ISSN: 0974-5823 Vol. 7 No. 5 May, 2022

International Journal of Mechanical Engineering

An Enhanced Neutrosophic Correlation with Long-Short Memory Unit using Whale Optimization for Intrusion Detection Model

[1] Soundarraj.K, [2]Ravichandran.M

Department of Computer Science, Sri Ramakrishna Mission vidyalaya College of Arts and Science, Coimbatore, Tamilnadu.

Abstract

In this contemporary computing infrastructure, it is a vital task to monitor and detect malicious network activity and it is accomplished by Network Intrusion Detection Systems (NIDS). Most of the existing literatures, NIDS are signature-based, with a set of rules used to determine what constitutes undesirable network traffic by monitoring patterns in that traffic. Even though such systems are highly effective against known threats like signature-based detection, they fail to detect unknown or known attacks that are modified to circumvent due to impreciseness. In this paper an enhanced impreciseness handling Long-short term memory is developed to improve the accuracy of network intrusion detection and reduce false alarm rate. To handle the impreciseness, this work developed a neutrosophic correlation based significant subset selection is accomplished. The instances in IDs dataset is represented in neutrosophic triplet values based on the truthiness, falsity and indeterminacy to eliminate the irrelevant and redundant features. The variant of recurrent neural network, known as Long-Short Term memory is used for understanding the long sequence of KDD cup 99 dataset to predict the data packet with traffic information as normal or anomalies. The vanishing problem in LSTM is overcome intelligently by adapting whale optimization algorithm. The prey searching behavior of the humpback whale is used for searching potential values to the parameters namely input weights and recurrent neuron weights and bias value of each cell state for enriching the learning rate of the LSTM to detect Intrusion detection more prominently. From the observed results, it proves the effectiveness of the proposed ELSTM-WOA by achieving better accuracy of intrusion detection with less false alarm compared with other existing models even in presence of class imbalance, impreciseness and vagueness in the dataset.

Keywords: network Intrusion detection, impreciseness, vagueness, Long-Short Term Memory, Neutrosophic Correlation, whale optimization

Introduction

Presently, owing to the widespread utilization of Internet of Things (IoT) principles, there are vast amounts of interconnected physical devices which encompass not only computers yet also automobiles, digital gadgets, sensors, and so forth [1]. Because of the network's massive scale and the Internet's unregulated nature, researchers have found it difficult to preserve both company information and conversations [2]. Even though many systems employ firewalls to avoid this, intrusion detection systems (IDSs), that are considered the next level of defense, which play a significant role in raising the security aspect of the system [2]. An intrusion detection system (IDS) is a software application that identifies network intrusions by employing different machine learning algorithms. The IDS monitor is a network or system for hostile behaviour and protects a computer network against illegal access from users, perhaps insiders [3]. The intrusion detector learning aims to design a predictive model which is also known as classifier that can discriminate among normal and abnormal/ intrusion packets.

In general, IDS is categorized into two types depending on the techniques used for detection. They are anomaly detection and misuse detection [4]. The anomaly detection generates normal activities of the network database, if there is any deviation from the normal activity then it alerts the user about the presence of intrusion in the network. The activities related to the assaults are maintained in the database by misuse detection, when it finds same type of attacking pattern incoming, then it is declared as attack. While comparing to misuse detection system, it is stated that anomaly detection is best suited for detecting unknown pattern of attacks.

As the volume of the network traffic pattern increase, the usage of data mining, machine learning and rule-based models play a significant role in intrusion detection. In the initial stage, data mining is the suggested model for network intrusion detection. The data mining approaches acquire knowledge from historical database. It assists to extract interesting pattern from a knowledge base and with the gain knowledge, it predicts future intrusion or abnormal packets entering insider the network. But still they have drawbacks when they deal with new attacks which use a new signature pattern which is not stored in historical or knowledge database. The data collected from network usually be imprecise, ambiguous and vague to understand. The volume of network traffic will also be increasing which results in overflow or overfitting problem by the conventional mining approaches.

Hence, in this proposed work the variant of Recurrent Neural Network which is known as Long-Short Term memory is used to understand the pattern of traffic or packet details. The LSTM with its ability of remembering long sequence detail of the instances, the training phase is potentially improvised by using whale optimization algorithm which selects the best values for parameters

involved in LSTM. The following sections discuss in detail about the working principle of proposed ELSTM-WOA for efficient intrusion detection.

Related Work

Mohammed et al [5] in their study aims to employ an improved intrusion detection system. Deep Neural network produced high network performance to identify unknown attack packages. Both binary and multiclass classification are accomplished for intrusion attack detection.

Yin et al [6] deigned a recurrent neural network for network intrusion detection. They used the cell state information for discovering the pattern of network traffic signals. The binary and multiclass classification on NSL-KDD for intrusion detection system is accomplished for instruction detection system.

Vigneswaran et al [7] developed a deep neural network for classifying the network traffic as normal or attacking. The ReLU based non-linear activation is applied in hidden layer. The output layer comprised of two neurons for predicting the normal and abnormal pattern of incoming data packets.

Yadav et al [8] in their work handled the intrusion detection by discovering any malicious activities. The machine learning algorithm is used to examine entire network activities to detect the anomalies behavior. In this study, the authors discuss about various datasets used for intrusion detection and the different machine learning algorithms used for predicting IDS.

Peng et al [9] designed a novel light weight random neural network that is developed to predict various cyber security attacks such as denial of service, fraudulent operation, etc. These attacks are prominently handled by the deep learning model

Kathryn et al [10] devised a mini batch k-means unsupervised learning model for IDS. The feature reducing is carried by applying principal component analysis and the k means++ with mini batch strategy clusters the instances as normal and abnormal.

Latif et al [11] examines various machine learning model performance in classifying the network attacks. Their advantages and disadvantages are discussed in this paper. Four different algorithms are used to detect binary and multiclass attacks by analyzing the network traffic.

