Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 6639714, 10 pages
https://doi.org/10.1155/2021/6639714

Research Article

The Anomaly- and Signature-Based IDS for Network Security Using Hybrid Inference Systems

Sajad Einy,[1] **Cemil Oz**,[1] **and Yahya Dorostkar Navaei**[2]

[1]*Computer Engineering Department, Sakarya University, Serdivan, Turkey*
[2]*Computer and Information Technology Engineering, Qazvin Islamic Azad University, Qazvin, Iran*

Correspondence should be addressed to Yahya Dorostkar Navaei; y.dorostkar@qiau.ac.ir

Received 30 December 2020; Revised 15 February 2021; Accepted 1 March 2021; Published 13 March 2021

Academic Editor: Ali Asghar Rahmani Hosseinabadi

With the expansion of communication in today's world and the possibility of creating interactions between people through communication networks regardless of the distance dimension, the issue of creating security for the data and information exchanged has received much attention from researchers. Various methods have been proposed for this purpose; one of the most important methods is intrusion detection systems to quickly detect intrusions into the network and inform the manager or responsible people to carry out an operational set to reduce the amount of damage caused by these intruders. The main challenge of the proposed intrusion detection systems is the number of erroneous warning messages generated and the low percentage of accurate detection of intrusions in them. In this research, the Suricata IDS/IPS is deployed along with the NN model for the metaheuristic's manual detection of malicious traffic in the targeted network. For the metaheuristic-based feature selection, the neural network, and the anomaly-based detection, the fuzzy logic is used in this research paper. The latest stable version of Kali Linux 2020.3 is used as an attacking system for web applications and different types of operating systems. The proposed method has achieved 96.111% accuracy for detecting network intrusion.

1. Introduction

In the last decade, the use of different types of networks (such as communication networks, social networks, mobile internet networks, and Internet of Things networks) has been welcomed by people all over the world, and many life affairs (such as buying and selling, medical consulting, business, and education) are done through these networks.

Providing security is one of the most important challenges for all of these networks, and different researchers have proposed different methods (such as firewalls, cryptography, and restricting access to the network) to overcome this challenge. The biggest weakness of all these methods is that they are not able to detect intrusions and attacks inside the network.

To eliminate this weakness, intrusion detection systems have been proposed that can detect internal and external attacks and intrusions. These systems monitor inbound network traffic to detect any abnormal behavior or abuse by users. These systems can play an important role in reducing the damage done by these intruders to the network structure or the data exchanged through this network due to the early detection of intrusions.

Intrusion detection systems can detect intrusions in two ways, including the following:

(a) Signature recognition: data related to the history of operations and transactions which are performed by authorized users and hackers in the networks are used, and patterns for normal behaviors and abnormal behaviors are created. By matching these patterns or signatures with the pattern of behavior of existing users, it is possible to identify unauthorized users or hackers from authorized users. The most important disadvantage of this method is that if the user's signature or pattern is new and there is no similar pattern or signature in the database before, we cannot determine whether it is an authorized or unauthorized user.

(b) Identify anomalies from machine learning methods (such as neural networks and SVM) to create a classification of users based on their behavior in the network (such as the amount of traffic generated, the content of files shared or downloaded, and network usage time). Two classes, normal and abnormal, are created to identify and classify users, based on which the system is trained. One of the advantages of this method is that it is possible to identify hackers who infiltrate the system with new methods.

The whole process performed in intrusion detection systems can be divided into two main parts. These parts are as follows:

(i) Selecting specific features from all the features registered and contained in the collected dataset: because the collected dataset about users' behaviors and interaction with different networks may include different types of features, therefore, using all of these features for identifying patterns or performing analyses is time-consuming and can be made complex or confuse the application of different techniques to detect the pattern of attackers and ordinary users, so to expedite the detection of hackers' abnormal behaviors in different networks, using selection techniques, a subset of the entire set of collected features will be created. For this matter, we select useful and effective features. In other words, in feature selection, we select a subset from the total features in the available datasets, and this subset can not only help reduce the time which is required to analyze these data and identify normal and abnormal behaviors but also greatly reduce the complexity of these data.

(ii) Using two techniques (including classification and anomaly detection) to separate normal users from attackers: all network users can be generally divided into two groups (normal users and attackers). Therefore, because the type and number of classes are known, it is possible to create patterns to identify the type of class of each user in the network environment by using different classification techniques (such as SVM and decision tree) and finally identify normal and abnormal behaviors of users in the network environment (such as creating abnormal traffic in the network).

This article is organized into six sections. In the second part, we will review some related works which are done by other researchers. In the third section, we will explain the proposed method. In the fourth section, we will show the implementation and evaluation results. In the fifth section, we will have more discussion about the proposed method and the results obtained. In the sixth section, we will bring the conclusion.

2. Related Works

Several kinds of research studies have been conducted by authors about the proposed techniques or methods for improving the accuracy of intrusion detection by IDSs. Furthermore, many authors have used feature selection methods for increasing precision in IDSs. In this work, we will be reviewing some of these research studies' goals and results.

