ISSN: 2229-6948(ONLINE)
DOI: 10.21917/ijct.2022.0393

MULTI-OBJECTIVE WHALE OPTIMIZED WITH RECURRENT DEEP LEARNING FOR EFFICIENT INTRUSION DETECTION IN HIGH SENSITIVE NETWORK TRAFFIC

P. Roshni Mol and C. Immaculate Mary
Department of Computer Science, Sri Sarada College for Women, India

Abstract

Intrusion detection plays a pivotal aspect in providing security for the information and the main technology lies in identifying different networks in an accurate as well as precise manner. By swift technological development, in recent years network systems are becoming highly susceptible to numerous revolutionary intrusion types. However, Deep learning based models are significant with accessible technique for detecting network intrusions of high network traffic. In this work, a method called, Multi-objective Whale Optimized with Recurrent Deep Learning (MWO-RDL) classifier for intrusion detection in high sensitive network traffic is proposed. With high sensitive network traffic as input, initially, Robust Scaler and Multi-objective Whale Optimization-based feature selection is designed for developing feature selection procedure as well as improving accuracy and finally the intrusion detection. The main idea behind the design of model is employed for assessing chosen feature subset that was explored in specified exploration space. Next, a classifier model called, Discrete Mutual Information-based Recurrent Deep Neural Learning is applied to the optimal selected features for classifying according to the characteristics of network traffic features into different type of attacks, normal traffic. Multi-objective Whale Optimized with Recurrent Deep Learning (MWO-RDL) is very suitable for modeling an intrusion detection model with high classification accuracy, intrusion detection rate and that its performance is comparatively better to that of traditional deep learning classification methods in multiclass classification. The MWO-RDL method minimizes false alarm rate of intrusion detection in a timely manner and bestows a novel research of high network traffic.

Keywords:

Deep Learning, Multi-objective Whale Optimization, Robust Scaler, Recurrent Deep Learning, Discrete Mutual Information

1. INTRODUCTION

Intrusion Detection System (IDS) is set off influential instrument that keeps track of malicious activities and activates alerts for suspicious attack detection. The intrusions make the system uncertain for network traffic owing to its nonlinear behavior. Also, IDSs have become significant and extensively utilized instruments for providing network security. In recent years, intrusion detection based on optimal classification methods has fascinated a broad span of attentiveness in retaliation to the increasing insistence of genuine and intelligent IDSs that are essential to detect advanced and discrete intrusion attacks.

Owing to 26 billion associated devices spread in a heterogeneous network, attacks are also making a swift change in a rapid manner. Also, cyber threats have also become notorious menace with the over dependence of government, military and commercial establishments even for their day-to-day chores.

Upon comparison to the machine learning techniques algorithms employing deep learning have proved their efficiency in intrusion detection. Attack detection using deep learning LSTM-PCA was proposed in [1]. Here, Principal Component Analysis (PCA) and Mutual information (MI) were utilized as dimensionality reduction and feature selection. Following which, accurate classification was made.

With the swift evolution in network traffic in the recent few years has resulted in the significance of flow-based intrusion detection. Auto Encoder and Variational Auto Encoder (AE-VAE) was proposed in [2] that concentrated on network traffic anomalous detection on unsupervised deep learning with semi supervised learning model.

Moreover, methods for intrusion detection based on anomaly that can identify unknown attacks were also combined. To be more precise, Auto Encoder and Variational Auto Encoder were employed in identifying unknown attacks with the aid of flow features and validity was ensured using Receiver Operating Characteristic (ROC) curve.

Motivated by the above said facts, in our work intrusion detection in highly sensitive network traffic is made by means of optimization-based feature selection model and deep classification into different types of attacks.

The contributions of the work include the following:

- To ensure data transformed-feature selection and classifier offering effective as well as precise intrusion detection.
- On the contrary data transformation and feature selection, Multi-objective Whale Optimization based Feature selection (Preprocessing) model was employed for determining correlation of chosen sub-features as well as advantageous to obtain effectiveness of training and testing phase.
- To enhance multi-objective classification, Recurrent Deep Neural Learning algorithm is developed by combining decisions to exhibit temporal dynamic behavior.
- The proposal was associated by existing methods on extensive test bed comprising of dataset namely, UNSW-NB15 intrusion detection dataset. Experiment results shows that the proposed solution excels parallel methods with various parameters.

The article is summarized by. Section 2 reviews related works of intrusion detection. Section 3 explains description of Multi-objective Whale Optimized with Recurrent Deep Learning (MWO-RDL) method in the area of intrusion detection, especially how deep learning methods facilitate the development of intrusion detection. Section 4 introduces Performance evaluation measures. Section 5 highlights MWO-RDL with a discussion about the experimental results. Section 6 describes the conclusion of this article.

2. RELATED WORKS

Network Intrusion Detection Systems (NIDS) are paramount in today's computing structure to assist track and identify obtain unnecessary and malicious network traffic. A novel methodology to detect malicious network traffic for utilization of deep packet IDS was proposed in [3]. With the deep packet inspection, the false positive rate was reduced drastically. One of the main problems of protection patterns was the false alarm that occur high network traffic data.

A novel reliable hybrid method by ABC was presented in [4] for ensure higher DR and lower FPR. With the swift enhancement in technology, network systems are found to be highly susceptible to many intrusion types. However, algorithms employing machine learning (ML) are found to be the most effective methods in identifying intrusions. In [5], predominant and pertinent features were identified. Followed by which, SSPLR and SVMs were employed therefore contributing to high intrusion detection.

In recent days, utilization of Internet assists in numerous fields like, e-commerce, digital marketing and entertainments. However, as far as cyber security is concerned, it has become highly susceptible to enormous advancement of expeditious inception. With unprecedented network technologies and numerous intrusion methods, conventional machine learning algorithms emerge incompetent.

An anomaly-based intrusion detection model employing convolutional neural networks was proposed in [6]. Similarity among IDS application association rule mining as well as SVM was presented in [7] with the objective of improving the accuracy. Though several supervised and unsupervised learning methods have been applied so far for detecting intrusion achieving good performance still remained a great question. To address this issue, an integration of feature selection and ensemble classifier was applied in [8] for attack recognition. Long Short Term Based Anomaly Detection (LSTBAD) via statistical strategy to minimize prediction error was proposed in [9]. Yet another hybrid network intrusion detection method employing machine learning algorithms to attain good accuracy as well as minimize false alarm rate was designed in [10].

