhyat, and H. M. Shukur, "A survey on the role of artificial intelligence, machine learning and deep learning for cybersecurity attack detection," in *Proceedings of the 7th International Engineering Conference "Research and Innovation Amid Global Pandemic", IEC 2021*, 2021, pp. 61–66, doi: 10.1109/IEC52205.2021.9476132.
- [12] Y. Feng, H. Akiyama, L. Lu, and K. Sakurai, "Feature selection for machine learning-based early detection of distributed cyber attacks," in *Proceedings - IEEE 16th International Conference on Dependable, Autonomic and Secure Computing, IEEE 16th International Conference on Pervasive Intelligence and Computing, IEEE 4th International Conference on Big Data Intelligence and Computing and IEEE 3rd Cyber Science and Technology Congress, DASC-PiCom-DataCom-CyberSciTec 2018*, 2018, pp. 181–186, doi: 10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00040.
- [13] M. R. Babu and K. N. Veena, "A survey on attack detection methods for IOT using machine learning and deep learning," in *2021 3rd International Conference on Signal Processing and Communication, ICSPC 2021*, 2021, pp. 625–630, doi: 10.1109/ICSPC51351.2021.9451740.
- [14] A. Al-Abassi, H. Karimipour, A. Dehghantaha, and R. M. Parizi, "An ensemble deep learning-based cyber-attack detection in industrial control system," *IEEE Access*, vol. 8, pp. 83965–83973, 2020, doi: 10.1109/ACCESS.2020.2992249.
- [15] M. F. Kabir and S. Hartmann, "Cyber security challenges: an efficient intrusion detection system design," in *Proceedings - 2018 International Young Engineers Forum, YEF-ECE 2018*, 2018, pp. 19–24, doi: 10.1109/YEF-ECE.2018.8368933.
- [16] A. Phadke, M. Kulkarni, P. Bhawalkar, and R. Bhattad, "A review of machine learning methodologies for network intrusion detection," in *Proceedings of the 3rd International Conference on Computing Methodologies and Communication, ICCMC 2019*, 2019, pp. 272–275, doi: 10.1109/ICCMC.2019.8819748.
- [17] P. R. Chandre, P. N. Mahalle, and G. R. Shinde, "Machine learning based novel approach for intrusion detection and prevention system: a tool based verification," in *Proceedings - 2018 IEEE Global Conference on Wireless Computing and Networking, GCWCN 2018*, 2019, pp. 135–140, doi: 10.1109/GCWCN.2018.8668618.
- [18] R. Raval, A. Maskus, B. Saltmiras, M. Dunn, P. J. Hawrylak, and J. Hale, "Competitive learning environment for cyber-physical system security experimentation," in *Proceedings - 2018 1st International Conference on Data Intelligence and Security, ICDIS 2018*, 2018, pp. 211–218, doi: 10.1109/ICDIS.2018.00042.
- [19] I. D. Barbu, C. Pascariu, I. C. Bacivarov, S. D. Axinte, and M. Firoiu, "Intruder monitoring system for local networks using python," in *Proceedings of the 9th International Conference on Electronics, Computers and Artificial Intelligence, ECAI 2017*, 2017, vol. 2017-January, pp. 1–4, doi: 10.1109/ECAI.2017.8166457.
- [20] D. Preethi and N. Khare, "Performance Evaluation of Shallow learning techniques and deep neural network for cyber security," 2020, doi: 10.1109/ic-ETITE47903.2020.128.
- [21] Y. Goyal and A. Sharma, "A semantic approach for cyber threat prediction using machine learning," in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, Mar. 2019, pp. 435–438, doi: 10.1109/ICCMC.2019.8819694.
- [22] D. Sisiaridis and O. Markowitch, "Reducing data complexity in feature extraction and feature selection for big data security analytics," in *Proceedings - 2018 1st International Conference on Data Intelligence and Security, ICDIS 2018*, 2018, pp. 43–48, doi: 10.1109/ICDIS.2018.00014.
- [23] Y. Xin *et al.*, "Machine learning and deep learning methods for cybersecurity," *IEEE Access*, vol. 6, pp. 35365–35381, 2018, doi: 10.1109/ACCESS.2018.2836950.
