DOI: 10.21917/ijsc.2020.0305

MODIFIED BEE COLONY WITH BACTERIAL FORAGING OPTIMIZATION BASED HYBRID FEATURE SELECTION TECHNIQUE FOR INTRUSION DETECTION SYSTEM CLASSIFIER MODEL

S. Kalaivani and G. Gopinath
School of Computer Science and Engineering, Bharathidasan University, India

Abstract

Feature selection (FS) plays an essential role in creating machine learning models. The unrelated characteristics of the data disturb the precision of the perfection and upsurges the training time required to build the model. FS is a significant process in creating the Intrusion Detection System (IDS). In this document, we propose a technique for selecting container functions for IDS. To develop the performance capacity of the modified Artificial Bee Colony (ABC) procedure, a hybrid method is presented in which the swarm behavior of the Bacterial Foraging Optimization (BFO) algorithm is entered into the Modified Bee Colony (MBC) procedure to perform a local search. The proposed Hybrid MBC-BFO algorithm is analyzed with three different classification techniques which are separately analyzed to verify the proposed performance. The classification techniques are Artificial Neural Networks (ANN), Recursive Neural Network (ReNN), and Recurrent Neural Network (RNNs). The proposed algorithm has passed several algorithms for selecting advanced functions in terms of detection accuracy, recall, precision, false positive rate, and F-score.

Keywords:

Swarm Intelligence, Modified Bee Colony, Bacterial Foraging Optimization, Feature Selection, IDS, KDDCUP'99

1. INTRODUCTION

Feature Selection (FS) is the process of choosing the factors that furnish the best inputs that would help to strongly structure the intended model [1]. The choice of capacities in this regard should be possible physically or by utilizing different frameworks and calculations. This is a huge advance in building strong Intrusion Detection Systems (IDS) by avoiding irrelevant aspects that might produce bogus alerts and alter the exactness of the framework [2].

An IDS is a product or gadget application for observing system traffic action to recognize vindictive substance or action. An IDS is likewise intended to give admonitions and reports on the revelation of vindictive substance. Despite the way that most IDS plans are planned to distinguish and report dubious movement on the system, it ought to be noticed that best in class plans exist that can square dubious traffic on the system [3].

IDS can be grouped into two classes as indicated by the location strategy [4]. The primary classification is a marked exploration. It utilizes a few examples on the system, for example, bytes, at that point thinks about them to a current mark database. The secondary classification is an internet searcher dependent on irregularity. It works by contrasting the conduct of the system and the introduced base and is truly appropriate for recognizing known and obscure assaults.

IDS forms a lot of information with bogus positives, which are unimportant and insignificant. These features not only slow down the search but also consume many IT resources. There should be a mechanism for choosing only the best features to improve accuracy, driving, and speed tests [5]. The selected activities solve some common problems with identifiers by identifying related activities. As shown in [4], the corresponding functions contain the necessary information, which greatly facilitates the classification process. When selecting functions from IDS, it is important to consider that this reduces processing costs, reduces memory, and improves understanding of test data.

Since FS is a concept of ML, it is mainly applied using a variety of algorithms. FS also uses statistical analysis, supports vector machines, NN, and data mining. Also, the selection of attributes requires a recognition device that can be divided into three groups: random, incremental, and reductive selection [1]. The selection appliance is used to regulate and select relevant attributes from a dataset. It should be noted that FS can be built using a variety of technologies, including intelligence models, cluster information, ANN, critical algorithms, and fuzzy and fuzzy sets [6]. Climate algorithms are mainly used to select functions in infiltration detection systems, particularly because of their high accuracy. In this case, information about weapons is an important technique for using and classifying meta-heuristic algorithms. Animal information is a common and artificial method of promoting the performance of insects and herds. It is used to resolve compound difficulties.

There are many precedents for IDS, but in most of them, the main hassle lies in low performance accuracy, detection, and FAR. Based on observations collected from these issues, this research strives to increase the performance of detection accuracy of IDS by introducing a new Metaheuristic Swarm Intelligence (MSI) based hybrid MBC-BFO algorithm. In this algorithm, the swarm behavior of the bacteria search (swarm mechanism) optimization algorithm [19] is presented into MBC [18] algorithm to perform local research. The proposed methodologies are splitting training and testing datasets for data pre-processing, feature selection and classification.