A systematic review on intrusion detection on deep learning was investigated in [11]. IDS provide one of the domains where neural networks are extensively validated for enhancing comprehensive security and privacy in computer network. An elaborate overview of recent literature concerning neural networks utilization in intrusion detection system was investigated in [12]. In [13], unsw-nb15 dataset was designed for analysis. Host based IDS was investigated in [14].

Hybrid data optimization methods were employed in [15] to obtain optimal feature subset and therefore improving detecting rate anomaly behaviors. However, with high class imbalance network traffic feature, intrusion detection has still to be improved. Supervised machine learning to address imbalance class issues was introduced in [16] that with the aid of Extremely Randomized Trees Classifier detected the attacks separately with elevated accuracy and minimum false alarm rate. For advanced metering infrastructure, Online Sequence Extreme Learning Machine was proposed in [17]. Deep learning was applied in [18] [19] to improve intrusion detection accuracy.

Based on the aforementioned materials and methods, in this work, Multi-objective Whale Optimized with Recurrent Deep Learning (MWO-RDL) classifier is proposed and elaborated in the forthcoming sections.

3. METHODOLOGY

With the objective of improving the detection ability of IDS in highly sensitive network traffic, we propose an efficient Multi-objective Whale Optimized with Recurrent Deep Learning (MWO-RDL) classifier based optimal feature selection model and deep learning classifier. During the experiments, MWO-RDL technique was validated and classifies traffic as well as different attacks. The Fig.1 illustrates proposed Multi-objective Whale Optimized with Recurrent Deep Learning (MWO-RDL) classifier for intrusion detection in high sensitive network traffic.

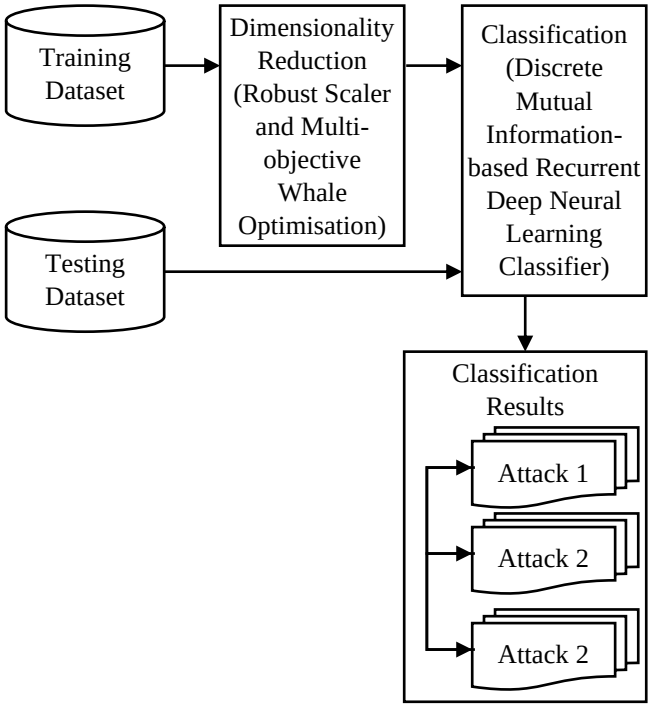

Fig.1. Multi-objective Whale Optimized with Recurrent Deep Learning (MWO-RDL)

As shown in the Fig.1, the proposed MWO-RDL consists of two phases. Data transformation as well as feature selection is achieved for dimensionality reduction using Multi-objective Whale Optimization model. Followed by which with the optimal feature selection, a Recurrent Deep Learning classifier is applied for classifying different types of attacks or normal for high sensitive network traffic. The proposed MWO-RDL method is illustrated in following sub-sections.

3.1 ROBUST SCALER AND MULTI-OBJECTIVE WHALE OPTIMIZED FEATURE SELECTION MODEL

Present day intrusion detection datasets [20] and method [1] certainly hold loss of unnecessary and irrelevant network traffic features [2] that in turn reduces the efficiency of network traffic and detect incomprehensible results. Hence, the initial step in our

proposed method is used for minimizing feature as well as chosen feature subset of dataset [20]. Thus, the proposed MWO-RDL method uses feature selection named Robust Scaler and MWO is used for minimizing amount of network traffic and utilized novel model termed as Recurrent Deep Neural Learning for classifying data. Besides, arbitrary fitness function is employed for minimizing features to ensure lower false alarm rate. The Fig.2 shows the structure of named Robust Scaler and Multi-objective Whale Optimization model.

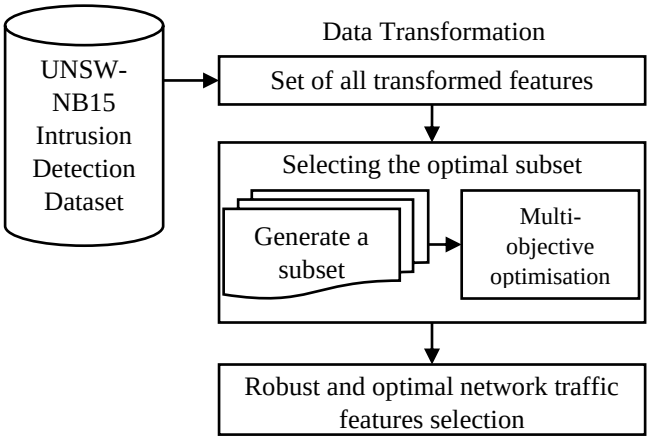

Fig.2. Structure of Robust Scaler and Multi-objective Whale Optimization

To start with, initially, the raw data is transformed into system appropriate using dataset. From data preprocessing, feature transformation and feature selection operation is performed on UNSW-NB15 intrusion detection dataset [20]. With this preprocessing the underlying data structure can be exposed in a better manner to respective algorithm and performance of robust predictive.

In the data or feature transformation process, nominal values present in the dataset. Owing to fact that IDS are contemplated as classification issue and certain classification techniques cannot handle nominal features. Certain features in the UNSW-NB15 intrusion detection dataset like, attack_cat, protocol_type, state, etc., are transformed from nominal to numeric values and the final UNSW-NB15 intrusion detection dataset contains the entire numeric values for the classification.