- [24] H. B. Alharbi, N. Abdulrazak Baghanim, and A. Munshi, "Cyber risk in internet of things world," 2020, doi: 10.1109/ICCAIS48893.2020.9096720.
- [25] F. Meneghello, M. Calore, D. Zucchetto, M. Polese, and A. Zanella, "IoT: Internet of threats? A survey of practical security vulnerabilities in real IoT devices," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8182–8201, 2019, doi: 10.1109/IIOT.2019.2935189.
- [26] Y. S. Park, C. S. Choi, C. Jang, D. G. Shin, G. C. Cho, and H. S. Kim, "Development of incident response tool for cyber security training based on virtualization and cloud," in *2019 International Workshop on Big Data and Information Security, IWBS 2019*, 2019, pp. 115–118, doi: 10.1109/IWBS.2019.8935723.
- [27] M. Wankhade and S. V. Kottur, "Security facets of cyber physical system," in *Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT 2020*, 2020, pp. 359–363, doi: 10.1109/ICSSIT48917.2020.9214079.
- [28] S. Sharma, A. Yadav, M. Panchal, and P. D. Vyavahare, "Classification of security attacks in wsns and possible countermeasures: a survey," *International Symposium on Advanced Networks and Telecommunication Systems, ANTS*, vol. 2019-Decem. IEEE, 2019, doi: 10.1109/ANTS47819.2019.9118119.
- [29] J. M. Torres, C. I. Comesaña, and P. J. García-Nieto, "Review: machine learning techniques applied to cybersecurity," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 10, pp. 2823–2836, 2019, doi: 10.1007/s13042-018-00906-1.




- [30] A. Saravanan and S. S. Bama, "A review on cyber security and the fifth generation cyberattacks," *Oriental journal of computer science and technology*, vol. 12, no. 2, pp. 50–56, 2019, doi: 10.13005/ojst12.02.04.
- [31] I. Wiafe, F. N. Koranteng, E. N. Obeng, N. Assyane, A. Wiafe, and S. R. Gulliver, "Artificial intelligence for cybersecurity: a systematic mapping of literature," *IEEE Access*, vol. 8, pp. 146598–146612, 2020, doi: 10.1109/ACCESS.2020.3013145.
- [32] I. H. Sarker, "Deep cybersecurity: a comprehensive overview from neural network and deep learning perspective," *SN Computer Science*, vol. 2, no. 3, 2021, doi: 10.1007/s42979-021-00535-6.
- [33] L. Wang and R. Jones, "Big data analytics of network traffic and attacks," in *Proceedings of the IEEE National Aerospace Electronics Conference, NAECON*, 2018, vol. 2018-July, pp. 117–123, doi: 10.1109/NAECON.2018.8556802.
- [34] K. Liu, S. Xu, G. Xu, M. Zhang, D. Sun, and H. Liu, "A review of android malware detection approaches based on machine learning," *IEEE Access*, vol. 8, pp. 124579–124607, 2020, doi: 10.1109/ACCESS.2020.3006143.
- [35] A. E. Ibor, F. A. Oladeji, and O. B. Okunoye, "A survey of cyber security approaches for attack detection, prediction, and prevention," *International Journal of Security and Its Applications*, vol. 12, no. 4, pp. 15–28, 2018, doi: 10.14257/ijisa.2018.12.4.02.
- [36] F. Li, X. Yan, Y. Xie, Z. Sang, and X. Yuan, "A review of cyber-attack methods in cyber-physical power system," in *APAP 2019 - 8th IEEE International Conference on Advanced Power System Automation and Protection*, 2019, pp. 1335–1339, doi: 10.1109/APAP47170.2019.9225126.
- [37] Y. Supriya, G. Kumar, D. Sowjanya, D. Yadav, and D. L. Kameshwari, "Malware detection techniques: a survey," in *PDGC 2020 - 2020 6th International Conference on Parallel, Distributed and Grid Computing*, 2020, pp. 25–30, doi: 10.1109/PDGC50313.2020.9315764.
- [38] L. Caviglione, "Trends and challenges in network covert channels countermeasures," *Applied Sciences (Switzerland)*, vol. 11, no. 4, pp. 1–16, 2021, doi: 10.3390/app11041641.