The rest of this paper is organized as follows: section 2 provides an overview of select precedents related to the proposed work. Section 3 describes the methodology, dataset description, data pre-processing, feature selection based on MSI algorithms, and classification. Section 4 gives performance metrics, experiment results and discussions, and section 5 presents the conclusion.

2. RELATED WORKS

Chen et al. [7] has built up an IDS work choice technique that utilizes a blend of collection calculations actualized utilizing shifting and exemplification strategies. The covering technique utilizes the direct connection coefficient calculation (FGLCC) utilizing the channel strategy calculation (CFA). The proposed technique utilizes the choice tree to make a grouping dependent on the KDD Cup 99 informational index. During the test arrangement, execution assessment depended on exactness, search speed, false positives and fitness capacities. Test results are contrasted with the results obtained by gauges utilizing 10-overlap cross legitimacy and other execution based calculations. The assessment results show that the planned consolidated calculation FGLCC-CFA gives a higher inquiry speed of 95.23%, precision of 95.03% and mistakes of 1.65% with a positive outcome.

Rikhtegar et al. [8] presented a FS model dependent on a cross breed learning system. The IDS instrument joins the choice and gathering of functionalities. The primary uses a SVM while the secondary uses the K-Medoids gathering calculation. Likewise, the methodology additionally utilizes the Nave Bayes classifier for the assessment procedure with the KDDCUP'99 dataset. The projected model is assessed utilizing three fundamental execution measures, including precision, discovery recurrence and caution speed. Execution measures are made utilizing correct positives, right recommendations and bogus positives. The test results were contrasted and the other three characters are choice techniques. Relative strategies incorporate predisposition K-Medoids + GFR + corp, inclination K-Medoids + corp, and 10-overlay + body cross-approval predispositions. The trial results demonstrated that the proposed half and half normalization approach would give better exactness (91.5%), location rate (90.1%), and bogus alert rate (6.36%).

Acharya and Singh [9] proposed another methodology that utilizes the IWD calculation to choose IDS usefulness. IWD is a naturally roused calculation that SVM uses to construct the classifier. IWD is likewise viewed as a multitude insight streamlining calculation that utilizes metric heuristics. This methodology was assessed utilizing the KDDCUP'99 informational collection with execution rules dependent on bogus alerts, recognition speed, and exactness. Besides, the consequences of the tests are contrasted and existing methodologies utilizing bio-inspired calculations. The test results showed that the IWD work choice calculation delivers a high discovery rate (91.35%), a superior accuracy (93.12%) and a low bogus caution rate (3.35%).

Eesa et al. [10] built up a FS model that employs the blend of the ID3 classifier calculation and the honey bee calculation. The ideal known as ID3-BA is intended to improve the choice of the subset of usefulness required in IDS. In this perfect, the Ape calculation is utilized to create the necessary subset of qualities, while the ID3 calculation is utilized to assemble the classifier. The ID3-BA model uses the KDDCUP'99 dataset which comprises 41 highlights for preparing and testing drives. The projected approach is assessed utilizing three unique models: the bogus caution rate (FAR), the recognition rate (DR) and the precision. Exploratory outcomes demonstrate that ID3-BA delivers high estimations of DR (91.02%), AR (92.002%) and lower esteems relating to FAR (3.917%). What's more, the outcomes show that utilizing a subset of capacities rather than all capacities delivers preferable rankings over lower DR, AR and FAR.

Zorarpacı and Özel [11] attitude was made using swarm information improvement methodologies. The system uses a crossbreed mix of differential progression computations and phony bumble bee colonization techniques. The evaluation and course of action process is performed using fifteen instructive assortments gained from the UCI vault. The evaluation system similarly included connections of results procured using the cream methodology and results got using various procedures, including assurance features. Techniques used to take a gander at feature decision join chi-square, information advancement, and CFS. Introduction standards depended on estimations of F-measure, accuracy, execution and search speed, and were tried utilizing a 10-crease grouping test. Simulation results demonstrated that the technique prompt higher order precision.