On the other hand, the feature selection process minimizes the number of input variables. Feature selection process is preferable for minimizing amount of input features for lesser time involved in modeling as well as enhances intrusion detection. The feature selection is performed owing to the high dimensional nature of dataset that in turn reduces the dimensionality of dataset and also selects the most pertinent network traffic features for each type of attack.

Let us consider dataset *DS* comprising of network set S_i, $\forall i$=1,2,3,…n with a class label C_j, $\forall j$=1,2,3,…m. Here 'm' represent amount of classes whereas 'n' indicates amount of instances with 'F={f_1,f_2,f_3,…,f_n}' representing the number of features in S_i. This is mathematically expressed as given below.

$$DS=\{S_i,C_j\},\forall i=1,2,3,\ldots n;\ \forall j=1,2,3,\ldots m \quad (1)$$

Then, with the above dataset *DS* representation inclusive of network set S_i and class label C_j, to map the input feature value to the same scale, Robust Scaler Transformation is performed as given below.

$$V' = \frac{F(V) - F(V_{\min})}{F(V_{\max}) - F(V_{\min})} \quad (2)$$

From the above Eq.(2), the resultant feature transformed or the normalized value V' is obtained based on the attribute value for each feature $F(V)$, minimum attribute for every feature $F(V_{min})$ as well as maximum attribute for every feature $F(V_{max})$ respectively. Next, with the feature transformation, feature selection is performed to identify subset of features from the feature transformed set that are representative for feature and the subset attributions form highly relevant to intrusion detection. In the Multi-objective Whale Optimization process of optimal selection of network traffic features, initialization of network traffic features is done. This is mathematically expressed as given below.

$$F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1n} \\ F_{21} & F_{22} & \cdots & F_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ F_{m1} & F_{m2} & \cdots & F_{mn} \end{bmatrix} \quad (3)$$

From the above Eq.(3), m refers to the number of whales (i.e., network traffic features) and n represents the dimensionality respectively. Next, the fitness of whale population or the network traffic features are evaluated. Then, for the corresponding network traffic feature two fitness is said to prevail. Then, for two objectives Obj_1 (i.e., maximizing intrusion detection rate) and Obj_2 (i.e., minimizing intrusion detection time) the evaluation is formulated as given below.

$$ObjF = \begin{bmatrix} Obj_1(F_1) & Obj_2(F_1) \\ Obj_1(F_2) & Obj_2(F_2) \\ \vdots & \vdots \\ Obj_1(F_n) & Obj_2(F_n) \end{bmatrix} \quad (4)$$

Then, with the two objective functions as given in above (4), the position update for selecting optimal feature is mathematically formulated based on arbitrary value instead of best value as given below.

$$PU=[CV_1\ AP(t)\text{-}CP(t)] \quad (5)$$

$$CP(t+1) = [(AP(t)\text{-}CV_2\ PU)] \quad (6)$$

From the Eq.(5) and Eq.(6), the optimal position update *PU* for selecting optimal network traffic features is evaluated based on the arbitrary position of network traffic obtained so far $AP(t)$, current position of network traffic $CP(t)$, coefficient vectors CV_1 and CV_2 respectively. The pseudo code representation of Robust Scaler and Multi-objective Whale Optimized Feature Selection is given below.

Algorithm 1: Robust Scaler and Multi-objective Whale Optimized Feature Selection

Input: Dataset *DS*, Network Traffic Features *NTF*= {*FF*,*BF*,*CF*, *TF*,*AGF*}, Flow Features *FF*={FF_1,FF_2,…,FF_n}, Basic Features *BF*={BF_1,BF_2,…,BF_n}, Content Features *CF*={CF_1,CF_2,…, CF_n}, Time Features *TF*={TF_1, TF_2,…,TF_n}, Additional Generated Features *AGF*={AGF_1,AGF_2,…,AGF_n}

Output: Optimal features '*OF*={OF_1,OF_2,…,OF_n}'

1: Initialize network set S_i, $\forall i$=1,2,3,…n, class label C_j, $\forall j$=1,2,3,…m.

2: Begin

3: For each Dataset *DS* with Network Traffic Features 'NTF'

4: Formulate dataset as in Eq.(1)

//Data (feature) transformation

5: Perform data or feature transformation as in Eq.(2)

//Feature selection

6: Initialize whale (network traffic feature) population as in Eq.(3)

7: Formulate objective function as in Eq.(4)

8: Estimate position update for selecting optimal feature as in Eq.(5) and Eq.(6)

9: Check if the traffic feature attributes go beyond the exploration space

10: Recompute the fitness of updated traffic feature attributes

11: Update if there are better optimal position update

12: Return (Features selected)

13: End for

14: End

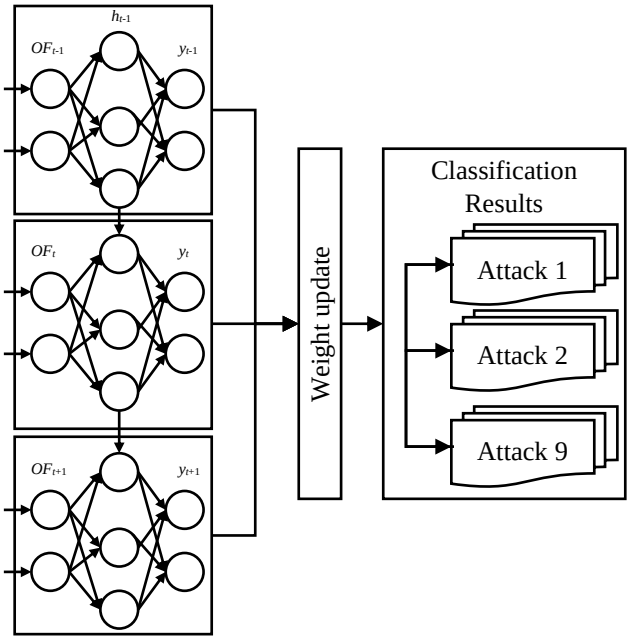

Fig.3. Structure of Discrete Mutual Information-based Recurrent Deep Neural Learning Classification model

As given in the above algorithm, enhancing or maximizing intrusion detection rate and minimizing intrusion detection time, a preprocessing model employing Robust Scaler function and Multi-objective Whale Optimization technique is used. With the UNSW-NB 15 network traffic dataset provided as input, two distinct processes, namely, feature transformation and feature preprocessing are performed. By employing Robust Scaler function for data transformation, network traffic features are mapped to the same scale so that ease in feature subset selection is ensured. Next, with the multi-objective, i.e., improving intrusion detection rate and time, optimal feature selection is made on the basis of the best position or the optimal feature for further classification.