Shenfield [12] designed artificial neural network for malicious network traffic analysis in order to gain deep knowledge about the data packets as normal or abnormal. It improves the process of network intrusion system by applying both cyber physical system and standard network traffic analysis.

Methodology: Neutrosophic Correlation and Long Short-Term Memory Unit using Whale Optimization for Intrusion Detection Model

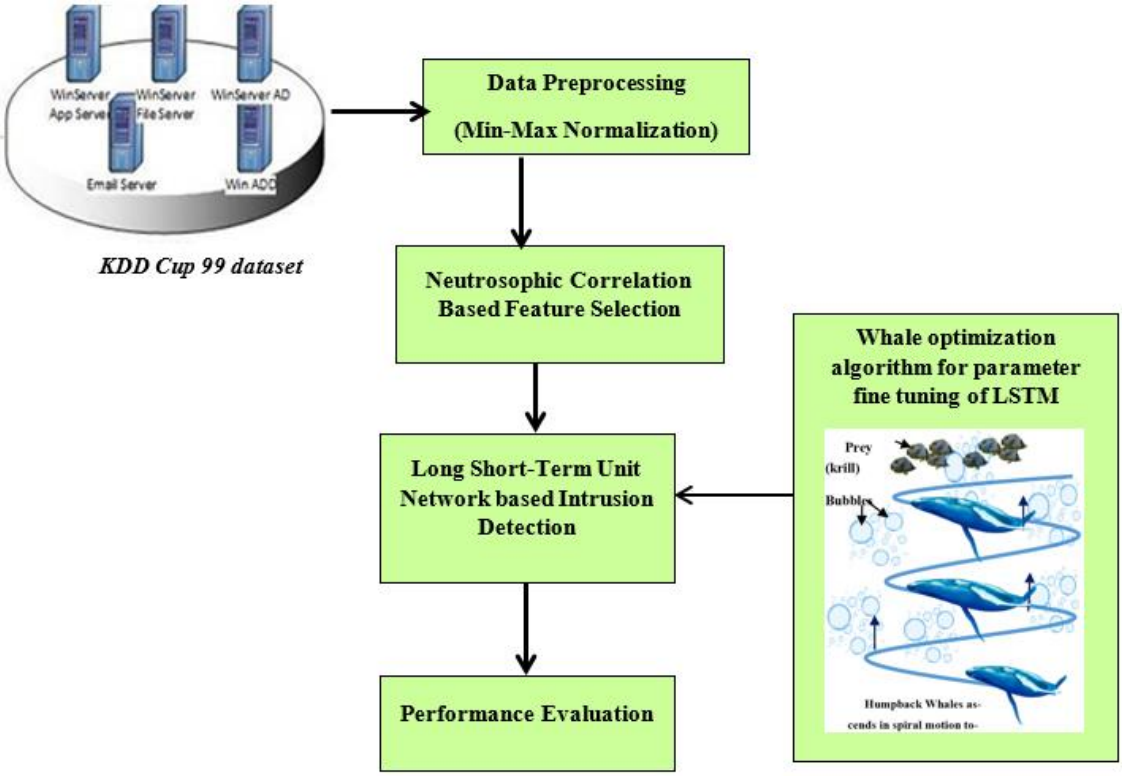

Fig 1: Overall working principle of Enhanced Neutrosophic Correlation and Long Short-Term Memory Unit using Whale Optimization for Intrusion Detection Model

The figure 1 shows the overall framework of the Neutrosophic Correlation and Long Short-Term Memory Unit using Whale Optimization for Intrusion Detection Model. The dataset is preprocessed to treat all the attributes with equal range of values by applying min-max normalization. The KDD Cup 99 [15] dataset which comprised of 41 attributes with 1,00,000 records for Intrusion detection. The irrelevancy and redundancy features are

eliminated using neutrosophic correlation-based feature subset selection. With the reduced subset of features, Long Short-Term Memory classifies the input packet details as normal or abnormal by improving its parameter values by adapting the whale optimization algorithm.

Neutrosophic Correlation based Feature Selection

Once the preprocessing of the instance of KDD Cup 99 dataset [15]t is completed, the features which produce more information to classify the normal packets and abnormal packets are accomplished by developing a novel impreciseness handling algorithm known as neutrosophic correlation feature selection.

In terms of multivalued Neutrosophic theory it characterizes real-time truthiness, falsity, and indeterministic knowledge. The neutrosophic theory directly handles the concerns of imperfections, ambivalence, vagueness, and discrepancy that are regarded as significant issues for determining a most significant feature subset to boost accurate intrusion detection. The dependability of fuzzy and classical theory is preserved in Neutrosophication with its changing grades of truthiness (T), indeterminacy (I) and falsity (F) with the condition T+F+I=1 [16]. Whilst T+F+I < 1 then neutrosophic becomes intuitionistic, and it can sustain paraconsistency when its three elements T+F+I > 1. Thus, NL has the quality of non-standard examination, which discriminates qualified falsehood represented as 0 and absolute untruth signified by $^{-}0$. The actual truth is embodied by 1 and the ultimate fact by 1^{+}. Each Instances in IDS dataset is represented in the triple format based on their degree of truthiness membership, Indeterminacy membership and Falsity membership which is signified as

$$E = \{r, \langle \check{T}_E(r), \check{I}_E(r), \digamma_E(r)\rangle : r \in D, \check{T}_E(d), \check{I}_E(d), \digamma_E(d) \subseteq [0,1]\}$$

The specified input values are transformed to the neutrosophical quality in elements of truthiness, falsity, and indeterminacy member $\langle \check{T}_{NU}(x), \check{I}_{NU}(x), \digamma_{NU}(x)\rangle$

$$\check{T}_{NU}(\mathbf{r}) = \begin{cases} u_{NU}(r); & x_1 \le r < x_2 \\ \check{T}_{NU}(r) \ ; & x_2 \le r < x_3 \\ v_{NU}(r); & x_3 < r \le x_4 \\ 1 \ ; & else \end{cases} \check{I}$$

$$\check{I}_{NU}(\mathbf{r}) = \begin{cases} d_{NU}(r); & y_1 \le r < y_2 \\ \check{I}_{NU}(r) \ ; & y_2 \le r < y_3 \\ e_{NU}(r); & y_3 < r \le y_4 \\ 1 \ ; & else \end{cases}$$

$$\digamma_{NU}(\mathbf{r}) = \begin{cases} k_{NU}(x); & z_1 \le r < z_2 \\ F_{NU}(x) \ ; & z_2 \le r < z_3 \\ l_{NU}(x); & z_3 < r \le z_4 \\ 1 \ ; & else \end{cases}$$