Farha and Singh [1] studied about feature selection. They proposed an intelligent hybrid technique using the concept of metaheuristic optimization. The proposed technique combines IWD (intelligent water drops) and ACO (ant colony optimization) for selecting favorite features from the data. The main goal of this paper is to concentrate on reducing the training time and creating an optimal feature subset. In other words, these authors attempted to optimize the process of feature selection from the available data. Finally, these authors applied their proposed technique for the selection of features from KDDCUP 99 data, and the results showed that their proposed technique achieved its aim.

Azam Davahli and Abaei [2] proposed a new model using concepts of GA (genetic algorithm) and mathematical equations of GWO (grey wolf optimizer) for developing SVM- (support vector machine-) based LIDS (lightweight IDS). This new model is called GABGWO. The new model to increase the performance of the LIDS, by applying two new crossover and mutation operators, attempts to find the most relevant traffic features and eliminates irrelevant ones. The performance of GABGWO is evaluated, and the results show that, by using this model, we can choose optimal traffic, decrease computation cost, and obtain high accuracy for LIDS. Furthermore, the results show that the composition of GA and GWO had better performance in comparison with other recent methods and existing FS algorithms.

Al-Yaseen [3] used the support vector machine (SVM) and firefly algorithm (FA) for proposing a new wrapper feature selection method. The aim of this research was removing irrelevant and repetitive features to reduce the dimension of the data and reduce the time of classification and, finally, improving the performance of the IDS. FA was inapplicable in diverse combination problems, and this research was used to produce feature subsets. The SVM model was used for evaluating these feature subsets. The main advantage of the proposed method is the improvement of the ability of FA to produce an optimal feature subset. These authors used the NSL-KDD dataset for the evaluation of their proposed method. Experimental results showed that the proposed method had 78.89% overall accuracy of the IDS. Furthermore, these results showed that the proposed method (FA-SVM) can reduce the number of features, improve the accuracy of identifying intruders, and reduce the number of false alarm rates.

Hadeel Alazzam [4] proposed a new wrapper selection algorithm and a new method for binarizing a continuous

pigeon optimizer for the IDS. The new algorithm aims to utilize the selection process by using the pigeon-inspired optimizer. This author compared the new method with the traditional methods for binarizing continuous swarm intelligent algorithms. In this research, a new method of discretization of a continuous algorithm was proposed that was based on the usage of cosine similarity. For evaluating this proposed algorithm, the author used three popular datasets including UNSW-NB15, NLS-KDD, and KDDCUP 99. The results showed that the proposed algorithm can successfully reduce the number of features needed to build the IDS, and it had high TPR, FPR, accuracy, and *F*-score. Furthermore, the proposed algorithm reduced the required time for building the decision.

Yukang Liu [5] proposed a novel feature selection algorithm called MH_SFS (metaheuristic-based sequential forward selection). The main goal of this algorithm was to reduce the dimension of anomaly detection methods. The proposed algorithm can improve the performance of the SFS (sequential forward selection) feature selection algorithm. For decreasing the computational cost in addition to the metaheuristic search process, this author added a filter selection function. The author used the KDDCUP 99 dataset for testing their proposed algorithm and compared the performance of this proposed algorithm with traditional SFS and IFFS algorithms over three anomaly detection models. Results showed that the proposed algorithm had high performance and a smaller number of features selected in comparison with other algorithms.

Calvet et al. [6] studied the researches which are done by other researchers related to the combination of machine learning methods and metaheuristics. They proposed a novel type of hybrid algorithm with the help of the concept of learning heuristics. Constraint Optimization Problem Dynamic Inputs (COPDIs) attempt to solve combinatorial optimization problems. Usually, in these COPDIs, the problem inputs are not fixed; furthermore, they vary predictably. Therefore, the solution is made according to some heuristic-based iterative processes. These authors said that the coordination between the metaheuristic algorithm and the learning mechanism may be required for solving problems that occur due to variations in the input. The results of testing their proposed novel type of hybrid algorithms showed that the learning methods updated the input models iterationaly which were used by the metaheuristic.

Lin et al. [7] studied big data which were created from various IoT applications. They said that, for finding an optimal feature subset from these big data for applying classification methods, we need to have a metaheuristic search algorithm. They proposed a modified version of CSO (cat swarm optimization) and called it MCSO. The main goal of the proposed algorithm was to improve search efficiency within the problem space. Experimental results showed that MCSO, with a decreasing number of features in subfeatures in UCI datasets, improves the classification accuracy in comparison to the original CSO. Furthermore, MCSO increased training time and had higher accuracy in comparison with the original CSO. Therefore, MCSO can be applied to real-time IoT applications.

Yuyang Zhou [8] proposed a new IDS framework which was based on feature selection and ensemble learning techniques. For selecting an optimal subset of features, the author proposed a heuristic algorithm called CFS-BA. This algorithm used the correlation between features for selecting the optimal subset of features. Furthermore, the author combined forest, RF (random forest), and C4.5 by forest PA (penalizing attribute) algorithms. For combining the probability distribution of the base learners for attack detection, voting techniques were used. The algorithm was tested with the help of 3 different IDS datasets (including CIC-IDS-2017, AWID, and NSL-KDD datasets). The results showed that the proposed algorithm had better performance (accuracy: 99.81%) compared with the other approaches.