3.2 DISCRETE MUTUAL INFORMATION-BASED RECURRENT DEEP NEURAL LEARNING CLASSIFIER

In this section, with the obtained optimal network traffic features, an efficient intrusion detection model in highly sensitive network traffic using Discrete Mutual Information-based Recurrent Deep Neural Learning Classifier model is designed. Recurrent Deep Neural Learning Classifier, nine distinct types of attacks and normal traffic are classified with high accuracy rate.

Given a sequence of inputs or optimal features $OF=\{OF_1,OF_2,\ldots,OF_n\}$, the Discrete Mutual Information-based Recurrent Deep Neural Learning estimates the sequence of hidden states $h=h_1,h_2,\ldots,h_n$ and a sequence of predictions or classification results $y=y_1,y_2,\ldots,y_n$ by learning the given below equations in an iterative fashion. First, the inputs to the hidden states are mathematically expressed as given below.

$$IH_i = W_{hOF}OF_i+W_{hh}h_{i-1}+b_h \quad (7)$$

Next, with the resultant inputs to the hidden states network traffic features, sequence of hidden states are estimated as given below.

$$h_i=SIGMOID(IH_i) \quad (8)$$

Followed by which the inputs to the output units are measured using the bias value as given below.

$$IO_i = W_{yh}h_i+b_y \quad (9)$$

Finally, sequence of network traffic predictions either with the presence of any types of attack or normal traffic are estimated as given below:

$$y_i=SIGMOID(IO_i) \quad (10)$$

From the above Eq.(7)-Eq.(10), 'W_{hOF}' represents the optimal features (i.e., input network traffic features) to hidden weight, 'W_{hh}' represents hidden to hidden weight, 'W_{yh}' represent hidden to output weight. Moreover, 'b_h' with 'b_y' denotes biases for hidden output layer. Finally, the activation function is represented in the form of '$SIGMOID$' and the classification function in the form of '$SOFTMAX$'.

From the above Eq.(7)-Eq.(10), W_{hOF} represents the optimal features (i.e., input network traffic features) to hidden weight, W_{hh} represents hidden to hidden weight, W_{yh} represent hidden to output weight. Moreover, b_h with b_y denotes biases for hidden output layer. Finally, the activation function is represented in the form of *SIGMOID* and the classification function in the form of *SOFTMAX*.

The objective function associated with Recurrent Deep Neural Learning for network traffic training pair (OF_i,y_i) is defined as $f(\theta)=MTI(OF,y)$, where *MTI* is the mutual traffic information which measures the deviation of the predictions y_i from the actual labels IF_i respectively. Let α be the learning rate and i represents the iterations. Then, the weight updates (i.e., network traffic updates) is estimated based on the mutual traffic information and is mathematically formulated as given below.

$$MFI\left(OF,y\right)=\sum\sum\left(OF,y\right)\log\frac{\text{Prob}\left(OF,y\right)\left(OF,y\right)}{\text{Prob}_{OF}\left(OF\right)\text{Prob}_{y}\left(y\right)} \quad (11)$$

From the above Eq.(11), the resultant mutual traffic information is estimated based on the joint probability network traffic feature function $Prob(OF,y)$ and the marginal probability network traffic feature function $Prob_{OF}$, $Prob_y$ respectively. The pseudo code representation of Discrete Mutual Information-based Recurrent Deep Neural Learning Classifier is given below.

Algorithm 2: Discrete Mutual Information-based Recurrent Deep Neural Learning Classifier

Input: Dataset *DS*, Optimal features $OF=\{OF_1,OF_2,\ldots,OF_n\}$

Output: Accurate attack type variant classification

1: Initialize *TSbytes*, *TSloss*, *TSload*, *TSpkts*, *TService*, *TScrip*, *TSport*, *TSmeansz*, *TSintpkt*, *TSwin*, *TSTCPb*, *TState*

2: Begin

3: For each Dataset *DS* with Optimal feature *OF*

4: Estimate inputs to the hidden states as in Eq.(7)

5: Estimate sequence of hidden states as in Eq.(8)

6: Estimate inputs to the output units as in Eq.(9)

7: Estimate sequence of network traffic predictions as in Eq.(10)

8: Update weight based on mutual traffic information as in Eq.(11)

9: Return(predicted results *y*)

10: If *SBytes=TSBytes*, *Sloss=TSloss*, *SLoad=TSLoad and Spkts=TSpkts*

11: Then *y* is Normal

12: Else *y* is Fuzzers

13: End if

14: If Service=TService

15: Then *y* is Normal

16: Else *y* is Analysis

17: End if

18: If *Srcip=TSrcip*, *Sport=TSport*

19: Then *y* is Normal

20: Else *y* is Backdoor

21: End if

22: If *SMeansz=TSMeansz*, *Sintpkt=TSintpkt*

23: Then *y* is Normal

24: Else *y* is DoS

25: End if

26: If *Srcip=TSrcip*, *Service=TService*

27: Then *y* is Normal

28: Else *y* is Exploit

29: End if

30: If *Srcip=TSrcip*, *SWin=TSWin* and *STCPb=TSTCPb*

31: Then *y* is Normal

32: Else *y* is Generic

33: End if

34: If *Service=TService*, *State=TState*

35: Then *y* is Normal

36: Else *y* is Reconnaisance

37: End if

38: If *Srcip=TSrcip*, *Sport=TSport* and *Service=TService*

39: Then *y* is Normal

40: Else *y* is Shell code

41: End if

42: If *Service=TService*

43: Then *y* is Normal

44: Else *y* is Worm

45: End if

46: End for

47: End

From Algorithm 2, classification of different types of attacks based on the description given in Table.1 is made by means of Recurrent Deep Neural Learning Classifier. First, the optimal feature selected is provided as input to the input unit. Next, forms the hidden layer where the learning is performed for prediction and accordingly weight estimation for each network traffic feature is obtained. Third, with the aid of Discrete Mutual Traffic Information, value of weight is updated. Finally, the classified results provide the type of attacks or the normal traffic with better accuracy.