The neutrosophic correlation among the attributes are discovered by applying the formula as shown below

$$\mathcal{P}(E,S) = \frac{NCR(E,S)}{\sqrt{NCR(E,E)} * NCR(S,S)}$$

Where

$$NCR(E,S) = \frac{1}{q}\sum_{j=1}^{q}\sum_{i=1}^{n}\{\check{T}_E^j(\mathbf{r_i})\check{T}_S^j(\mathbf{r_i}) + \check{I}_E^j(\mathbf{r_i})\check{I}_S^j(\mathbf{r_i}) + \digamma_E^j(\mathbf{r_i})\digamma_S^j(\mathbf{r_i})\}$$

$$NCR(E,E) = \frac{1}{q}\sum_{j=1}^{q}\sum_{i=1}^{n}\{\check{T}_E^j(\mathbf{r_i})\check{T}_E^j(\mathbf{r_i}) + \check{I}_E^j(\mathbf{r_i})\check{I}_E^j(\mathbf{r_i}) + \digamma_E^j(\mathbf{r_i})\digamma_E^j(\mathbf{r_i})$$

$$NCR(S,S) = \frac{1}{q}\sum_{j=1}^{q}\sum_{i=1}^{n}\{\check{T}_S^j(\mathbf{r_i})\check{T}_S^j(\mathbf{r_i}) + \check{I}_S^j(\mathbf{r_i})\check{I}_S^j(\mathbf{r_i}) + \digamma_S^j(\mathbf{r_i})\digamma_S^j(\mathbf{r_i})$$

Here r refers to the record in the dataset, q is the number of features in the KDD Cup 99 dataset and E and S refers to two different features. The neutrosophic correlation matrix is generated to discover the independent feature's which contribute high in classifying the data packets as normal or attacking. Depending on the membership values the most relevant features are selected for classification using LSTM-WOA instead of using whole dataset to improve the accuracy of detection rate and reduce the complexity.

Recurrent Neural Networks

In recent years, there has been a surge in the use of deep architectures, one of which is Recurrent Neural Networks (RNN) [13]. RNN is employed in a variety of applications such as word embedding, language modelling, handwriting and speech recognition. RNN like traditional neural network models, it also includes connections between hidden nodes created as a directed graph with sequence of events as shown in the figure 2. The approach can exhibit temporal dynamic characteristics and employs its internal state which is coined as memory, to process input with variable sequence of length.

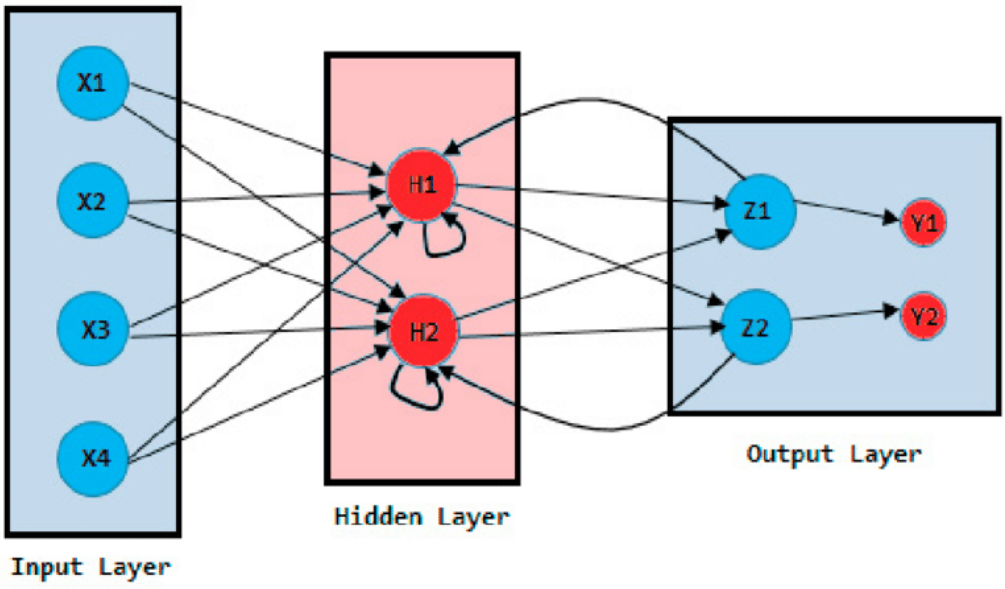

Fig 2 : General Structure of Recurrent Neural Network

RNN is characterized by the flexibility among its forward paradigm of ANN, which incorporates an internal memory. It is recurrent by definition since it performs the same function with each input and the output relates to the previous input calculation. It duplicates the output and sends it back to the recurrent neural network after it is formed. In order to reach a decision, it considers the current input and its outcome, which it learns from the specifics of prior inputs. In contrast to traditional networks, RNNs process input sequences using their memory, known as internal state. Individual inputs passed to the conventional networks are autonomous, whereas inputs participating in the prediction process in recurrent networks are intimately connected to one another.