3. PROPOSED SYSTEM

Reducing dimensionality as a preprocessing step for ML is actual in eliminating irrelevant and redundant data, increasing learning accuracy and refining understanding of results. However, the recent increase in data size represents a major challenge for many methods of selecting and extracting existing functionality in terms of efficiency. In the area of ML and pattern recognition, dimensionality reduction is a significant area in which several methods have been projected. In this document, to progress the presentation of the modified ABC algorithm, we introduce a hybrid MBC-BFO algorithm in which the swarm behavior of the bacteria search optimization algorithm is presented in the MBC algorithm to perform local research. The proposed MBC-BFO hybrid algorithm is analyzed with three different classification techniques which are analyzed separately to verify the proposed performance. The three classification techniques are ANN, recursive neural networks, and RNN. The proposed algorithm has succeeded in several algorithms for selecting advanced work functions in terms of precision, recall, RPF, and F-score. The Fig.1 displays the proposed Hybrid FS block diagram.

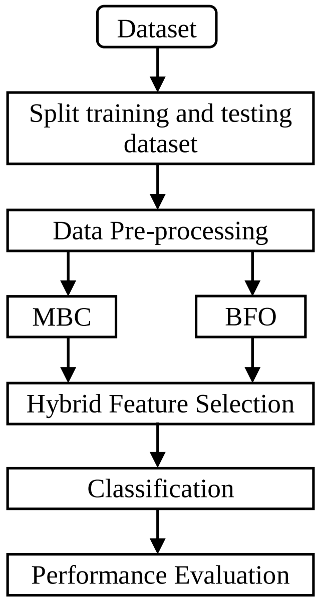

Fig.1. Proposed Block Diagram

3.1 DAPRA KDDCUP'99 DATASET

DARPA dataset was originally established in 1998 with the goal of refining offline intruder detection. Additionally, the 1998 DARPA dataset was established to enhance research and investigation in the field of intruder detection. KDDCUP'99 is an enhanced version of the 1999 DARPA dataset that is used for the improvement of IDS to distinguish between bad and good connections.

The KDDCUP'99 dataset comprises samples of normal attacks and influences, and a 10% KDD training dataset with a goal to train classifiers and a KDD test dataset designed for testing. Attacks can be classified into four main groups and datasets, and the percentages of the total percentage of a given category in a given dataset are revealed in Table.1.

Table.1. Amount of instances in the KDDCUP'99

Attacks	Whole Dataset		10% Training Set		Test Set	
	No. of Instances	(%)	No. of Instances	(%)	No. of Instances	(%)
Normal	492,708	19.86	97,278	19.69	60,593	19.48
DOS	3,883,370	79.30	391,458	79.24	229,853	73.94
Probe	41,102	0.84	4,107	0.83	4,166	1.34
R2L	1,126	0.02	1,126	0.23	16,347	5.26
U2L	52	0.00	52	0.01	70	0.02

- *Denial of Service Attacks (DoS)*: This is an attacker's effort to limit the use of the network by interrupting the availability of the service to the intended users.
- *Probing Attacks*: Happens when the attacker tests the scheme network to gather data about the scheme in order to use it to bypass the organization security check.
- *User to Root Attacks (U2R)*: This happens when an attacker gains access to a regular user account and attempts to gain initial access through a security system.
- *Remote to Local attacks (R2L)*: The attacker does not have a version on the local system, but is trying to gain access to it by sending network packets to exploit susceptibilities and gain access as a local user.

The complete dataset of the KDD Cup 99 contains 4,898,431 instances for training and 311,029 for testing individual connection records, each consisting of 41 functions labeled as normal or attack. However, these paper experiments are used at 10% in the full KDDCUP dataset for the correct training and test set. It is significant to remember that the test dataset comprises novel types of attacks that do not exist in the training set. This dataset comprises 24 types of attacks, while the test set comprises 14 supplementary attacks.