4. EXPERIMENTAL SETUP

Multi-objective Whale Optimized with Recurrent Deep Learning (MWO-RDL) classifier for intrusion detection is performed using the Python source code. To measure the MWO-RDL method, the network traffic generated via UNSW-NB 15 network traffic dataset is used. First, dataset details are provided. Followed by which, qualitative analysis is presented and finally, experiments are conducted on various factors with respect to distinct numbers of network traffics are proposed.

4.1 DATASET DETAILS

In this work, UNSW-NB15 intrusion detection dataset [20] has been used for estimating the performance. The dataset was determined by synthetic surroundings by University of New South Wales (UNSW) cyber security laboratory in 2015. Dataset extracted the unique data for training IDS. In training set, dataset comprise 45 features and 9 attacks for 82,332 records. Moreover, UNSW-NB15 intrusion detection dataset was produced as Cyber Range Lab of the ACCS. Nine types of attacks controlled in UNSW-NB15 dataset are, Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms with 45 distinct features. These distinct features are nominal and non-nominal values namely float, binary, timestamp as well as integer. These features were grouped into Flow, Basic, Content, Time and, Labeled features. From Table.1 dataset description are illustrated below,

Table.1. Dataset Description - Flow features

Name	Description
Srcip	Source IP address
Sport	Source port number
Dstip	Destination IP address
Dsport	Destination port number

Proto	Protocol type
State	Indicates state
Dur	Duration
Sbytes	Source bytes
Dbytes	Destination bytes
Sttl	Source to destination time to live
Dttl	Destination to source time to live
Sloss	Source packet retransmitted
Dloss	Destination packet retransmitted
Service	Service type
Sload	Source bits per second
Dload	Destination bits per second
Spkts	Source to destination packet count
Dpkts	Destination to source packet count
Swin	Source TCP window advertisement value
Dwin	Destination TCP window advertisement value
Stcpb	Source TCP base number
Dtcpb	Destination TCP base number
Smeansz	Mean of flow packet size transmitted by source
Dmeansz	Mean of flow packet size transmitted by destination
Trans_depth	Represents pipelined depth
Res_bdy_len	Actual uncompressed content size
Sjit	Source jitter
Djit	Destination jitter
Stime	Record start time
Ltime	Record last time
Sintpkt	Source inter packet arrival time
Dintpkt	Destination inter packet arrival time
Tcprtt	TCP round trip time
Synack	Time between syn and acknowledge
Ackdat	Time between acknowledge and data
Attack_cat	Category of attack
Label	0 for normal and 1 for attack

Attack types from Table.1 (i.e., dataset) can be classified into nine groups. The nine groups attack types, their description and the features used to describe the attack are listed in Table.2.

Table.2. Attack Types

Attack type	Description	Features used	Threshold used
Fuzzers	An attack where attacker acquire security faults in operating system by providing enormous data to crash.	Sbytes Sloss Sload Spkts	TSbytes TSloss TSload TSpkts
Analysis	Intrusions that infiltrate web via ports, emails and web scripts	Service	TService
Backdoor	Bypassing normal authentication,	Srcip Sport	TSrcip TSport
DoS	Interrupts computer resources via memory as busy to mitigate authorized requests.	Smeansz Sintpkt	TSmeansz TSintpkt
Exploit	Sequence of instructions that take advantage of a bug by unsuspected host behavior.	Srcip Service	TSrcip TService
Generic	Establishes against block-cipher to collide without configuration of block cipher	Srcip Swin STCPb	TSrcip TSwin TSTCPb
Reconnaissance	Obtain network information for security evasion.	Service State	TService TState
Shell code	Attacker penetrates piece of code to control compromised machine	Srcip Sport Service	TScrip TSport TService
Worm	Replica to spread to other computers	Service	TService

4.2 QUALITATIVE ANALYSIS

From highly sensitive network traffic, intrusion detection of proposed method, Multi-objective Whale Optimized with Recurrent Deep Learning (MWO-RDL) is provided. First, indispensable method is selected for applying Robust Scaler and Multi-objective Whale Optimized Feature Selection model to obtain the optimal network traffic reduced feature subset in the feature selection phase. From Table.3 features for UNSW-NB 15 network traffic dataset is illustrated.

Table.3 Selected features for the UNSW-NB 15 network traffic dataset

Feature	Description
Sbytes	Source to destination bytes
Sloss	Source packets retransmitted
Sload	Source bits per second

Spkts	Source to destination packet count
Service	http, ftp, ssh, dns
Srcip	Source IP address
Sport	Source port number
Smeansz	Mean of flow packet size transmitted by source
Sintpkt	Source inter packet arrival time
Swin	Source TCP window advertisement
STCPb	Source TCP sequence number

State	State and protocol

By implementing Robust Scaler and Multi-objective Whale Optimized Feature Selection model along, the method is found to minimize the dimensionality in a drastic manner and discard the irrelevant features from the dataset. Next, with the objective of efficiently enhancing the classification performance of IDS, a Discrete Mutual Information-based Recurrent Deep Learning classifier is employed which classifies from the optimal feature selected into different types of attack and normal traffic. The results obtained are provided in Table.4.

Table.4 Discrete Mutual Information-based Recurrent Deep Learning Classifier results