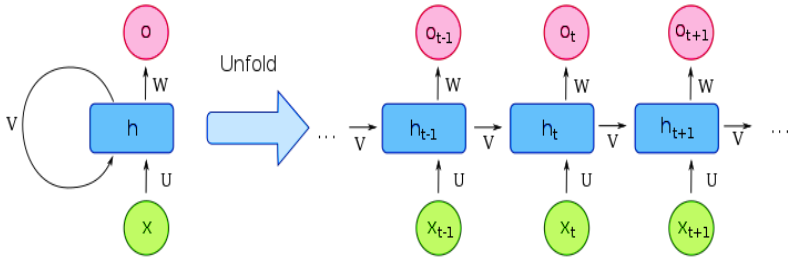

Fig 3: Internal Function of Recurrent Neural Network

It is observed from the figure 3, the input is X_{t-1} and its output value is O_{t-1}m the hidden parameter is $h_{t-1.}$ In the next sequence, the input X_t and hidden value h_{t-1} is used for computing the O_t is the output value, hidden value h_t is the hidden value influenced by both the inputs X_t and $h_{t-1.}$ This process continues for remaining sequence also. RNN follows likewise to keep memorizing the framework during training phase. The present state of RNN is represented as formulated in the equation

$$h_t = f(h_{t-1,}X_t)$$

where h_t is the current state, X_t is the input state and h_{t-1} is the previous state of the RNN. T

The process after applying tanh activation function that is described as follows

$$h_t = f\text{-}tanh(Vh_{t-1} + UX_t)$$

where V is the recurrent node's weight, U denotes an input neuron, h_{t-1} denotes the weight of the previous layer's hidden states and *F*-tanh signifies the activation function, whose range is [-1,1]. The output O_t of the RNN is represented as shown in the equation.

$$O_t = W_t h_t$$

Where W_t denotes the weight of the current input state.

Long Short-Term Memory-Recurrent Neural Network

Although RNN provides good accuracy, it fails to work with long dependencies due to the weak vanishing nature of past distance gradient values. In this research work the variant of RNN which is known as Long Short-Term memory (LSTM) is used for Intrusion detection because it has the ability to learn patterns with lengthy dependencies. Thus, LSTM outperforms state-of-the- art RNN [14]. LSTM is a complicated deep learning model with recurrent neural architecture that, unlike ordinary RNN, uses back prorogation to offer feedback connections. It can process an entire stream of data at once.

A basic LSTM consists of a cell with three separate gates: input, output, and forget. The cell reminisces values at random intervals, and these three gates equalize the flow of information into and out of the cell. LSTM is best suited for processing, prediction, and categorization of temporal data in this work intrusion detection when it is performed. The primary goal of LSTM is to deal with the vanishing gradient problem that occurs during the learning phase of a normal RNN. The difficulty with present RNNs is that when back propagation is used, their gradient disappears because their computations become entangled in the process because they use finite precision values. The adaptation of LSTM units in RNN will solve the gradient vanishing problem because it has the capacity to allow gradients to run unaltered. The Cell of LSTM with three different gates is illustrated in the figure 4.

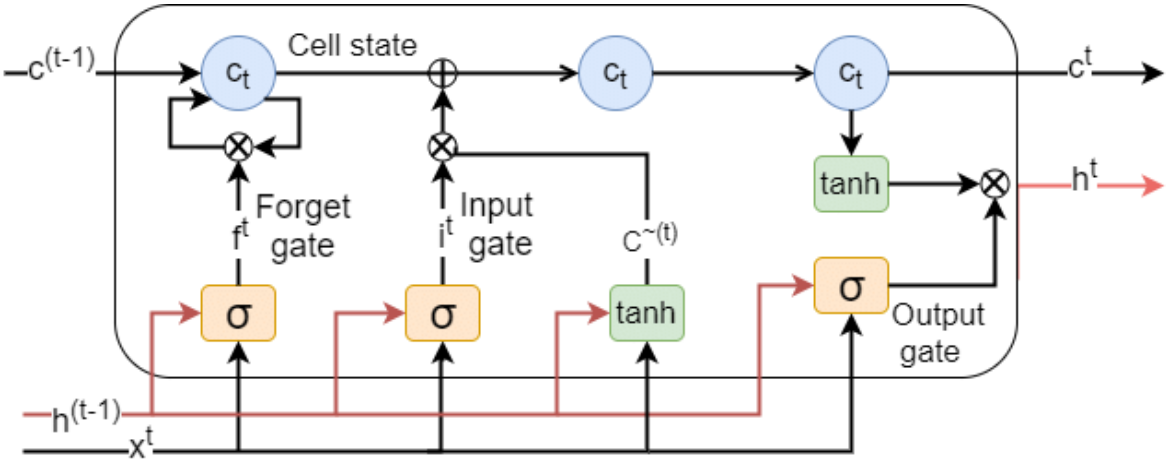

Fig 4: Cell of LSTM

Input Gate: This gate is responsible for identifying the type of input sequence that must be removed in order to change memory. The transfer function converts the range of input value to 0 to 1, the tanh function generates a weight value that agrees to determine their rank ranging from -1 to 1. It is expressed by the equations below.

$$ig\ (\mathcal{I}_t) = \sigma_g(\mathcal{W}_i \mathcal{X}_t + \mathcal{R}_i \mathcal{h}_{t-1} + \mathcal{b}_i)$$

Where $\mathcal{I}_t$ refers to the input value at t period, $\mathcal{W}_i$ refers to weight of the input value $\mathcal{X}_t$, $\mathcal{R}_i \mathcal{h}_{t-1}$ represents recurrent neuron $\mathcal{h}_{t-1}$ weight and $\mathcal{b}$ is the bias.