3.2 DATA PRE-PROCESSING

It consists of three chief phases: label transfer, duplication removal, and data normalization. When transferring the label and transferring data, all symbolic data are transported to numerical values. In addition, the entries in the class column are converted into binary classes 0 or 1, where 0 indicates a normal record while 1 indicates an attack record regardless of the type of attack.

It is significant to eliminate duplicate records from the training set to prevent classifiers from being affected by the most recurrent records and to avoid learning of rare records such as the U2L attack. To avoid this, duplicate KDDCUP'99 records have been deleted and the amount of records in the training set after deleting dismissed data is 145,584 illustrations.

$$X_{normalized} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

3.3 FEATURE SELECTION

This process is significant to eliminate the sum of features required to create IDS. In this survey, the container method is used to select the characteristics. The wrapper approach uses blind search to find a subset of functions and randomly searches for the best subset that cannot be guaranteed without obtaining all of the possible subsets. Therefore, the selection of features in this approach is NP-difficult and the search for each iteration tends to be intractable for the user. The advantage of the surround approach [12] is that it selects an almost perfect subset and the error rate in this method is lower than in other methods. The enveloping approach to the selection of functionalities is classified as sequential search and heuristic search.

3.3.1 Modified Bee Colony (MBC) based Feature Selection:

The ABC algorithm examines a new source of food for the bees used and communicates this information to the bees that they see. The viewer uses the food sources used by the bees to explore the other bees. The search equation suggested by the ABC algorithm is good in search, but weak in functionality, so it affects the speed with which the algorithm is integrated. In other words, when a new solution called solution v_i is created as a solution, only one parameter of the main solution v_i is changed, resulting in a slower convergence rate. To recover the use of the ABC procedure, the modified search is considered the best global solution in the Eq.(2). Change the search equation of the ABC algorithm and define it in Eq.(2).

$$v_{ij} = z_{ij} + \varnothing_{ij}\left(z_{ij} - z_{kj}\right) \quad (2)$$

where $\varnothing \in \{1,2,\ldots,S_N\}$ is an arbitrarily chosen record that is unique in relation to i;$j \in \{1,2,...,D\}$is a haphazardly chosen list; w_j is component j of the best worldwide arrangement; $\varnothing_{ij} \in [-1,1]$ and $\varnothing_{ij} \in [-1,-1]$ are consistently circulated irregular numbers.

Differential appraisal (SD) is a population based algorithm whose primary technique is to make new areas for an individual by ascertaining vector contrasts between other arbitrarily chose individuals from the populace. "DE/Current-to-Rand/1" is a type of a DE change methodology that can be utilized to viably keep up populace decent variety dependent on the randomization of the pursuit condition. In light of this honey bee stage transformation methodology, another exploration condition is recommended that is utilized as follows.

$$v_{ij} = z_{ij} + \varnothing_{ij}\left(z_{ij} - z_{kj}\right) + \varnothing_{ij}\left(w_j - z_{ij}\right) \quad (3)$$

where $k \in \{1,2,\ldots,S_N\}$is an arbitrarily chosen file other than i;$j \in \{1,2,...,D\}$ is a record chosen aimlessly; w_j is the j^{th} component of the best worldwide arrangement; $\varnothing_{ij} \in [-1,1]$ and $\varnothing_{ji} \in [-1,-1]$ are consistently dispersed irregular numbers; $\varnothing_{ij}$ and $\varnothing_{ji}$ negative or both positive, which can keep a similar pursuit course.

Algorithm 1 shows a general pseudocode for the altered ABC improvement approach [18].