Srcip	**port**	**service**	**state**	**spkts**	**sbytes**	**sload**	**sloss**	**sinpkt**	**swin**	**stcpb**	**smean**	**attack_cat**	**label**
1	udp	-	INT	2	496	180363632	0	0.011	0	0	248	Normal	0
2	udp	-	INT	2	1762	881000000	0	0.008	0	0	881	Normal	0
3	udp	-	INT	2	1068	854400000	0	0.005	0	0	534	Normal	0
4	udp	-	INT	2	900	600000000	0	0.006	0	0	450	Normal	0
5	udp	-	INT	2	2126	850400000	0	0.01	0	0	1063	Normal	0
250	gre	-	INT	2	156	69333328	0	0.009	0	0	78	Exploits	1
251	gre	-	INT	2	156	69333328	0	0.009	0	0	78	Exploits	1
252	gre	-	INT	2	156	69333328	0	0.009	0	0	78	Exploits	1
253	udp	-	INT	2	196	49000000	0	0.016	0	0	98	Exploits	1
254	tcp	http	FIN	10	830	26616.30664	2	24.947111	255	3306255269	83	Exploits	1
244	ospf	-	INT	20	1280	10551.125	0	48.525633	0	0	64	Reconnaissance	1
245	ospf	-	INT	20	1280	10551.125	0	48.525633	0	0	64	Reconnaissance	1
260	udp	-	INT	2	168	224000000	0	0.003	0	0	84	Reconnaissance	1
261	tcp	-	FIN	10	564	6121.92041	2	73.760444	255	702207080	56	Reconnaissance	1
262	tcp	http	FIN	10	1148	38065.08594	2	24.145778	255	1345084615	115	Reconnaissance	1
278	tcp	-	FIN	10	674	9273.44043	2	56.548667	255	2589843549	67	Fuzzers	1
282	tcp	-	FIN	20	17266	192446.9531	7	33.942421	255	894453596	863	Fuzzers	1
283	tcp	-	FIN	10	490	6223.012207	2	61.171222	255	2164953844	49	Fuzzers	1
284	tcp	-	FIN	10	738	6380.013672	2	88.893111	255	2847876758	74	Fuzzers	1
285	tcp	-	FIN	10	1506	14156.46484	2	77.807333	255	21155452	151	Fuzzers	1
339	tcp	http	FIN	10	800	3930.656494	2	162.822667	255	3499888983	80	DoS	1
337	tcp	http	FIN	10	958	4800.321777	2	159.804111	255	750893150	96	DoS	1
360	ospf	-	INT	20	1280	344.803925	0	1484.90175	0	0	64	DoS	1
361	ospf	-	INT	20	1280	344.803925	0	1484.90175	0	0	64	DoS	1
362	ospf	-	INT	20	1280	344.803925	0	1484.90175	0	0	64	DoS	1
578	udp	-	INT	2	322	128800000	0	0.01	0	0	161	Shellcode	1
583	tcp	-	FIN	10	518	6938.601562	2	57.501778	255	1595099619	52	Shellcode	1
589	tcp	-	FIN	10	628	20437.08594	2	22.025222	255	2758651049	63	Shellcode	1
688	udp	-	INT	2	144	57600000	0	0.01	0	0	72	Shellcode	1
724	tcp	-	FIN	10	724	15012.92285	2	33.523	255	3978417235	72	Shellcode	1
784	br-sat-mon	-	INT	2	200	100000000	0	0.008	0	0	100	Analysis	1

809	vmtp	-	INT	2	200	100000000	0	0.008	0	0	100	Analysis	1
860	sprite-rpc	-	INT	2	200	160000000	0	0.005	0	0	100	Analysis	1
1118	ospf	-	INT	20	1040	16138.58203	0	25.776736	0	0	52	Analysis	1
1171	chaos	-	INT	2	200	266666656	0	0.003	0	0	100	Analysis	1
1587	tcp	http	FIN	10	1282	7321.504395	2	138.352778	255	560953154	128	Worms	1
2854	udp	-	INT	2	92	73600000	0	0.005	0	0	46	Worms	1
3663	tcp	http	FIN	10	1306	13862.04004	2	71.071667	255	94971482	131	Worms	1
4263	udp	-	INT	2	2050	1366666624	0	0.006	0	0	1025	Worms	1
5298	tcp	http	FIN	10	1302	10094.81055	2	96.318667	255	1102451960	130	Worms	1
246	ospf	-	INT	20	1280	10551.125	0	48.525633	0	0	64	Backdoor	1
483	chaos	-	INT	2	200	266666656	0	0.003	0	0	100	Backdoor	1
513	ospf	-	INT	20	1040	16138.58203	0	25.776736	0	0	52	Backdoor	1
524	dcn	-	INT	2	200	88888888	0	0.009	0	0	100	Backdoor	1
694	tcp	-	FIN	10	454	13761.31348	2	25.132	255	2995528503	45	Backdoor	1
11942	udp	dns	INT	2	114	50666664	0	0.009	0	0	57	Generic	1
11943	udp	dns	INT	2	114	152000000	0	0.003	0	0	57	Generic	1
11944	udp	dns	INT	2	114	50666664	0	0.009	0	0	57	Generic	1
11945	udp	dns	INT	2	114	152000000	0	0.003	0	0	57	Generic	1
11946	udp	dns	INT	2	114	152000000	0	0.003	0	0	57	Generic	1

4.3 DISCUSSION

In this section, intrusion detection comparative analysis for high network sensitive traffic using the proposed Multi-objective Whale Optimized with Recurrent Deep Learning (MWO-RDL) method with existing methods, LSTM-PCA [1] with Auto Encoder and Variational Auto Encoder (AE-VAE) [2] are presented. Performance metrics analyzed are intrusion detection rate, intrusion detection time, and false alarm rate and classification accuracy with respect to distinct network traffic collected at different time intervals.

4.3.1 Performance Analysis of Intrusion Detection Rate:

Intrusion detection rate refers to the rate of intrusion being detected based on the network traffic features. Higher being the intrusion being detected at the early stage, more efficient is the intrusion detection rate is. The intrusion detection rate is mathematically formulated as given below.

$$IDR = \sum_{i=1}^{n} \frac{F_{ADI}}{F_i} * 100 \quad (12)$$

From the above Eq.(12), the intrusion detection rate *IDR* is measured based on the network traffic features involved in simulation F_i and the features accurately detected with intrusion F_{ADI}. IDR is measured in percentage (%). From Table.5 lists calculated values of intrusion detection rate from Eq.(12) using the three methods, MWO-RDL, LSTM-PCA [1] and AE-VAE [2] is illustrated.

Table.5. Intrusion Detection Rate

Network traffic features	Intrusion detection rate (%)		
	MWO-RDL	LSTM-PCA	AE-VAE
50	94.0	81.0	67.0
100	95.0	78	66.0
150	94.6	77.5	71.3
200	96.5	75	74.0
250	94.0	70.3	68.4
300	80.6	67	64.6
350	67.4	63.7	60.8
400	61.5	60.7	59.5
450	52.8	50.5	52.0
500	51.4	49.5	49.0

The Table.5 shows the intrusion detection rate with respect to 500 distinct network traffic features. From the From Table.5, it is inferred that increasing the network traffic features causes a slight downward trend towards the intrusion detection rate. This is due to the reason that increasing the network traffic features causes an increase in the types of features to be classified for detecting intrusion and obviously causing a decrease in the intrusion detection rate. However, comparative analysis with 50 network traffic features saw 47 features accurately detected with intrusion using MWO-RDL, 41 and 33 using [1] and [2].