- [39] S. Sabhadiya, J. Barad, and J. Gheewala, "Android malware detection using deep learning," in *Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019*, 2019, vol. 2019-April, pp. 1254–1260, doi: 10.1109/icoei.2019.8862633.
- [40] M. Baykara and Z. Z. Gürel, "Detection of phishing attacks," in *6th International Symposium on Digital Forensic and Security, ISDFS 2018 - Proceeding*, 2018, vol. 2018-Janua, pp. 1–5, doi: 10.1109/ISDFS.2018.8355389.
- [41] J. Ahmed and Q. Tushar, "Covid-19 pandemic: a new era of cyber security threat and holistic approach to overcome," in *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering, CSDE 2020*, 2020, doi: 10.1109/CSDE50874.2020.9411533.
- [42] S. Singh, N. Yadav, and P. K. Chuarasia, "A review on cyber physical system attacks: issues and challenges," in *Proceedings of the 2020 IEEE International Conference on Communication and Signal Processing, ICCSP 2020*, 2020, pp. 1133–1138, doi: 10.1109/ICCSP48568.2020.9182452.
- [43] C. Aravindan, T. Frederick, V. Hemamalini, and M. V. J. Cathirine, "An extensive research on cyber threats using learning algorithm," in *International Conference on Emerging Trends in Information Technology and Engineering, ic-ETITE 2020*, 2020, doi: 10.1109/ic-ETITE47903.2020.337.
- [44] V. Mullet, P. Sondi, and E. Ramat, "A review of cybersecurity guidelines for manufacturing factories in industry 4.0," *IEEE Access*, vol. 9, pp. 23235–23263, 2021, doi: 10.1109/ACCESS.2021.3056650.
- [45] I. Tasevski and K. Jakimoski, "Overview of SQL injection defense mechanisms," in *2020 28th Telecommunications Forum, TELFOR 2020 - Proceedings*, 2020, doi: 10.1109/TELFOR51502.2020.9306676.
- [46] R. K. Jothi, S. Balaji B, N. Pandey, P. Beriwal, and A. Amarajan, "An efficient SQL injection detection system using deep learning," in *Proceedings of 2nd IEEE International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2021*, 2021, pp. 442–445, doi: 10.1109/ICCIKE51210.2021.9410674.
- [47] T. Zoppi, A. Ceccarelli, and A. Bondavalli, "Unsupervised algorithms to detect zero-day attacks: strategy and application," *IEEE Access*, vol. 9, pp. 90603–90615, 2021, doi: 10.1109/ACCESS.2021.3090957.
- [48] I. H. Sarker, "CyberLearning: Effectiveness analysis of machine learning security modeling to detect cyber-anomalies and multi-attacks," *Internet of Things (Netherlands)*, vol. 14, p. 100393, 2021, doi: 10.1016/j.iot.2021.100393.
- [49] K. Di Lu, G. Q. Zeng, X. Luo, J. Weng, W. Luo, and Y. Wu, "Evolutionary deep belief network for cyber-attack detection in industrial automation and control system," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 11, pp. 7618–7627, 2021, doi: 10.1109/TII.2021.3053304.
- [50] I. H. Sarker, A. S. M. Kayes, S. Badsha, H. Alqahtani, P. Watters, and A. Ng, "Cybersecurity data science: an overview from machine learning perspective," *Journal of Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00318-5.
- [51] K. Sathya, J. Premalatha, and S. Suwathika, "Reinforcing cyber world security with deep learning approaches," in *Proceedings of the 2020 IEEE International Conference on Communication and Signal Processing, ICCSP 2020*, 2020, pp. 766–769, doi: 10.1109/ICCSP48568.2020.9182067.
- [52] K. Yang, J. Zhang, Y. Xu, and J. Chao, "DDoS attacks detection with autoencoder," in *Proceedings of IEEE/IFIP Network Operations and Management Symposium 2020: Management in the Age of Softwarization and Artificial Intelligence, NOMS 2020*, 2020, doi: 10.1109/NOMS47738.2020.9110372.
- [53] A. Saber, M. Abbas, and B. Fergani, "A DDoS attack detection system: applying a hybrid genetic algorithm to optimal feature subset selection," in *ISIA 2020 - Proceedings, 4th International Symposium on Informatics and its Applications*, 2020, doi: 10.1109/ISIA51297.2020.9416558.