Algorithm 1: A Modified ABC optimization

Initialization:

Initialize the primary population and estimate the fitness;

Compute the primary fitness value, *f*(*sol*);

Set best solution, *solbest*←*sol*;

Set supreme sum of iteration, *Numofite*;

Set the population dimensions;

//where population size=# of onlooker bees = #of employed bees;

abnd_lst = abandoned list; // extendable list

Limit: set maximum number of trails;

Iteration ← 0;

Improvement:

do while (*iteration*<*NumofIte*)

for *i*=1: # of employed bees

Choice a random solution

Put on random neighborhood structure

end for

for *i*=1: #of onlooker bees

*sol** ← selection solution from population based on roulette wheel choice

Apply a random neighbourhood structure on *sol**

If (*f*(*sol*)< f(*sol**))

*sol** ← *sol*;

else

Limit_list *i*++;

end if

end for

for *i*=1:population size

If (limit_list(*i*) ≥ *limit*)

abnd_lst ← insert *sol*(*i*)

Solbest<best solution found so far;

for *i*=1: # of **scout bees**

solNew (*i*) ← new solution by the scout bee

rand = a random number from 1 to the number of routes in the dataset

solNew (*i*).*rout* (*rand*) ← *solbest*.*Rout*(*rand*)

reschedule (*solnew*(*i*))//keep the solution feasible choice solution from abnd_1st based on roulette wheel selection and replace it with *solNew* (i);

apply all the neighborhood structure on the *solNew* (*i*);

end for

end if

end for

iteration ++;

end do

3.3.2 BFO based Feature Selection:

Utilizing E.Coli Scavenging Theory of Behavior in light of science and material science, Passino and Liu [10] applied a progression of amassing and bacterial scrounging practices, additionally scanned for the correspondence between the E.Coli control framework and search execution [19]. All microscopic organisms attempt to build the focus slope of the food independently. Every bacterium is created by the four phases (chemotaxis, swarm, generation, disposal, and dispersal) to locate a superior food source. The BFO are accounted for in [11] [13]. The procedures followed in this process are as listed below:

- *Chemotaxis*: E.coli bacteria can transfer biologically in two diverse ways. It often swims in the same way or can fall and alternate among these two types of operations during its lifespan.
- *Swarming*: When enthused by a high level of succinate, the cells release an attractive aspartate, which enables them to assemble into different subgroups, and therefore move in teams with high bacterial density.
- *Reproduction*: The weakest bacteria eventually die, but every healthy bacterium produces its offspring in the same place. It maintains a stable layer shape.
- *Elimination and Dispersal*: In the local atmosphere, the life of a population of bacteria may change gradually or suddenly due to flu. To pretend this occurrence in BFO, some bacteria are randomly killed with a very low probability, while new substitutes are randomly initialized in the research space.

3.3.3 Hybrid MBC-BFO Feature Selection Technique:

MBC is a set of optimization algorithms that have proven to be more efficient than other algorithms such as PSO and ACO. Since its discovery, it has been of great interest to researchers in various fields because of its small control parameters, high research capability, and ease of implementation. Although MBC is well scanned, the main disadvantage is its poor performance, which in some cases can cause a problem with the quality of the solution.

In the authoritative MBC calculation, the condition for refreshing the arrangement of a fundamental ABC presents a few issues, for example, wastefulness during a neighborhood search in the arrangement space. In actuality, in spite of the fact that the BFO calculation has a moderate combination, it has the best ability to arrive at the worldwide ideal. It must be said that the multitude system of the BFO calculation makes the microorganisms to bunch together and later move into a concentric model with high bacterial thickness. In this manner, microscopic organisms that have arrived at the best food-borne way should leave other microbes a similar way with the goal that they can discover their goal all the more rapidly and precisely. So, to improve the proficiency of the MBC, we propose the multitude component of the BFO calculation in the period of the pre-owned honey bees and in the period of observer honey bees of the MBC calculation. The main phases of the hybrid MBC-BFO algorithms are listed in Algorithm 2 and the swarm mechanism pseudocode is described in Algorithm 3.

Algorithm 2: Main steps of the Hybrid MBC-BFO algorithm

Step 1: Initialize variables and randomize positions;

Step 2: **While** ((Iter < MaxCycle))

Step 3: /*Employed BeesÿPhase*/

For(i=1:(FoodNumber))

Produce a new food source;

Evaluate the fitness of the new food source;

Swarming mechanism defined in Table.2.