From this result, the intrusion detection rate using MWO-RDL was improved compared with LSTM-PCA and AE-VAE. The reason behind the enhancement was owing to Robust Scaler function during data transformation that in turn performed significant mapping, therefore assisting intrusion detection rate. Intrusion detection rate using MWO-RDL was enhanced as 9% and 13% compared with LSTM-PCA and AE-VAE respectively.

4.3.2 Performance Analysis of Intrusion Detection Time:

The second parameter of significance for intrusion detection in high sensitive network traffic is the intrusion detection time. This parameter denotes time taken for detecting the intrusion. The intrusion detection time is shown by,

$$IDR = \sum_{i=1}^{n} F_i * Time[ID] \quad (13)$$

From the above Eq.(13), the intrusion detection time *IDT* is calculated on distinct network traffic features F_i with time taken for detecting intrusion *Time*[*ID*]. It is measured in terms of milliseconds (ms). The Table.6 given below lists the calculated values of intrusion detection time from Eq.(13) using the three methods, MWO-RDL, LSTM-PCA [1] and AE-VAE [2] respectively.

Table.6. Intrusion Detection Time

Network Traffic Features	Intrusion Detection Time (ms)		
	MWO-RDL	LSTM-PCA	AE-VAE
50	3.0	3.5	3.75
100	3.2	4.5	5.0
150	3.7	6.25	6.75
200	3.92	7.5	8.2
250	4.1	10.0	11.1
300	4.7	11.5	12.5
350	5.1	13.2	13.75
400	5.9	15.0	16.25
450	6.5	17.25	18.75
500	7.2	18.75	20.0

The Table.6 shows intrusion detection time conducted for ten different simulation runs. In above figure, intrusion detection time is proportional to the network traffic features. Specifically, enhancing network traffic features determines improve in type of network traffic features to be monitored for simulation and this in turn increases the intrusion detection time also. However, simulations conducted with 50 network traffic features the intrusion detection time using MWO-RDL was observed to be 0.06ms, 0.07ms using [1] and 0.075ms using [2] respectively. With this, the intrusion detection time using MWO-RDL was found to be comparatively lesser than LSTM-PCA and AE-VAE. Due to utilization of Robust Scaler and Multi-objective Whale Optimized Feature Selection algorithm is enhanced. By applying this algorithm, only after the feature or data transformation the optimal features were selected. With this the intrusion detection time using MWO-RDL was said to be reduced by 50% compared to [1] and 53% compared to [2].

4.3.3 Performance Analysis of False Alarm Rate:

Third parameter of most significance for intrusion detection is the false alarm rate owing to reason that the ordinary traffic should not be estimated as the attack traffic. It represents percentage ratio of ordinary network traffic being falsely identified by attack variant. This is mathematically expressed as given below.

$$FAR = \sum_{i=1}^{n} \frac{NF_{attack}}{F_i} * 100 \quad (14)$$

In Eq.(14), false alarm rate *FAR* is calculated on network traffic features involved in the simulation F_i and the normal traffic features falsely detected as attack variant NF_{attack}. False alarm rate is measured in percentage (%). From Table.7, False alarm rate from Eq.(14) using the three methods, MWO-RDL, LSTM-PCA [1] and AE-VAE [2] is illustrated as:

Table.7. False Alarm Rate

Network Traffic Features	False Alarm Rate (%)		
	MWO-RDL	LSTM-PCA	AE-VAE
50	6.0	19.0	33.0
100	5.0	22.0	34.0
150	3.4	21.5	28.6
200	3.5	25.0	26.0
250	6.0	29.7	31.59
300	19.3	33.0	35.3
350	33.4	36.3	39.3
400	38.5	39.3	41.5
450	47.2	49.5	48
500	48.6	51.5	51

The Table.7 given above shows the results of false alarm rate with respect to distinct 500 network traffic features obtained at different time intervals. The false alarm rate was neither found to be increasingly proportional nor decreasingly proportional to the network traffic features provided as input. However, with simulations conducted for 50 network traffic features 15 features being detected with attack variants and 1 normal traffic feature falsely detected as attack variant using MWO-RDL, 3 normal traffic features falsely detected as attack variant using [1] and 4 normal traffic features falsely detected as attack variant using [2], the false alarm rate was observed to be 6.0%, 19.0% and 33.0% respectively. With these results the false alarm rate using MWO-RDL was comparatively lesser than [1] and [2]. The reason behind the minimization of false alarm rate was resultant to new arbitrary fitness function. By applying this function, the number of features wrongly identified as attack type using MWO-RDL was reduced by 46% compared to [1] and 50% compared to [2].

4.3.4 Performance Analysis of Classification Accuracy:

Finally, classification accuracy or the accuracy rate involved during the process of intrusion detection is measured. This parameter measures the accuracy involved while performing the classification of network traffic type into either attack variants or normal traffic. The classification accuracy is mathematically expressed as given below.

$$CA = \sum_{i=1}^{n} \frac{F_{CA}}{F_i} * 100 \quad (15)$$

In Eq.(15), classification accuracy *CA* was measured on network traffic features by intrusion detection process F_i and the network traffic features classified accurately into either attack type or normal traffic F_{CA}. Classification accuracy is calculated in %. From Table.8 lists calculated values of classification accuracy rate from Eq.(15) using the three methods, MWO-RDL, LSTM-PCA [1] and AE-VAE [2] is shown by,

Table.8 Classification Accuracy

Network Traffic Features	Classification accuracy (%)		
	MWO-RDL	LSTM-PCA	AE-VAE
50	90.0	60	58
100	99.0	67.6	70
150	98.6	76	70.6
200	98.0	75.5	74.5
250	96.0	68.4	65.2
300	94.2	65.3	63.3
350	85.2	62	62.2
400	79.2	59.5	57.4
450	77.4	58.9	56.8
500	70.3	58.2	56.9

The Table.8 illustrates classification accuracy measure using three different methods, MWO-RDL, LSTM-PCA [1] and AE-VAE [2]. From the Table.8, it is inferred that the classification accuracy drops slightly using all the three methods with a slight increase in the network features. However, with simulations conducted with 50 network traffic features, 45 features were accurately classified accurately into either attack type or normal traffic using MWO-RDL, 30 and 29 using [1] and [2]. From this result, the classification accuracy was observed to be 90%, 60% and 58% respectively. The improvement in classification accuracy using MWO-RDL method was owing to the application of Discrete Mutual Information-based Recurrent Deep Neural Learning Classification algorithm. By applying this algorithm, on the basis of the discrete mutual information between the optimal selected features, the classification of different types of attack or normal traffic were made. With this the classification accuracy using MWO-RDL method was found to be improved by 36% compared to [1] and 50% compared to [2] respectively.