- [54] K. N. Zakaria, A. Zainal, S. H. Othman, and M. N. Kassim, "Feature extraction and selection method of cyber-attack and threat profiling in cybersecurity audit," in *2019 International Conference on Cybersecurity, ICoCSec 2019*, 2019, pp. 1–6, doi: 10.1109/ICoCSec47621.2019.8970786.
- [55] M. Dehghani *et al.*, "Cyber attack detection based on wavelet singular entropy in AC smart islands: false data injection attack," *IEEE Access*, vol. 9, pp. 16488–16507, 2021, doi: 10.1109/ACCESS.2021.3051300.
- [56] R. R. Zebari, S. R. M. Zeebaree, and K. Jacksi, "Impact analysis of HTTP and SYN flood DDoS attacks on apache 2 and IIS 10.0 web servers," in *ICOASE 2018 - International Conference on Advanced Science and Engineering*, 2018, pp. 156–161, doi: 10.1109/ICOASE.2018.8548783.
- [57] J. Akram and L. Ping, "How to build a vulnerability benchmark to overcome cyber security attacks," *IET Information Security*, vol. 14, no. 1, pp. 60–71, 2020, doi: 10.1049/iet-ifs.2018.5647.
- [58] K. Cabaj, Z. Kotulski, B. Książkowski, and W. Mazurczyk, "Cybersecurity: trends, issues, and challenges," *Eurasip Journal on Information Security*, vol. 2018, no. 1, 2018, doi: 10.1186/s13635-018-0080-0.
- [59] D. S. Terzi, R. Terzi, and S. Sagiroglu, "Big data analytics for network anomaly detection from netflow data," *2nd International Conference on Computer Science and Engineering, UBMK 2017*. IEEE, pp. 592–597, 2017, doi: 10.1109/UBMK.2017.8093473.
- [60] I. H. Sarker, M. H. Furhad, and R. Nowrozy, "AI-driven cybersecurity: an overview, security intelligence modeling and research directions," *SN Computer Science*, vol. 2, no. 3, 2021, doi: 10.1007/s42979-021-00557-0.

- [61] R. Das and T. H. Morris, "Machine learning and cyber security," in *2017 International Conference on Computer, Electrical and Communication Engineering, ICCECE 2017*, 2017, doi: 10.1109/ICCECE.2017.8526232.
- [62] S. H. Ahn, N. U. Kim, and T. M. Chung, "Big data analysis system concept for detecting unknown attacks," in *International Conference on Advanced Communication Technology, ICACT*, 2014, pp. 269–272, doi: 10.1109/ICACTION.2014.6778962.
- [63] S. Dhir and Y. Kumar, "Study of machine and deep learning classifications in cyber physical system," in *Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT 2020*, 2020, pp. 333–338, doi: 10.1109/ICSSIT48917.2020.9214237.
- [64] D. Liu, H. Zhang, H. Yu, X. Liu, Y. Zhao, and G. Lv, "Research and application of APT attack defense and detection technology based on big data technology," in *ICEIEC 2019 - Proceedings of 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication*, 2019, pp. 701–704, doi: 10.1109/ICEIEC.2019.8784483.
- [65] D. Zhou, Z. Yan, Y. Fu, and Z. Yao, "A survey on network data collection," *Journal of Network and Computer Applications*, vol. 116, pp. 9–23, 2018, doi: 10.1016/j.jnca.2018.05.004.
- [66] J. Hou, P. Fu, Z. Cao, and A. Xu, "Machine learning based DDos detection through netflow analysis," in *Proceedings - IEEE Military Communications Conference MILCOM*, 2019, vol. 2019-Octob, pp. 565–570, doi: 10.1109/MILCOM.2018.8599738.
- [67] V. Kumar, D. Sinha, A. K. Das, S. C. Pandey, and R. T. Goswami, "An integrated rule based intrusion detection system: analysis on UNSW-NB15 data set and the real time online dataset," *Cluster Computing*, vol. 23, no. 2, pp. 1397–1418, 2020, doi: 10.1007/s10586-019-03008-x.
- [68] K. Shaikat, S. Luo, V. Varadarajan, I. A. Hameed, and M. Xu, "A survey on machine learning techniques for cyber security in the last decade," *IEEE Access*, vol. 8, pp. 222310–222354, 2020, doi: 10.1109/ACCESS.2020.3041951.