Opeyemi Osanaiye [9] combined three different filter methods (including ReliefF, chi-squared, and gain ratio) to propose a new feature selection method. Their main aim was reducing the system complexity and increasing the classification accuracy (using decision tree algorithm and J48). The proposed method made a subset with 14 important features out of 41 original features in the NSL-KDD dataset. The performance of the proposed method was evaluated by considering accuracy, detection rate, and false alarm rate. Experimental results showed that the proposed method can reduce effective features and have high classification accuracy and detection rate and less power consumption in comparison with other filtering methods.

Melike Günay [10] proposed the modification of FA-based feature selection. The author improved the performance of traditional FA by using a KNN classifier and adding an extra feature selection step. Four different datasets of IDS were used, various subfeatures were made for each dataset, and the performance of classification methods that are applied to these subsets was compared. The experimental results showed that the proposed FA can decrease the dimension of features, decrease memory usage (approximately 50%), and save the time of computations. Furthermore, the proposed FA improved the accuracy of classification.

To optimize the performance of intrusion detection systems, in [11], a new integrated learning approach has been proposed that is scalable and adaptive in nature. In this approach, a random forest machine learning algorithm was used to build the classified model. The proposed method was implemented with the NSL-KDD 40558 dataset, and its evaluation results showed that the accuracy of intrusion detection by applying this method was 99.91%.

Li et al. [12] proposed an improved krill swarm algorithm based on the linear nearest neighbor lasso step (LNNLS-KH) to solve problems related to the low performance and high number of false positives in intrusion detection systems. In this method, the linear nearest neighbor lasso step optimization was performed, which increased the accuracy of intrusion detection. The results of the implementations performed by these researchers showed that the LNNLS-KH algorithm selected an average of 7 features in the NSL-KDD dataset and 10.2 features in the CICIDS2017 dataset, and features that were not applicable were easily removed. This algorithm also performed well in the optimal

fitness iteration curve and convergence speed and showed a false positive rate.

To improve the classification accuracy and speed of the NNGE algorithm and reduce computational resource consumption, Li et al. [12] introduced a method of vertical and horizontal data reduction named as VHDRA (Vertical and Horizontal Intelligent Dataset Reduction Approach). VHDRA provides the following features:

(i) Reducing the properties in the dataset vertically by selecting the most important properties and over-reducing the NNGE rectangles

(ii) Tracking the mode and method of data extraction horizontally using a method called STEM, which is a method for monitoring states and data extraction

(iii) The results of the implementation of the proposed method showed that this method can reduce the number of features in the dataset and increase the accuracy of intrusion detection compared to other methods

Using the Unified Intrusion Anomaly Detection (UIAD) method, Kamarudin et al. [13] presented a new way to develop the function of intrusion detection systems to detect unknown attacks on web servers. The proposed method used three components (preprocessing, statistical analysis, and classification). Implementation of the proposed method using DARPA 1999 and ISCX 2012 datasets showed that it has 95% accuracy in detecting unknown attacks, and the detection rate of false alarms is about 1%.

Maple [14] conducted extensive studies on the use of intrusion detection systems in communication networks used in the Internet of Things (such as WSN, MANET, and CPS). The researcher said that most intrusion detection systems are not quite suitable due to the limitations of the resources used in the Internet of Things. Therefore, we need to develop appropriate intrusion detection systems for the Internet of Things.

Guoquan Li at al. [15] has reviewed some of the most important problems in using intrusion detection systems, creating a large number of false alarms and the accuracy of accurate intrusion detection. In this research, the methods of dealing with this problem have been researched. To this end, they focused on data fusion (DF) techniques to detect network intrusions and examined a set of benchmarks to compare the performance of DF techniques. In this study, the efficiency of many classification techniques (such as RF, C4.5, NN, and SVM) in creating more efficient detection systems was demonstrated.

Khraisat et al. [16] proposed a hybrid HIDS (combining the C5 decision tree classification and a support vector machine (SVM)). The purpose of this intrusion detection system was to overcome the problems of traditional intrusion detection systems in detecting known and unknown threats. They also proposed a framework aimed at identifying known intrusions, increasing the accuracy of intrusion detection, and reducing false alarm rates. In this framework, two methods were used to identify anomalies (to identify unknowns) and to identify signatures (for known threats). The proposed HIDS was evaluated using different datasets, and the evaluation results showed that it had better performance compared to SIDS and AIDS in terms of accuracy of diagnosis and low false alarm rate.

Snehi [17] discussed about different signature-based influence systems and their benefits. The results of their studies showed that a set of signatures and basic patterns show the relative importance of each feature of the intrusion detection system and help system administrators identify cyber attacks and network and computer system threats. Therefore, 80% of intrusions can be easily and quickly identified using signature-based detection methods.

David Mudzingwa [18] tried to study the methods used to create intrusion detection systems and intrusion prevention systems. The researcher focused more on anomaly-based, signature-based, state-of-the-art, and composition-based protocol analyses. The anomaly-based method was involved in detecting new threats without any updates or input for users, but most IDPSs on the market used a combination of several basic methods. This study also provided methods for easy comparison and evaluation of IDPS methods used by IDPS products in the market.