5. CONCLUSION

Deep learning techniques have demonstrated their potentiality to detect intrusion in many areas of research in the recent years. Despite the design of several techniques, intruders use novel and innovative mechanisms to instigate distinct attacks type. Though numerous endeavors attempt to detect and classify these attack variants but however are not free from drawbacks like, false alarm rate, accuracy, intrusion detection rate. This article proposes and develops an intrusion detection method for highly sensitive network traffic using MWO-RDL classifier. First, Feature Transformation and Feature Selection based on Machine Learning Techniques called, Robust Scaler and Multi-objective Whale Optimization based Feature selection was proposed to retrieve optimal and relevant feature required for further processing, Next, Discrete Mutual Information-based Recurrent Deep Neural Learning Classifier was employed for classifying different types of attacks, normal traffic respectively. Our proposed method showed significant results with various parameters comparison with existing intrusion detection methods.

REFERENCES

[1] F. Laghrissi, S. Douzi, K. Douzi and B. Hssina, "Intrusion Detection Systems using Long Short-Term Memory (LSTM)", *Journal of Big Data*, Vol. 8, pp. 1-6, 2021.

[2] S. Zavrak and M. Iskefiyeli, "Anomaly-based Intrusion Detection from Network Flow Features using Variational Autoencoder", *IEEE Access*, Vol. 10, No. 8, pp. 108346-108358, 2020.

[3] A. Shenfield and D. Day D, "Intelligent Intrusion Detection Systems using Artificial Neural Networks", *ICT Express*, Vol. 1, No. 4, pp. 95-99, 2018.

[4] M. Mazini, B. Shirazi and I. Mahdavi, "Anomaly Network-Based Intrusion Detection System using a Reliable Hybrid Artificial Bee Colony and AdaBoost Algorithms", *Journal of King Saud University-Computer and Information Sciences*, Vol. 31, No. 4, pp. 541-553, 2019.

[5] A.S. Alzahrani, R.A. Shah, Y. Qian and M. Ali, "A Novel Method for Feature Learning and Network Intrusion Classification", *Alexandria Engineering Journal*, Vol. 59, No. 3, pp. 1159-1169, 2020.

[6] I. Ullah and Q.H. Mahmoud, "Design and Development of a Deep Learning-Based Model for Anomaly Detection in IoT Networks", *IEEE Access*, Vol. 9, pp. 103906-103926, 2021.

[7] O.A. Sarumi, A.O. Adetunmbi and F.A. Adetoye, "Discovering Computer Networks Intrusion using Data Analytics and Machine Intelligence", *Scientific African Journal*, Vol. 9, pp. 1-9, 2020.

[8] Y. Zhou and M. Dai, "Building an Efficient Intrusion Detection System based on Feature Selection and Ensemble Classifier", *Computer Networks*, Vol. 19, pp. 1-15, 2020.

[9] Z. Ji, J. Gong and J. Feng, "A Novel Deep Learning Approach for Anomaly Detection of Time Series Data", *Scientific Programming*, Vol. 2021, pp. 1-12, 2021.

[10] N. Thomas Rincy and Roopam Gupta, "Design and Development of an Efficient Network Intrusion Detection System using Machine Learning Techniques", *Wireless Communications and Mobile Computing*, Vol. 2021, pp. 1-11, 2021.

[11] J. Lansky, "Deep Learning-Based Intrusion Detection Systems: A Systematic Review", *IEEE Access*, Vol. 9, pp. 101574-101599, 2021.

[12] A. Drewek Ossowicka, M. Pietrołaj and J. Ruminski, "A Survey of Neural Networks usage for Intrusion Detection Systems", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 12, No. 1, pp. 497-514, 2021.

[13] N. Moustafa and J. Slay, "UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems (UNSW-NB15 Network Data Set)", *Proceedings of International Conference on Military Communications and Information Systems*, pp. 1-6, 2015.

[14] R.A. Bridges and Q. Chen, "A Survey of Intrusion Detection Systems Leveraging Host Data", *ACM Computing Surveys*, Vol. 52, No. 6, pp. 1-35, 2019.

[15] Jiadong Ren, Jiawei Guo, Wang Qian,Huang Yuan, Xiaobing Hao and Hu Jingjing, "Building an Effective Intrusion Detection System by Using Hybrid Data Optimization Based on Machine Learning Algorithms", *Security and Communication Networks*, Vol. 13, No. 1, pp. 1-13, 2019.

[16] Soulaiman Moualla, Khaldoun Khorzom and Assef Jafar, "Improving the Performance of Machine Learning-Based Network Intrusion Detection Systems on the UNSW-NB15 Dataset", *Computational Intelligence and Neuroscience*, Vol. 2021, pp. 1-13, 2021.

[17] Y. Li, R. Qiu and S. Jing, "Intrusion Detection System using Online Sequence Extreme Learning Machine (OS-ELM) in Advanced Metering Infrastructure of Smart Grid", *PloS One*, Vol. 13, No. 2, pp. 1-13, 2018.

[18] M.J. Kang and J.W. Kang, "Intrusion Detection System using Deep Neural Network for in-Vehicle Network Security", *PloS One*, Vol. 11, No. 6, pp. 1-13, 2016.

[19] F. Laghrissi, S. Douzi and K. Douzi, "Intrusion Detection Systems using Long Short-Term Memory (LSTM)", *Journal on Big Data*, Vol. 8, pp. 65-79, 2021.

[20] UNSW Dataset, Available at https://research.unsw.edu.au/projects/unsw-nb15-dataset, Accessed at 2020.