Greedy selection;

End For

Step 4: Calculate the probability P;

Step 5: /*Onlooker Bees Phase*/

For(i=1:(FoodNumber)) Parameter P is determined randomly;

Onlooker bees find food sources depending on P;

Produce a new food source;

Evaluate the fitness of the new food source;

Swarming mechanism; Greedy selection;

End For

Step 6: /*Scout Bees Phase*/

If(any employed bee becomes scout bee) Parameter P is determined randomly;

The scout bees find food sources depending on P;

End If

Step 7: Memorize the best solution;

Iter=Iter+1;

End While

Algorithm 3: Swarming mechanism of the BFO algorithm

Step 1: Initialize variables

Step 2: Let $m = 0$;

Step 3: While $m < N_s$

Step 4: If (the mutant solution is better than the current Solution) Update the solution by the mutant solution;

Step 5: End If Let $m = m + 1$;

Step 6: Else, let $m = N_s$.

3.4 CLASSIFICATION

Classifications are data organized in predefined groups. This is done using a model based on previous data. A classification algorithm learns from the training set and creates a predictive model to classify the “normal” or “attack” connection. In this proposed system, there are three different classification techniques which are analyzed separately to verify the proposed performance. The three techniques are ANN, Recursive Neural Networks (ReNN), and Recurrent Neural Network (RNN).

3.4.1 Recurrent Neural Networks:

It is clear that the construction of the RNN model involves two fragments such as forward and backward propagation [14]. Advanced forwarding is responsible for scheming output values and backward propagation is accountable for promoting accumulated debris to inform weight, which is not significantly diverse from the normal formation of the NN.

The objective function associated with RNN for a single training pair is defined as [26], where L is a distance function that events the deviation of the predictions and of the real labels. Let η be the learning rate and k be the amount of ongoing iterations with an assumption of sequence of labels.

3.4.2 Artificial Neural Networks (ANN):

FFNN are known as a renowned class of ANN-based neural simulations that are capable of creating and approximating complex copies based on their higher-level parallel layer structure. The basic elements of Feed Forward Neural Network (FFNN) processing are a series of neurons. These neurons propagate in several fully attached charged layers. One of the generalized examples of FFNN is the multilayer perceptron (MLP). In MLP [15], the initial processing elements are arranged in a unidirectional fashion. In these networks, information evolves on the basis of communications between three types of coincident levels: the entry, hidden and exit levels.

The nets between these levels are associated with certain weight values which vary between [-1,1]. Two functions can be performed on each MLP node, called addition and activation functions. The product of the input values, weight values and polarization values can be obtained based on the sum function described in the Eq.(4)

$$S_j = \sum_{i=1}^{n} \omega_{ij} I_i + \beta_j \tag{4}$$

where n represents the entire sum of inputs, I_j is the input variable I_i, β_j is a bias value, and w_{ij} reveals the weight of the connection. In the next step, an activation function based on the result of the equation is activated as in Eq.(4). Numerous activation methods can be used in MLP which, according to the collected works, the most used is the S-shaped sigmoid function. This function can be designed based on the Eq.(5)

$$f_j(x) = \frac{1}{1 + e^{-S_j}} \tag{5}$$

Therefore, the final output of neuron j is obtained using the Eq.(6)

$$y_i = f_j\left(\sum_{i=1}^{n} \omega_{ij} I_i + \beta_j\right) \tag{6}$$

After building the final ANN structure, the learning process to refine and develop the weighting vectors of the network is encouraged. These weight vectors must be updated to approximate the results and optimize the total error of the network.

3.4.3 Recursive Neural Network (ReNN):

Like a great neural system structure, standard RNN comprehends the task of deduction in complex representation.

Congregations of Self-Assertive Size [16]: At the point when a sentence is given, RNN examines it in a paired semantic tree and figures the vector show of each word. During the learning time of direct spread, the ReNN computes the principle vectors in rising request. The piece condition is the accompanying condition as in Eq.(7):

$$p_1 = f\left(W\begin{bmatrix} c_2 \\ c_3 \end{bmatrix} + b\right)$$

$$p_2 = f\left(W\begin{bmatrix} c_1 \\ p_1 \end{bmatrix} + b\right) \tag{7}$$

where f is the activation function; $W \in R^{d\times 2d}$ is the weight matrix, where d is the dimensionality of the vector; and b is bias. Hence, each parent vector p_i is given as a characteristic to a softmax classifier as defined in the Eq.(8) to calculate the probability of the label:

$$y^p = \text{softmax}(W_s.p) \tag{8}$$

where $W_s \in R^{3\times d}$ is the classification matrix. In this recursive process, the vector and the classification result of the node will gradually converge. After giving the leaf node vector, the RNN can finally map the semantic demonstration of the entire tree to the root vector.