Liu Hua Yeo [19] provided an overview of different types of intrusion detection systems, how they are classified, and different algorithms used to identify unusual activities to operate in intrusion detection systems. The main focus of this research was on anomaly-based and signature-based intrusion detection systems. Then, the researcher tried to compare different methods of penetration techniques and also described different methods and the importance of IDSs in information security.

Singh [20] proposed a hybrid IDS by combining two approaches in one system. The hybrid IDS combining packet header anomaly detection (PHAD) and network traffic anomaly detection (NETAD) created using misuse-based IDS Snort, an open-source project, is a system based on anomaly detection. The proposed hybrid intrusion detection system has simulated using DARPA intrusion detection evaluation dataset that is available in MIT Lincoln laboratories repository. The results of this evaluation showed that the proposed system had a high efficiency in identifying intrusions created by creating anomalies and abuses.

3. Proposed Solution for Intrusion Detection Systems

In this research paper, the proposed solution for the network security of any organization against various types of attacks is given. In these attacks, the web application, operating system vulnerabilities, or any software vulnerabilities are considered. Structured Query Language (SQL) injection attack, Cross-Site Scripting (XSS) attack, broken authentication, and any payload for ransomware attacks are mostly focused. For this, the FreeBSD 12 Unix operating system is

used as a gateway or firewall along with Suricata as IDS/IPS. The two types of machine learning methods are used for the creation of Suricata signatures to block the malicious traffic on the targeted network. For the metaheuristic-based feature selection, the neural network, and the anomaly-based detection, the fuzzy logic is used in this research paper. The latest stable version of Kali Linux 2020.3 is used as an attacking system for web applications and different types of operating systems. In these client systems, Windows 10/7 and Ubuntu 20.04 are being used. To access web applications, Chrome, Firefox, and Edge browsers are utilized. This lab environment is deployed in VirtualBox for the proof of concept. The basic overview of any organization for its network security is given in Figure 1.

3.1. The Neural Fuzzy Logic Inference System. Here, in the proposed solution, the Suricata IDS/IPS is deployed along with the NN model for the metaheuristic's manual detection of malicious traffic in the targeted network. By utilizing this method for the detection of malicious traffic, the different attacks with higher identification proportion and lower false alert have been recognized. The NN model is one of the widely utilized methods and has been fruitful in tackling various composite reasonable inconveniences and different metaheuristic streamlining algorithms for advancement since firstly, it relies upon a somewhat direct start and is not hard to recognize. Secondly, it does not require exact information, and thirdly, it has the ability to avoid neighborhood optima. Fourthly, it can be utilized in an expansive extent of issues covering contrasting disciplines. Nature-pushed metaheuristic figuring settles upgrade issues by copying natural or physical miracles. They can be accumulated in three main functions: headway-based, material science-based, and swarm-based procedures. Progression-based procedures are impelled by the laws of trademark advancement. This method begins with indiscriminately made people which is categorized by ages. The quality reason for these systems is that the individuals are joined to the groups with same age and the manner of individuals ar changed by time. This empowers the general population to be redesigned in the wealth of the strategy for a very long time. ANN-based IDS exist in two points of view: I) lower revelation exactness, particularly for low-visit attacks, e.g., distant to the neighborhood and client to root; II) more fragile identification steadiness [21]. The fuzzy logic is used for the detection of anomaly-based attacks. The traffic with attachment files or unknown protocols will be processed by the fuzzy logic-based traffic anomaly detection. The detailed internal working of the proposed model is explained in Figure 2.

The patterns or signatures of various types of attacks related to web applications, operating systems (OSs), and any mobile apps will be trained in the NN model. After the processing of the trained data from the NN method, the new rule sets will be created for various types of malicious traffic such as SQL injection attacks, XSS attacks, broken authentication, insecure traffic, old versions of any application or framework, vulnerable OS, and any known vulnerability of mobile applications. The output of the NN model will be input to the Suricata IDS/IPS for preventing attacks on the targeted network. For the more efficient performance of the proposed solution, the open-source blacklist Internet Protocol (IP) address sources are used. The blacklist IP address sources are stored in the database at the frequency of updating these IP addresses after every 4 hours. The program has been used for storing blacklist IP which addresses the custom PHP and MySQL community version.

4. Implementation and Results

The lab has been deployed in a virtual environment for the proof of concept regarding the proposed solution for the IDS/IPS using ML methods. This has been deployed on VirtualBox open-source software to avoid damage to any real server or users. Due to this, the privacy of any web application users or operating system users is protected. FreeBSD 12.1 is deployed as a gateway and firewall for a virtual lab environment. On this, the proposed machine learning method for manual neural networks and anomaly fuzzy logic inference systems are installed. Along with this, the Suricata IDS/IPS, FreeBSD built-in packet filtering (PF), and customized program developed in PHP for blacklist IP address databases from different open-source resources are used. Ubuntu 20.04 is deployed for the Damn Vulnerable Web Application (DVWA) [22] for testing the proposed solution in this research paper. This web application has all the top ten vulnerabilities along with advanced SQL injection attacks. It is depicted in Figure 3.