Forget Gate: This gate determines the information to be eliminated from the input block. The sigmoid function makes the choice by considering the past state h_{t-1}, and the current input value is denoted by $\mathcal{X}_t$ for each cell state C_{t+1}. It is mathematically defined as stated in the following equation.

$$fg\ (f_t) = \sigma_g\big(\mathcal{W}_f \mathcal{X}_t + \mathcal{R}_f \mathcal{h}_{t-1} + \mathcal{b}_f\big),$$

Output gate: Both input and block memory are used to generate the model's output. The Sigmoid activation function includes value assignment. The tanh activation produces weightage, which is compounded by the sigmoid output.

$$CC\ (C_t) = \sigma_c(\mathcal{W}_c \mathcal{X}_t + \mathcal{R}_c \mathcal{h}_{t-1} + \mathcal{b}_c),$$
$$og(o_t) = \sigma_g(\mathcal{W}_o \mathcal{X}_t + \mathcal{R}_o \mathcal{h}_{t-1} + \mathcal{b}_o),$$

Where σ_g signifies activation function of each gate and $\mathcal{W}_i, \mathcal{W}_f, \mathcal{W}_c$ and $\mathcal{W}_0$ matrix are the weight values of input gate, forget gate, cell state and the output gate. The $\mathcal{b}_i, \mathcal{b}_f, \mathcal{b}_c$ and $\mathcal{b}_0$ denote bias vectors. The forget gate governs how considerable amount of previous memory values should be deleted from the cell state. Similarly, the input gate provides a fresh prior intake to the cell state. The C_t refers tot eh cell state and H_t is the output at t time of the LSTM is defined as

$$C_t = ft \odot ct\ \odot 1 + it\ \odot gt,$$
$$H_t = ot \odot \sigma c(ct)$$

Where $\odot$ denotes the Hadamard product (elements-wise multiplication of vectors).

About Whale Optimization Algorithm

The Whale Optimization Algorithm (WOA) is a metaheuristic model based on the inspiration of humpback whale hunting behaviour [5]. These killer whales are grazing on krill or fish flocks at the sea's surface. They proceed to the 12 meters floor of a prey and generate 9 shaped trail bubbles when they've recognized its target. The whales swim uphill in the direction of the sea surface, tracking the bubbles to encircle the prey. The WOA entails three distinct processes: prey encircling, prey probing, and prey attack. This approach is ideally suited for determining the optimal values for the weight parameters of input neurons, recurrent neurons, and bias values for intrusion detection. The figure 5 shows the encircling nature of the whale behaviour.

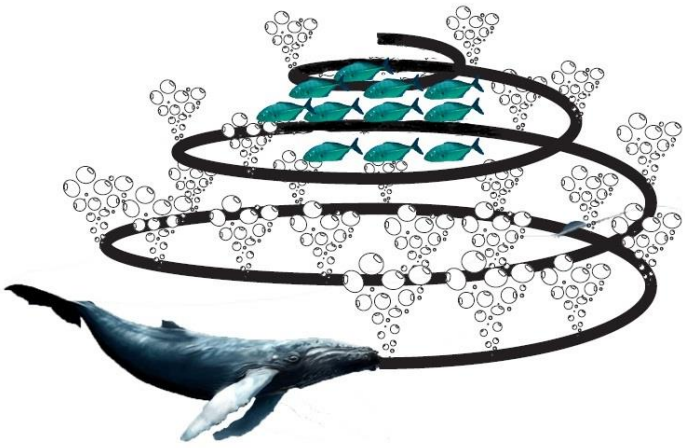

Fig 5: Whale prey encircling with bubble net

Strategy for searching prey

Humpback whales actively explore out prey in an arbitrary manner, which is accomplished by varying the direction B. When |B| < 1, utilization is carried out by adjusting the whale's position to the best feasible individuality, if |B| >= 1, position manipulation is carried out by randomly selecting independent location to attain the best solution it is mathematically represented as

$$Z_{t+1}=Z^{*}(t)- F.G$$

$$F=|S.\ Z^{*}(t) -Z(t)|$$

Where Z(t) refers to position vector and $Z^{*}(t)$ optimum solution is the among the present population, the random selected individual position vector of whale.

Algorithm for Enhanced Neutrosophic Correlation and Long-Short Memory Unit using Whale Optimization

Input: KDD Cup 99 dataset: DS

Procedure

Begin

For i = 1 to n // number of instances in Ds

For j = 1 to m // number of attributes in each instance

// Apply Min-Max Normalization

$$\text{Norm}(R_{i,j}) = \frac{R_{i,j}-\text{Min}(R_{i,j=1..m})}{\text{Max}(R_{i,j=1..m})-\text{Min}(R_{i,j=1..m})}$$

End {For}

End {For}

// Neutrosophic Correlation based Potential Feature subset selection

For E = 1 to m

For S = 1 to m

// Convert each instance to neutrosophic format

$$E = \{r, \langle \check{T}_E(r), \check{I}_E(r), \check{F}_E(r)\rangle : r \in D, \check{T}_E(d), \check{I}_E(d), \check{F}_E(d) \subseteq [0,1]\} \text{ using equations}$$

// Apply Neutrosophic Correlation based ranking

$$\mathcal{P}(E,S) = \frac{NCR(E,S)}{\sqrt{NCR(E,E)} * NCR(S,S)}$$

// Select the potential feature subset

End

Subset = R''_{rank_list}

// LSTM-WOA based Intrusion Detection System

Call LSTM-WOA ($\mathcal{X}_{i=1..n}$)

For i = 1 to Subset

//Assignment of weight using Whale Optimization Algorithm

$Z_{t+1}=Z^*(t)$- F.G

$F=|S. Z^*(t) -Z(t)|$

// Computer Input Gate

$ig\ (\mathcal{I}_t) = \sigma_g(\mathcal{W}_i\mathcal{X}_t + \mathcal{R}_i\mathcal{h}_{t-1} + \mathcal{b}_i)$

// Compute Forget Gate

$fg\ (f_t) = \sigma_g(\mathcal{W}_f\mathcal{X}_t + \mathcal{R}_f\mathcal{h}_{t-1} + \mathcal{b}_f)$

// Computer Cell State

$CC\ (C_t) = \sigma_c(\mathcal{W}_c\mathcal{X}_t + \mathcal{R}_c\mathcal{h}_{t-1} + \mathcal{b}_c)$

// Compute Output Gate

$og(o_t) = \sigma_g(\mathcal{W}_o\mathcal{X}_t + \mathcal{R}_o\mathcal{h}_{t-1} + \mathcal{b}_o)$