4. RESULTS AND DISCUSSION

In this section, the effectiveness of the MBC method is validated using various parameters, such as accuracy, precision, recall, and false positive percentage. The experiments are performed in the repository of the KDDCUP'99 dataset listed in Table.1.

The KDDCUP'99 [17] datasets contain four different categories such as Normal, R2L, DOS, U2R and Probe. MBC is an autonomous field that is tested by these datasets to detect implicit and explicit aspects for more effective results. The evaluation of the parameters of the proposed hybrid method and its validated results are reported in the following sections. The experiments are implemented using Python 3.7.3 on a computer with a 2.2 GHz Intel Core i5 processor with 8.00 GB of RAM.

4.1 PERFORMANCE METRICS

There are numerous solutions to assess FS procedures. The size chosen depends on the nature of the application. Most researchers measure your ID using TPR and FPR performance measures. In this unit, we define all key performance indicators that are used to appraise the projected method. All selected measurements can be considered using the exit from the confusion table.

A confusion table is signified by four main parameters:

- *True Positive* (*TP*): Sum of Attack examples classified properly.
- *True Negative* (*TN*): Sum of Normal examples classified properly.
- *False Negative* (*FN*): Sum of Attack examples wrongly classified as normal.
- *False Positive* (*FP*): Sum of normal examples wrongly classified as an attack.

Definition of performance metrics and formulas according to the confusion matrix as follows [17]:

- *Sensitivity*: Measures the percentage of effective attacks.
- *Accuracy*: It measures the percentage of correct ranked classes out of the total sum of rankings.
- *False Positive Rate*: It measures the percentage of normality identified as an attack.
- *F-score*: Degree the accuracy of the model by seeing both accuracy and recall.

4.2 PERFORMANCES ANALYSIS

Table.5. Performance analysis of different feature selection techniques with various classifiers by using 10% training set KDDCUP dataset

Techniques	PR	R	F-M	FPR	ACC
MBC-ANN	97.50	97.03	97.50	11.08	97.90
MBC-Recurrent	97.50	97.23	97.85	10.50	97.80
MBC-Recursive	97.65	97.56	97.58	10.80	98.20
BFO-ANN	97.12	97.36	97.01	11.05	97.23
BFO-Recurrent	97.32	97.54	97.42	11.35	97.55
BFO-Recursive	97.86	97.64	97.35	11.60	97.60
Hybrid-ANN	97.90	97.89	97.87	08.30	98.01
Hybrid-Recurrent	97.12	97.64	98.01	08.20	98.05
Hybrid-Recursive	98.05	98.25	98.23	08.04	98.30

The Table.5 illustrate the performance analysis of the selection of various characteristics with three classifiers such as ANN, recurrent and recursive methods using the 10% learning set of the KDDCUP'99 dataset. The precision obtained in the experiment when MBC is combined with the ANN classifier is 97.90%, the recurrent neural network is 97.80% and the recursive neural network is 98.20%. When BFO is combined with the ANN classifier, it is 97.23%, the recurrent neural network is 97.55% and the recursive neural network is 97.60%. Therefore, the proposed hybrid method achieved accuracy when combined with the ANN classifier is 98.01%, the recurrent neural network is 98.05% and the recursive neural network is 98.23%.

Based on the result, the proposed hybrid approach (MBC-BFO) with the recursive neural network (RNN) approach achieved higher classification accuracy than existing ones using a 10% learning set.