The Windows 10 client system is also installed for securing it against attacks from the hackers through Suricata IDS/IPS by implementing the ML methods. The hacking machine is also deployed with the installation of the Kali Linux 2020.3 version. This will be used for exploiting OS and web application vulnerabilities.

To launch an attack on the DVWA website and on Windows 10 OS, Kali Linux is used. With the help of Kali Linux, the security signature of Suricata IDS/IPS is created by using ML methods. In this process, two kinds of techniques are used. One is a neural network for well-known attacks such as SQL injection, XSS, broken authentication, and OS system vulnerabilities. The SQL injection attack is launched with the help of the SQLMAP open-source tool in Kali Linux. In the first command, the database name is extracted from the target web application, but due to the Suricata IDS/IPS as a firewall, it is not possible to obtain the database name or enumerate the targeted web application regarding database information as shown in Figure 4. Secondly, the usernames and passwords are extracted from this database of the web application. The result of advanced queries of SQL injection by using –temper or– random-agent has not been included here because it is unable to get required information due to repetition of the same type of result in a paper image of that result.

The XSS attack is launched with the help of the XSSer tool of pen-testing in Kali. Due to the proposed firewall system, this attack is also blocked as an error of 302 or 301 as is highlighted in Figure 5. Because of the failed attack, it

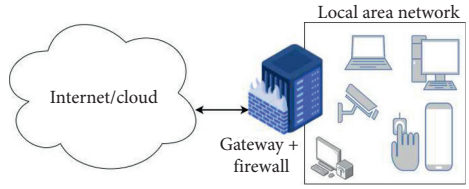

Figure 1: The basic overview of network security.

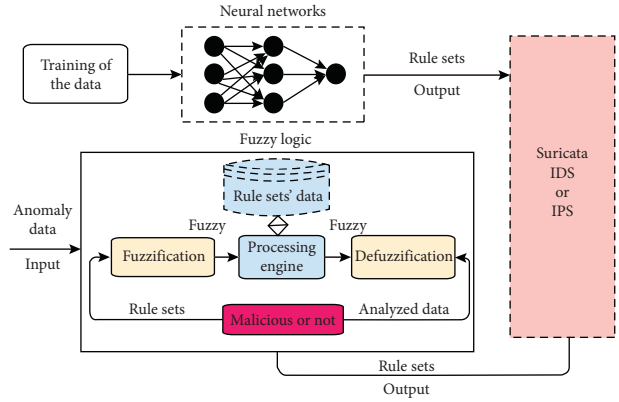

Figure 2: Internal working of the proposed solution for the IDS/IPS.

has generated a few errors also. The Metasploit tool has been used for exploiting the Windows 10 client systems. In this paper, the main target is to exploit the Simple Message Block (SMB) CVE-2018-0749 vulnerability. The Common Vulnerabilities and Exposures (CVE) (SMB) [23] is utilized to look into common vulnerabilities of Windows client systems. The target system is installed in a virtual environment to avoid harm to any real client. The attack launch on the target system is shown in Figure 6. This attack also fails due to the deployed firewall as security with the utilization of ML methods to generate the Suricata signatures.

The core work in this research paper is the implementation of Suricata IDS/IPS for the security of any organization network. This security will be from layer 4 to layer 7 applications. The novelty in this paper is that it is deployed on FreeBSD 12.1, the best Unix operating system for firewalls, which itself has the best packet filter (PF) firewall functionality. This PF will work for network-level filtering of malicious traffic or which administrator wants to block. The known attack's signatures for Suricata are generated with the help of this method of metaheuristic neural network. In this method, the data regarding web application and operating system attacks will be trained into the NN for the signatures, and these will be input into Suricata. Secondly, the fuzzy logic is used for the anomaly-based traffic on the network. In this process, unknown attacks or any attached files will be analyzed by the fuzzy logic systems. After analysis, this system will decide whether the traffic is malicious or legitimate. If it is malicious, the signature for the IDS/IPS will be generated to protect the clients of the targeted organization. The few malicious types of data are highlighted in red color in this raw logs JavaScript Object Notation (JSON) format as shown in Figure 7.

These JSON format logs are converted into XML, and from this format, these are converted into a tabular form. In Table 1, the success of full login or access to other resources at the network is shown.

In Table 2, the data are presented from the Suricata IDS/IPS logs. These logs are related to many types of events that have generated at a firewall in which remote login on Windows client system and web application. In these attacks, there are SMB exploits, SQL injection attack, XSS, brute force attack on web application login.

4.1. Confusion Matrix of the Proposed Hybrid Inference System. The confusion matrix is known as the error matrix for the analysis of statistical data. As in this paper, the hybrid inference system is used for the signature generation and anomaly-based traffic analysis for the Suricata intrusion detection system or intrusion prevention systems. Various types of attacks have been launched in the locally developed lab such as SQLi, XSS, Exploitation of Windows OS vulnerabilities, and distributed denial-of-service (DDoS) attack on the targeted network. These mentioned attacks are prevented by the proposed system for the targeted network. For the evaluation of results, the confusion matrix has been used as defined in Table 3.