// Compute the complete LSTM Cell

$C_t = ft \odot ct \bigodot 1 + it \bigodot gt,$

// Compute recurrent node

$H_t = ot \odot \sigma c(ct)$

End

$Y_{(i=1..n)}$ = LSTM-WOA ($\mathcal{X}_{i=1..n}$)

End

Output: Normal or attacking

The algorithm explains about the complete working principle of the proposed Enhanced Neutrosophic Correlation with Long-Short Memory Unit using Whale Optimization for Intrusion Detection Model

Experimental Results and Discussions

This section describes about performance analysis of proposed ELSTM-WOA algorithm deployed using python code for intrusion detection system. The dataset used in this work is collected from KDD cup 99 dataset with 1,00,000 records and 41 attributes along with class labels. The features details are 10 features with basic characters, 12 belongs to content features and 19 of them related to traffic features. By applying neutrosophic correlation-based feature selection, the subset with best features obtained are protocol type, src_bytes, dst_bytes, srv_count, diff_srv_rate and dst_host_srv_count. In this work, instead of using 41 attributes, after feature reduction only seven potential features are used of predicting Intrusion. The efficiency of ELSTM-WOA for intrusion detection is compared with three existing models DNN, MLP and SVM. The metrics used of measuring the performance of the classification models are accuracy, precision, recall and Error rate.

Table 1: Performance Comparison of four different IDS Models

	Precision	***Accuracy***	***Recall***	***Error Rate***
SVM	76.2	74.7	75.8	0.258
MLP	82.9	80.4	81.2	0.201
DNN	89.3	87.5	88.3	0.196
ELSTM-WOA	98.6	97.8	98.2	0.027

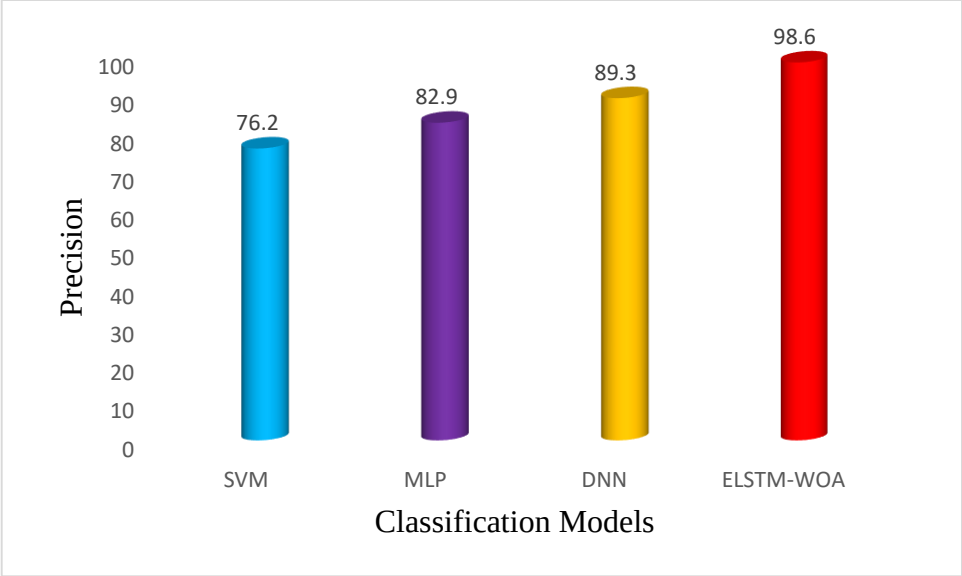

Fig 6: Results based on Precision

The Intrusion Detection system of four different prediction modes based on their precision value is shown in the figure 6. The correctly predicted ratio is high while using the proposed ELSTM-WOA compared to the other three existing models. The reason is ELSTM improves its learning rate of discovering the pattern of abnormal packets which are less in ratio compared to the normal packets. The neutrosophic correlation-based features selection is done to determine the potential features and those are fed as input to the LSTM. The earlier convergence to the results is avoided by applying the whale optimization algorithm. Hence, ELSTM-WOA produced highest precision rate in IDS.

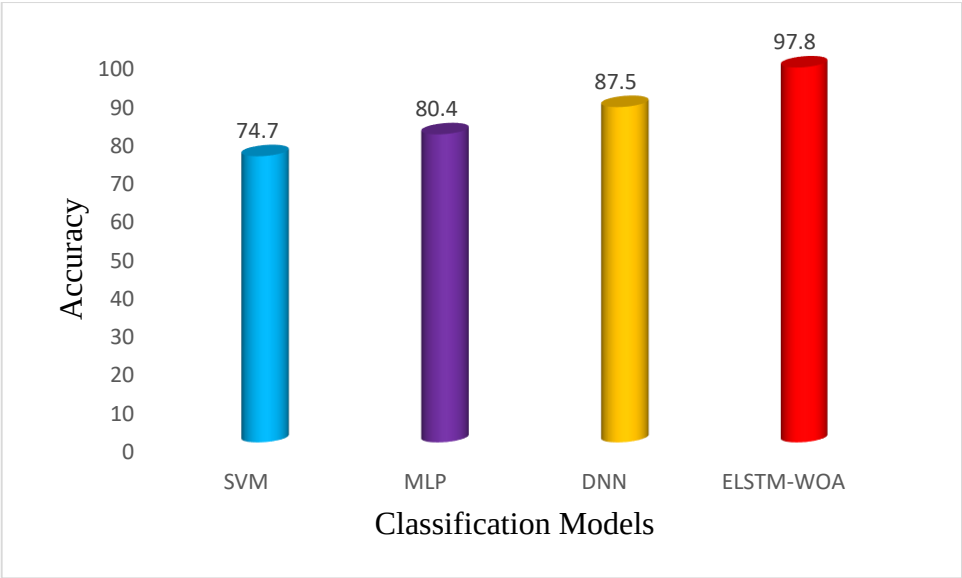

Fig 7: Results based on Accuracy rate of IDS models

The results shown in the figure 7 reveals that the proposed ELSTM-WOA more accurately predicts the abnormal packets more precisely than the SVM, MLP and DNN. The impreciseness and ambiguity among the features in the dataset are perfectly handled by neutrosophic correlation-based feature selection. The process of fine tuning three different parameters involved in LSTM is accomplished by whale optimization algorithm. The whale's hunting behavior is used to searching the best values to be assigned to the parameters of recurrent weights and input state weights to influence the accuracy rate in Intrusion detection system. The existing models SVM, MLP and DNN faces issue of class imbalance and the impreciseness of the attacking patterns in IDS lacks their performance to reach better accuracy value.