- [69] C. Callegari, E. Bucchianeri, S. Giordano, and M. Pagano, "Real time attack detection with deep learning," in *Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks workshops*, 2019, vol. 2019-June, doi: 10.1109/SAHCN.2019.8824811.
- [70] A. A. Darem, F. A. Ghaleb, A. A. Al-Hashmi, J. H. Abawajy, S. M. Alanazi, and A. Y. Al-Rezami, "An adaptive behavioral-based incremental batch learning malware variants detection model using concept drift detection and sequential deep learning," *IEEE Access*, vol. 9, pp. 97180–97196, 2021, doi: 10.1109/ACCESS.2021.3093366.
- [71] S. R. M. Zeebaree, K. Jacksi, and R. R. Zebari, "Impact analysis of SYN flood DDoS attack on HAProxy and NLB cluster-based web servers," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 1, pp. 505–512, 2020, doi: 10.11591/ijeecs.v19.i1.pp505-512.
- [72] S. Huang, E. H. Liu, Z. W. Hui, S. Q. Tang, and S. J. Zhang, "Challenges of testing machine learning applications," *International Journal of Performability Engineering*, vol. 14, no. 6, pp. 1275–1282, 2018, doi: 10.23940/ijpe.18.06.p18.12751282.
- [73] K. M. Sudar, M. Beulah, P. Deepalakshmi, P. Nagaraj, and P. Chinnasamy, "Detection of distributed denial of service attacks in SDN using Machine learning techniques," in *2021 International Conference on Computer Communication and Informatics, ICCCI 2021*, 2021, doi: 10.1109/ICCCI50826.2021.9402517.
- [74] M. Baykara, U. Gurturk, and R. Das, "An overview of monitoring tools for real-time cyber-attacks," in *6th International Symposium on Digital Forensic and Security, ISDFS 2018 - Proceeding*, 2018, vol. 2018-Janua, pp. 1–6, doi: 10.1109/ISDFS.2018.8355339.
- [75] M. Weiss, "From prediction to anticipation of cyber attacks," *IDRBT JOURNAL OF*, vol. 01, pp. 1–11, 2018.
- [76] M. Komar, V. Dorosh, G. Hladiy, and A. Sachenko, "Deep neural network for detection of cyber attacks," in *2018 IEEE 1st International Conference on System Analysis and Intelligent Computing, SAIC 2018 - Proceedings*, 2018, doi: 10.1109/SAIC.2018.8516753.
- [77] M. Husák, J. Komárková, E. Bou-Harb, and P. Čeleda, "Survey of attack projection, prediction, and forecasting in cyber security," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 1, pp. 640–660, 2019, doi: 10.1109/COMST.2018.2871866.
- [78] M. Husák and J. Kašpar, "Towards predicting cyber attacks using information exchange and data mining," in *2018 14th International Wireless Communications and Mobile Computing Conference, IWCMC 2018*, 2018, pp. 536–541, doi: 10.1109/IWCMC.2018.8450512.
- [79] A. M. S. N. Amarasinghe, W. A. C. H. Wijesinghe, D. L. A. Nirmana, A. Jayakody, and A. M. S. Priyankara, "AI based cyber threats and vulnerability detection, prevention and prediction system," in *2019 International Conference on Advancements in Computing, ICAC 2019*, 2019, pp. 363–368, doi: 10.1109/ICAC49085.2019.9103372.
- [80] G. J. Priya and S. Saradha, "Fraud detection and prevention using machine learning algorithms: a review," in *Proceedings of the 7th International Conference on Electrical Energy Systems, ICEES 2021*, 2021, pp. 564–568, doi: 10.1109/ICEES51510.2021.9383631.
- [81] A. Rashid, M. J. Siddique, and S. M. Ahmed, "Machine and deep learning based comparative analysis using hybrid approaches for intrusion detection system," in *3rd International Conference on Advancements in Computational Sciences, ICACS 2020*, 2020, doi: 10.1109/ICACS47775.2020.9055946.
- [82] N. Mishra and S. Pandya, "Internet of things applications, security challenges, attacks, intrusion detection, and future visions: a systematic review," *IEEE Access*, vol. 9, pp. 59353–59377, 2021, doi: 10.1109/ACCESS.2021.3073408.