The system accuracy of 96.11% has not been achieved due to the false positive ratio of network traffic detection by the Suricata IDS/IPS. As the results of the proposed solution As the results of the proposed solution in this paper, it shows overall good accuracy of detection or prevention of attacks on the targeted network. The limitations of this proposed system can be mentioned as per limited type of datasets. Furthermore, this proposed system is deployed within a

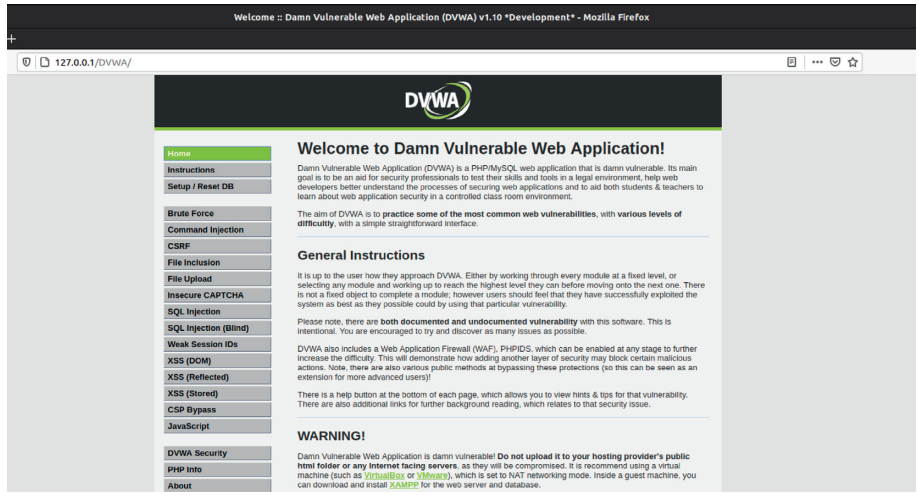

Figure 3: The vulnerable web application deployed in the virtual lab.

Figure 4: The SQL injection attack by using SQLMAP.

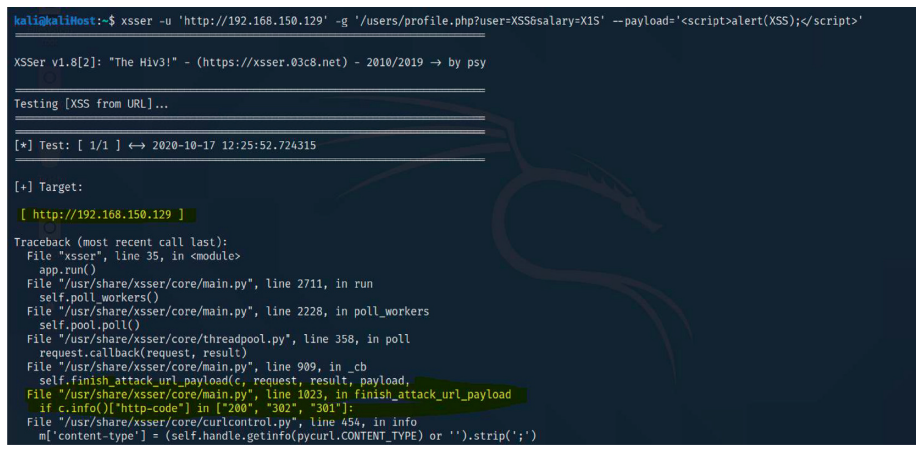

Figure 5: The XSSer tools for the XSS attack on the target web application.

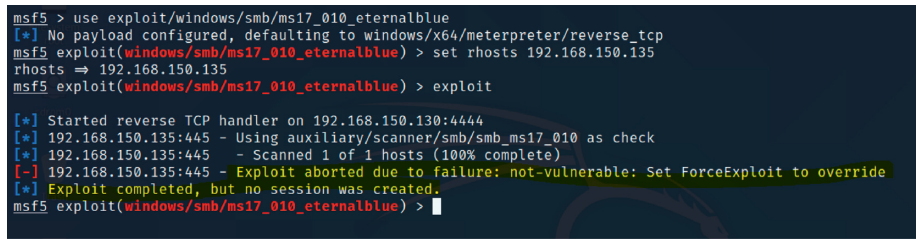

Figure 6: Attack launched on the Windows client.

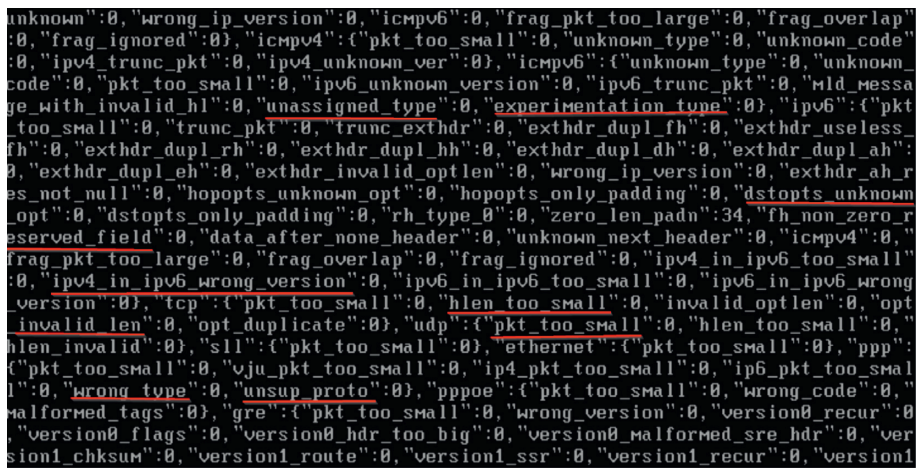

Figure 7: JSON format logs of Suricata IDS/IPS.