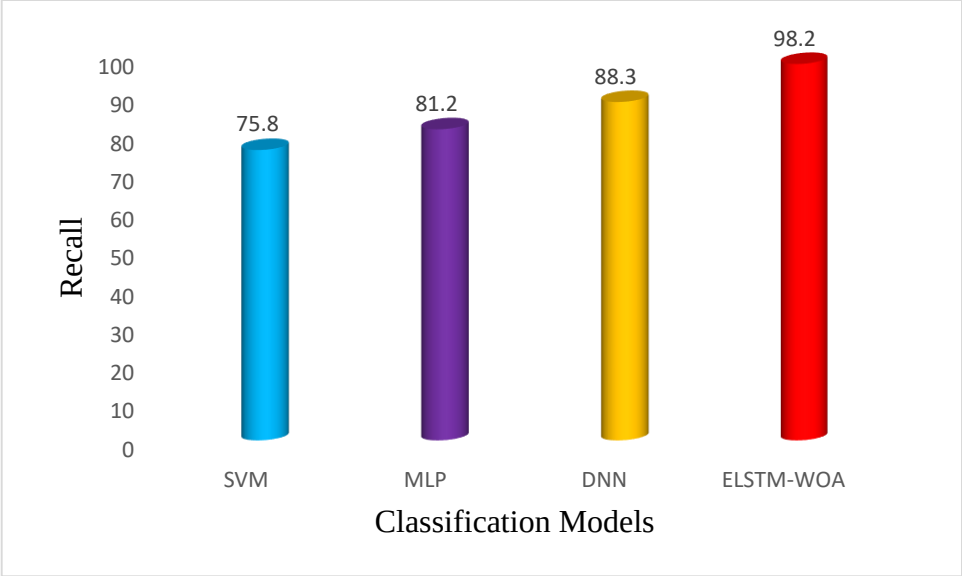

Fig 8: Results based on Recall

The recall value of four different classification models involved in Intrusion detection system is explored in the figure 8. The support vector machine cannot handle the huge volume of IDS dataset, so it results in overfitting and produced least result compared to other three models. The MLP algorithm performance relies on the parameters weight and bias, when it follows the credit assignment known as gradient descend method. The adjustment of the weights is trial and error hence its learning rate is very less. The deep neural network also follows the gradient descent method is used for weight values. Thus, these three existing models produced less rate of recall compared to the proposed model ELSTM-WOA. This is because, instead of using all the features, the highly informative features are discovered using neutrosophic correlation-based feature subset selection. The LSTM learning rate is improved by applying the searching strategy of WOA for assigning better values to weight parameters. Hence, ELSTM-WOA produced better recall rate for Intrusion Detection System even in the presence of impreciseness and ambiguity factors.

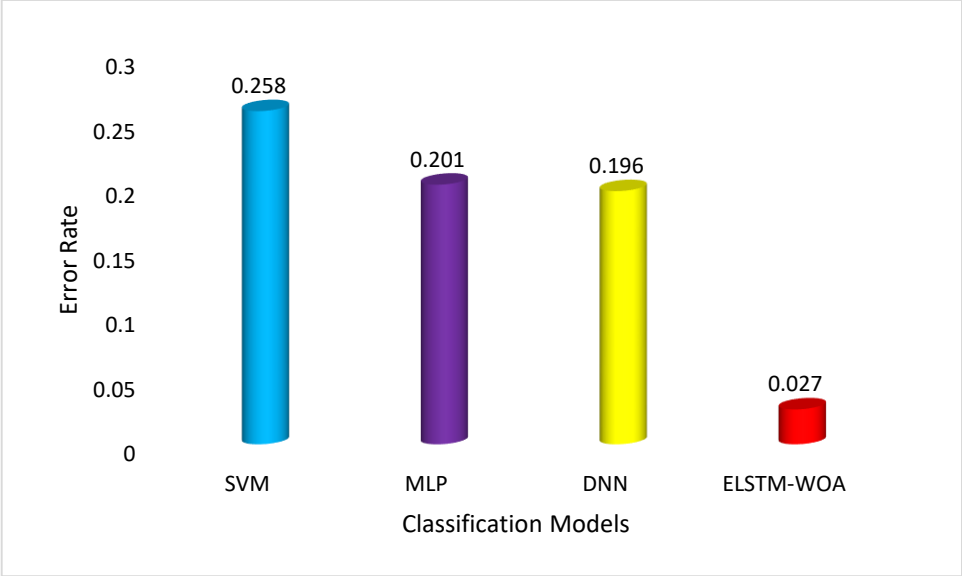

Fig 9: Results based on Error rate

The figure 9 illustrates the performance of four different classification models error rate while predicting the normal and abnormal packets for intrusion detection system. The SVM, MLP and DNN produced more error rate compared to the proposed ELSTM-WOA. The expected and the actual output is maximum correctly predicted by the ELSTM-WOA, because of enhancing the process of Long-short Term memory with two main factors. The neutrosophic correlation-based subset feature selection is carried out for handling the ambiguity and impreciseness among the instances of IDS dataset. The learning rate of the ELSTM is improvised by adapting the whale optimization algorithm to reduce the error rate among the observed and actual values.