- [83] M. Sergey, S. Nikolay, and E. Sergey, "Cyber security concept for internet of everything (IoE)," in *2017 Systems of Signal Synchronization, Generating and Processing in Telecommunications, SINKHROINFO 2017*, 2017, doi: 10.1109/SINKHROINFO.2017.7997540.
- [84] P. Parkar and A. Bilimoria, "A survey on cyber security IDS using ML methods," in *Proceedings - 5th International Conference on Intelligent Computing and Control Systems, ICICCS 2021*, 2021, pp. 352–360, doi: 10.1109/ICICCS51141.2021.9432210.
- [85] O. H. Jader, S. R. M. Zeebaree, and R. R. Zebari, "A state of art survey for web server performance measurement and load balancing mechanisms," *International Journal of Scientific and Technology Research*, vol. 8, no. 12, pp. 535–543, 2019.
- [86] K. Atefi, H. Hashim, and T. Khodadadi, "A hybrid anomaly classification with deep learning (DL) and binary algorithms (BA) as optimizer in the intrusion detection system (IDS)," in *Proceedings - 2020 16th IEEE International Colloquium on Signal Processing and its Applications, CSPA 2020*, 2020, pp. 29–34, doi: 10.1109/CSPA48992.2020.9068725.
- [87] R. R. Zebari, S. R. M. Zeebaree, A. B. Sallow, H. M. Shukur, O. M. Ahmad, and K. Jacksi, "Distributed denial of service attack mitigation using high availability proxy and network load balancing," in *3rd International Conference on Advanced Science and Engineering, ICOASE 2020*, 2020, pp. 174–179, doi: 10.1109/ICOASE51841.2020.9436545.
- [88] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, and S. Venkatraman, "Robust intelligent malware detection using deep learning," *IEEE Access*, vol. 7, pp. 46717–46738, 2019, doi: 10.1109/ACCESS.2019.2906934.
- [89] L. Wang and R. Jones, "Big data analytics in cyber security: network traffic and attacks," *Journal of Computer Information Systems*, vol. 61, no. 5, pp. 410–417, 2021, doi: 10.1080/08874417.2019.1688731.
- [90] P. O. Obitade, "Big data analytics: a link between knowledge management capabilities and superior cyber protection," *Journal of Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0229-9.




- [91] D. Arivudainambi, V. K. Varun, S. C. S., and P. Visu, "Malware traffic classification using principal component analysis and artificial neural network for extreme surveillance," *Computer Communications*, vol. 147, pp. 50–57, 2019, doi: 10.1016/j.comcom.2019.08.003.
- [92] X. A. Larriva-Novo, M. Vega-Barbas, V. A. Villagra, and M. Sanz Rodrigo, "Evaluation of cybersecurity data set characteristics for their applicability to neural networks algorithms detecting cybersecurity anomalies," *IEEE Access*, vol. 8, pp. 9005–9014, 2020, doi: 10.1109/ACCESS.2019.2963407.
- [93] Y. Zhang, X. Chen, D. Guo, M. Song, Y. Teng, and X. Wang, "PCCN: parallel cross convolutional neural network for abnormal network traffic flows detection in multi-class imbalanced network traffic flows," *IEEE Access*, vol. 7, pp. 119904–119916, 2019, doi: 10.1109/ACCESS.2019.2933165.
- [94] C. M. Chen, S. H. Wang, D. W. Wen, G. H. Lai, and M. K. Sun, "Applying convolutional neural network for malware detection," in *2019 IEEE 10th International Conference on Awareness Science and Technology, iCAST 2019 - Proceedings*, 2019, doi: 10.1109/ICAwST.2019.8923568.
- [95] X. Fang, M. Xu, S. Xu, and P. Zhao, "A deep learning framework for predicting cyber attacks rates," *Eurasip Journal on Information Security*, vol. 2019, no. 1, 2019, doi: 10.1186/s13635-019-0090-6.
- [96] A. E. Ibor, F. A. Oladeji, O. B. Okunoye, and O. O. Ekabua, "Conceptualisation of cyberattack prediction with deep learning," *Cybersecurity*, vol. 3, no. 1, 2020, doi: 10.1186/s42400-020-00053-7.