Table 1: Successful user login on Windows 10 client.

Action	Date and time	Status	Source IP	Destination IP	Username	Risk level
Logged in	Oct 01, 2020, 1:24:43 AM	User login success	192.168.150.128	192.168.150.135	N/A	5
Logged in	Oct 01, 2020, 1:20:03 AM	User login success	192.168.150.128	192.168.150.135	N/A	5
Logged in	Oct 01, 2020, 1:13:13 AM	User login success	192.168.150.128	192.168.150.135	N/A	5
Logged in	Oct 01, 2020, 1:07:13 AM	User login success	192.168.150.128	192.168.150.129	N/A	5
Logged in	Oct 01, 2020, 1:00:53 AM	User login success	192.168.150.128	192.168.150.129	N/A	5

Table 2: User login fail attempts of Windows client and web application.

Action	Suricata IDS/IPS	Date and time	Status	Source IP	Destination IP	Username or attack type	Risk level
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:53 AM	Failure	89.46.223.240	192.168.150.135	POSTGRES	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:43 AM	Failure	118.107.76.23	192.168.150.135	TEST1	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:43 AM	Failure	192.168.150.128	192.168.150.135	ADMINISTRATOR	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:43 AM	Failure	212.42.214.3	192.168.150.135	LENOVO	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:43 AM	Failure	124.109.54.218	192.168.150.129	SQLi	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:43 AM	Failure	192.168.150.128	192.168.150.129	SQLi	4
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:43 AM	Failure	212.42.214.3	192.168.150.129	XSS	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:34 AM	Failure	192.168.150.128	192.168.150.129	XSS	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:34 AM	Failure	212.42.214.3	192.168.150.129	SQLi	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:34 AM	Failure	112.161.27.203	192.168.150.129	SERVER	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:34 AM	Failure	192.168.150.128	192.168.150.129	ADMINISTRATOR	4
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:34 AM	Failure	212.42.214.3	192.168.150.129	MASTER	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:34 AM	Failure	192.168.150.128	192.168.150.129	ACER	3
Failed to log in	192.168.150.131	Oct 01, 2020, 1:29:34 AM	Failure	89.46.223.240	192.168.150.129	Web applicatiin	3

Table 3: The confusion matrix of the hybrid inference system for the IDS.

	Normal	SQLi	XSS	Windows OS attacks	DDoS	Classification overall	Precision (%)
Normal	500	0	0	0	0	500	100
SQLi	0	900	50	0	50	1000	90
XSS	0	0	925	25	50	1000	92.5
Windows OS attacks	0	0	0	1000	0	1000	100
DDoS	0	0	0	0	1000	1000	100
Truth overall	500	900	975	1025	1100	4500	
User accuracy (%)	100	100	94.872	97.561	90.909		

Overall accuracy = 96.111% is obtained.

virtual system lab created for weak operating systems and applications on which these tests were applied. The run time of the algorithm is within an expected response from the proposed system. For big datasets, the delay may be faced by this system, but it can optimize in future to improve the response time within the required time limits. As per the strategy proposed in the past areas of this paper, the initially produced fuzzy model on the extrema data of each solutions. To accomplish a decent harmony between the framework execution and rule intricacy, distinctive fix sizes and similarity limits are attempted. The subsequent fluffy frameworks utilizing diverse fix sizes are assessed by rule intricacy and the nearby guess mistake. This research has different types of datasets and attack types which were tested on the proposed system. To the best of authors' knowledge, this type of work has not been done before.

5. Conclusions

Nowadays, with the advancement of science and technology and the increase in the level of awareness and knowledge of human beings, as seen day by day, there are widespread attacks on various commercial, communication, information, and other networks. Some of these attacks target important and sensitive data and information contained in these networks, and some other attacks try to generate extra traffic in the communication routes with the resources and servers which are providing different services for users; thus, they block the access of authorized users to those services either making the servers unavailable or disrupting the network performance for several minutes or hours. One way to identify hackers is to use intrusion detection systems (IDSs) that try to detect or predict the possibility of attacks in various ways. One of these ways is to check the network traffic and try to identify the attacks by creating a lot of traffic in the network. In this paper, we considered this type of attack and proposed a new method for deploying Suricata IDS/IPS with the help of the NN model for the meta-heuristic's manual detection of malicious traffic in the targeted network. In the proposed solution, the traffic with attachment files or unknown protocols will be processed by the fuzzy logic as based on detected as anomaly traffic. For implantation of the proposed solution for the IDS/IPS using ML methods, a virtual environment is deployed. The results showed that our proposed method can detect different types of attacks (including in these attacks, SMB exploits, SQL injection attack, XSS, and brute force) on web application login. The overall accuracy of our proposed method is about 96.11% when experimented on the particular dataset.