Conclusion

This work introduced an imprecise handling model to overcome the existing problem in intrusion detection with class imbalance. To improve the detection rate of the prediction model developed in this proposed work, the relevant feature subset which produced relevant information about the pattern of the network traffic is identified by devising neutrosophic correlation-based feature subset

selection. The Neutrosophic correlation handles the impreciseness among the KDD cu 99 dataset. The reduced feature subset enhances the depth understanding of traffic pattern by LSTM but its learning rate is empowered by adapting the knowledge of Whale optimization algorithm. WOA with its expert knowledge of prey hunting is used to fine tune parameters of LSTM to maintain the lengthy sequence of data pattern. The results proved that the enhanced LSTM-WOA effectively handles the issue of impreciseness and vagueness in intrusion detection by increasing the accuracy of abnormal packets more precisely compared to other state of the art algorithms such as DNN, MLP and SVM.

References

[1]. J. A. Khan, N. Jain, "A survey on intrusion detection systems and classification techniques", *Int. J. Sci. Res. Sci. Eng. Technol.*, vol. 2, no. 5, pp. 202-208, 2016

[2]. O. Can, O. K. Sahingoz, "An intrusion detection system based on neural network," in 2015 23nd Signal Processing and Communications Applications Conference (SIU), May 2015, pp. 2302–2305

[3]. Vipin, Das Vijaya, Pathak Sattvik, Sharma Sreevathsan, Srikanth, T, Gireesh. (2010). Network Intrusion Detection System Based on Machine Learning Algorithms. International Journal of Computer Science & Information Technology (IJCSIT), Vol 2, No 6, December 2010.

[4]. Liu, Jingyu Yang, Dongsheng Lian, Mengjia Li, Mingshi, Research on Classification of Intrusion Detection in Internet of Things Network Layer Based on Machine Learning. 106-110. (2021).

[5]. Mohammed Maithem, Ghadaa A. Al-sultany, Network intrusion detection system using deep neural networks 2021 Journal of Physics: Conference Series 1804 (2021) 012138

[6]. C. Yin, Y. Zhu, J. Fei and X. He, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks," in *IEEE Access*, vol. 5, pp. 21954-21961, 2017, doi: 10.1109/ACCESS.2017.2762418.

[7]. R. K. Vigneswaran, R. Vinayakumar, K. P. Soman, P. Poornachandran, "Evaluating Shallow and Deep Neural Networks for Network Intrusion Detection Systems in Cyber Security," 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), July 2018, DOI:10.1109/ICCCNT.2018.8494096

[8]. M. K. Yadav and K. P. Sharma, "Intrusion Detection System using Machine Learning Algorithms: A Comparative Study," 2021 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC), 2021, pp. 415-420, Published 21 May 2021, DOI:10.1109/ICSCCC51823.2021.9478086

[9]. Peng K, Leung VC, Huang Q. Clustering approach based on mini batch Kmeans for intrusion detection system over Big Data. IEEE Access.2018.DOI:10.1109/ACCESS.2018.2810267

[10]. Kathryn-Ann Tait, Jan Sher Khan, Fehaid Alqahtani, Awais Aziz Shah, Fadia Ali Khan, Mujeeb Ur Rehman, Wadii Boulila, Jawad Ahmad, Intrusion Detection using Machine Learning Techniques: An Experimental Comparison, Conference: 2021 International Congress of Advanced Technology and Engineering (ICOTEN), pp 1-10, 978-1-6654-1224-7, July 2021, DOI:10.1109/ICOTEN52080.2021.9493543.

[11]. S. Latif, Z. Zou, Z. Idrees and J. Ahmad, "A Novel Attack Detection Scheme for the Industrial Internet of Things Using a Lightweight Random Neural Network," in *IEEE Access*, May 2020, 8(1):89337 – 89350, DOI:10.1109/ACCESS.2020.2994079

[12]. A. Shenfield, D. Day, A. Ayesh, Intelligent intrusion detection systems using artificial neural networks, ICT Express, vol. 4, no. 2, pp. 95–99, 2018, https://doi.org/10.1016/j.icte.2018.04.003

[13]. Mikolov, T., Karafiát, M, Burget, L., Cernocky, J., Khudanpur, S., Conference: INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, September 26-30, 2010.

[14].Trivedi. I.N. et al, "A Novel Adaptive Whale Optimization Algorithm for Global Optimization," Indian Journal of Science and Technology, Vol 9(38), October 2016, DOI:10.17485/ijst/2016/v9i38/101939.

[15]. F. Smarandache, "Neutrosophic set a generalization of the intuitionistic fuzzy sets," International Journal of Pure and Applied Mathematics, vol. 24, pp. 287–297, 2005.

[16] F. Kuang, W. Xu, and S. Zhang, ''A novel hybrid KPCA and SVM with GA model for intrusion detection,'' Appl. Soft Comput., vol. 18, pp. 178–184, May 2014.

[17] R. R. Reddy, Y. Ramadevi, and K. V. N. Sunitha, ''Effective discriminant function for

intrusion detection using SVM,'' in Proc. Int. Conf. Adv. Comput., Commun. Inform.

(ICACCI), Sep. 2016, pp. 1148–1153.

[18] W. Li, P. Yi, Y. Wu, L. Pan, and J. Li, ''A new intrusion detection system based on KNN classification algorithm in wireless sensor network,'' J. Elect. Comput. Eng., vol. 2014, Jun. 2014, Art. no. 240217.

[19] B. Ingre and A. Yadav, ''Performance analysis of NSL-KDD dataset using ANN,'' in Proc. Int. Conf. Signal Process. Commun. Eng. Syst., Jan. 2015, pp. 92–96.

[20] N. Farnaaz and M. A. Jabbar, ‘‘Random forest modeling for network intrusion detection system,’’ Procedia Comput. Sci., vol. 89, pp. 213–217, Jan. 2016.

[21] J. Zhang, M. Zulkernine, and A. Haque, ‘‘Random-forests-based network intrusion detection systems,’’ IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 38, no. 5, pp. 649–659, Sep. 2008.

[22] J. A. Khan and N. Jain, ‘‘A survey on intrusion detection systems and classification techniques,’’ Int. J. Sci. Res. Sci., Eng. Technol., vol. 2, no. 5, pp. 202–208, 2016.