- [97] N. Elmrahit, F. Zhou, F. Li, and H. Zhou, "Evaluation of machine learning algorithms for anomaly detection," in *International Conference on Cyber Security and Protection of Digital Services, Cyber Security 2020*, 2020, doi: 10.1109/CyberSecurity49315.2020.9138871.
- [98] K. Shaukat, S. Luo, S. Chen, and D. Liu, "Cyber threat detection using machine learning techniques: a performance evaluation perspective," in *1st Annual International Conference on Cyber Warfare and Security, ICCWS 2020 - Proceedings*, 2020, doi: 10.1109/ICCWS48432.2020.9292388.
- [99] S. Singh, M. P. Singh, and R. Pandey, "Phishing detection from URLs using deep learning approach," in *Proceedings of the 2020 International Conference on Computing, Communication and Security, ICCCS 2020*, 2020, doi: 10.1109/ICCCS49678.2020.9277459.
- [100] Z. Zhang, Q. Liu, S. Qiu, S. Zhou, and C. Zhang, "Unknown attack detection based on zero-shot learning," *IEEE Access*, vol. 8, pp. 193981–193991, 2020, doi: 10.1109/ACCESS.2020.3033494.
- [101] H. Sedjelmaci, "Attacks detection approach based on a reinforcement learning process to secure 5G wireless network," in *2020 IEEE International Conference on Communications Workshops, ICC Workshops 2020 - Proceedings*, 2020, doi: 10.1109/ICCWshops49005.2020.9145438.
- [102] N. Mane, A. Verma, and A. Arya, "A pragmatic optimal approach for detection of cyber attacks using genetic programming," in *20th IEEE International Symposium on Computational Intelligence and Informatics, CINTI 2020 - Proceedings*, 2020, pp. 71–76, doi: 10.1109/CINTI51262.2020.9305844.
- [103] R. Zuech, J. Hancock, and T. M. Khoshgoftaar, "Detecting web attacks using random undersampling and ensemble learners," *Journal of Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-021-00460-8.
- [104] V. Kumar and D. Sinha, "A robust intelligent zero-day cyber-attack detection technique," *Complex and Intelligent Systems*, vol. 7, no. 5, pp. 2211–2234, 2021, doi: 10.1007/s40747-021-00396-9.
- [105] C. C. Ugwu, O. O. Obe, O. S. Popoola, and A. O. Adetunmbi, "A distributed denial of service attack detection system using long short term memory with singular value decomposition," in *Proceedings of the 2020 IEEE 2nd International Conference on Cyberspace, CYBER NIGERIA 2020*, 2021, pp. 112–118, doi: 10.1109/CYBERNIGERIA51635.2021.9428870.
- [106] S. U. Yang, "Research on network malicious behavior analysis based on deep learning," in *IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2021, pp. 2609–2612, doi: 10.1109/IAEAC50856.2021.9390796.
- [107] O. Aslan and A. A. Yilmaz, "A new malware classification framework based on deep learning algorithms," *IEEE Access*, vol. 9, pp. 87936–87951, 2021, doi: 10.1109/ACCESS.2021.3089586.

BIOGRAPHIES OF AUTHORS



Azar Abid Salih    is an assistant lecturer at department of information technology management, technical college of administration, Duhok Polytechnic University, Kurdistan Region-Iraq. Now he is Ph.D. student at Technical College of Informatics-Akre Duhok Polytechnic University. He received his B.Sc. at 2009 in computer science at University of Duhok. He holds M.Sc. degree at 2015 in computer science, University of Zakho. His research areas are machine learning, data science, cyber security, and data mining. He can be contacted at email: azar.abid@dpu.edu.krd.



Maiwan Bahjat Abdulrazzaq    is associate professor at department of computer science, faculty of science, University of Zakho, Kurdistan Region-Iraq. He received his Ph.D. in computer science with a specialization in machine learning in 2013. He has supervised and co-supervised more than 10 masters and 3 Ph.D. students. He has authored or coauthored more than 20 publications. His research interests include artificial intelligence, machine learning, data mining, data warehousing, and the internet of things. He can be contacted at email: maiwan.abdulrazzaq@uoz.edu.krd.