Data Availability

The simulated data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Farha and S. Singh, "Feature selection technique for intrusion detection system based on iwd and aco," *International Journal of Advanced Research in Computer Science*, vol. 8, no. 9, pp. 270–275, 2017.

[2] M. S. Azam Davahli and G. Abaei, "A lightweight Anomaly detection model using SVM for WSNs in IoT through a hybrid feature selection algorithm based on GA and GWO," *Journal of Computing and Security*, vol. 7, no. 1, pp. 63–79, 2020.

[3] W. L. Al-Yaseen, "Improving intrusion detection system by developing feature selection model based on firefly algorithm and support vector machine," *IAENG International Journal of Computer Science*, vol. 46, no. 4, pp. 1–7, 2019.

[4] A. S. Hadeel Alazzam, "A feature selection algorithm for intrusion detection system based on Pigeon Inspired Optimizer," *Expert Systems With Applications*, vol. 148, pp. 1–14, 2020.

[5] Z. X. Yukang Liu, "A novel meta-heuristic-based sequential forward feature selection approach for anomaly detection systems," in *Proceedings of the 2016 International Conference on Network and Information Systems for Computers*, pp. 218–227, Wuhan, China, April 2016.

[6] L. Calvet, J. D. Armas, D. Masip, and A. A. Juan, "Learn-heuristics: hybridizing metaheuristics with machine learning for optimization with dynamic inputs," *Open Mathematics*, vol. 15, no. 1, pp. 261–280, 2017.

[7] K.-C. Lin, Y.-H. Huang, J. C. Hung, and Y.-T. Lin, "Feature selection and parameter optimization of support vector machines based on modified cat swarm optimization," *International Journal of Distributed Sensor Networks*, pp. 1–9, 2014.

[8] G. C. Yuyang Zhou, "Building an efficient intrusion detection system based on feature selection and ensemble classifier," *Computer Networks*, vol. 174, pp. 1–21, 2020.

[9] O. O. Opeyemi Osanaiye, "Feature selection for intrusion detection system IN a cluster-based heterogeneous wireless sensor network," *Facta Universitatis*, vol. 32, no. No 2, pp. 315–330, 2019.

[10] Z. O. Melike Günay, "A modified firefly algorithm-based feature selection method and artificial immune system for intrusion detection," *Uludağ University Journal of The Faculty of Engineering*, vol. 25, no. 1, pp. 269–287, 2020.
[11] A. Chiche and M. Meshesha, "Towards a scalable and adaptive learning approach for network intrusion detection," *Journal of Computer Networks and Communications*, vol. 2021, Article ID 8845540, 9 pages, 2021.
[12] X. Li, P. Yi, W. Wei, Y. Jiang, and L. Tian, "LNNLS-KH: a feature selection method for network intrusion detection," *Security and Communication Networks*, vol. 2021, Article ID 8830431, 22 pages, 2021.
[13] M. H. Kamarudin, C. Maple, T. Watson, and N. S. Safa, "A new unified intrusion anomaly detection in identifying unseen web attacks," *Security and Communication Networks*, vol. 2017, Article ID 2539034, 18 pages, 2017.
[14] Z. A. Maple, "recent advancements in intrusion detection systems for the Internet of Things," *Security and Communication Networks*, vol. 2019, Article ID 4301409, 19 pages, 2019.
[15] Z. Y. Guoquan Li, Z. Yan, Y. Fu, and H. Chen, "Data fusion for network intrusion detection: a review," *Security Measurements of Cyber Networks Issue*, vol. 2018, p. 18, Article ID 821061, 2018.
[16] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman, and A. Alazab, "Hybrid intrusion detection system based on the stacking ensemble of C5 decision tree classifier and one class support vector machine," *Electronics*, vol. 9, no. 1, p. 173, 2020.
[17] J. Snehi, "Diverse methods for signature based intrusion detection schemes adopted," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 9, no. 2, pp. 44–49, 2020.
[18] R. A. David Mudzingwa, "A study of methodologies used in intrusion detection and prevention systems (IDPS)," *Proceedings of IEEE Southeastcon*, 2012.
[19] X. C. Liu Hua Yeo, "Understanding modern intrusion detection systems: a survey," *Information & security: An International Journal*, vol. 46, no. 2, pp. 155–167, 2020.
[20] S. Singh, "A hybrid intrusion detection system design for computer network security," *International Journal of Engineering Sciences & Research Technology*, vol. 7, no. 4, pp. 339–343, 2018.
[21] R. Beghdad, "Critical study of neural networks in detecting intrusions," *Computers & Security*, vol. 27, no. 5, pp. 168–175, 2008.
[22] DVWA. (2020). Damn Vulnerable Web Application. http://www.dvwa.co.uk/.
[23] SMB T M (2020). https://www.cvedetails.com/cve/CVE-2020-0749/.