Table.6. Performance analysis of different feature selection techniques with various classifiers by using 100% whole KDDCUP dataset

Techniques	PR	R	F-M	FPR	ACC
MBC-ANN	91.45	90.89	91.90	10.08	92.10
MBC-Recurrent	90.80	91.80	91.70	10.50	92.05
MBC-Recursive	91.70	91.90	91.95	10.00	92.20
BFO-ANN	90.25	90.21	90.07	10.05	90.80
BFO-Recurrent	90.21	90.28	90.75	10.35	90.65
BFO-Recursive	90.45	90.43	90.90	9.60	91.23
Hybrid-ANN	93.05	93.23	94.07	08.30	94.09
Hybrid-Recurrent	94.24	94.78	94.90	07.20	94.90
Hybrid-Recursive	94.80	94.85	94.95	06.04	95.05

The Table.6 illustrate the analysis of the performance of the selection of different characteristics with three classifiers such as ANN, the recurrent and recursive methods using the KDDCUP'99 dataset at 100%. The precision obtained in the experiment when MBC is combined with the ANN classifier is 92.1%, the recurrent neural network is 92.05% and the recursive neural network is 92.2%. When BFO is combined with the ANN classifier, it is

90.8%, the recurrent neural network is 90.65% and the recursive neural network is 91.23%. Therefore, the achieved accuracy of the proposed hybrid method when combined with the ANN classifier is 94.09%, the recurrent neural network is 94.9% and the recursive neural network is 95.05%.

From the result, the proposed hybrid approach (MBC-BFO) achieved higher classification accuracy than the existing approach using 100% complete KDDCUP'99.

5. CONCLUSION

In this paper, we propose a new hybrid algorithm for selecting MBC-BFO functionalities based on a nature-inspired optimization for the IDS classifier model. Its goal is to reduce the number of features required to create a robust IDS while preserving a high detection rate, and accuracy with low false alarms. The benchmark components (10% and 100%) of KDDCUP'99 dataset is used to train and test the classification model for better detection frequency. In addition, we experimented with different analysis strategies for the features selection with various classification techniques, such as machine learning and the deep learning method, taking different dimensions from the reference dataset. The result of the experiment shows that the proposed meta-heuristic hybrid based on the swarm intelligence of the MBC-BFO algorithms offers better performance for high detection accuracy, precision, recall, F-Measures and low false positive rate. In the future, we aim at applying this hybrid technique to improve the accuracy of IDSs in detecting different attacks by using other datasets. It is about measuring your ability to work on various characteristics of the data. Also, its future goal is to reduce execution time and complexity.

REFERENCES

[1] H. Liu and H. Motoda, "*Feature Selection for Knowledge Discovery and Data Mining*", Kluwer Academic Publishers, 2012.

[2] X. Tang, Y. Dai and Y. Xiang, "Feature Selection based on Feature Interactions with Application to Text Categorization", *Expert Systems with Applications*, Vol. 120, pp. 207-216, 2019.

[3] K. Scarfone aand P. Mell, "Guide to Intrusion Detection and Prevention Systems (IDPS)", Technical Report, National Institute of Standards and Technology, pp. 1-78, 2012.

[4] S. Mohammadi, H. Mirvaziri, M. Ghazizadeh Ahsaee and H. Karimipour, "Cyber Intrusion Detection by Combined Feature Selection Algorithm", *Journal of Information Security and Applications*, Vol. 44, No. 2, pp. 80-88, 2019.

[5] S. Zaman and F. Karray, "Features Selection for Intrusion Detection Systems based on Support Vector Machines", *Proceedings of 6th IEEE International Conference on Consumer Communications and Networking*, pp. 1-8, 2009.

[6] S. Maza and M. Touahria, "Feature Selection Algorithms in Intrusion Detection System: A Survey", *KSII Transactions on Internet and Information Systems*, Vol. 12, No. 10, pp. 1-14, 2018.

[7] K. Chen, F.Y. Zhou and X.F. Yuan, "Hybrid Particle Swarm Optimization with Spiral-Shaped Mechanism for Feature Selection", *Expert Systems with Applications*, Vol. 128, pp. 140-156, 2019.

[8] M. Keshtgary and N. Rikhtegar, N., "Intrusion Detection Based on a Novel Hybrid Learning Approach", *Journal of AI and Data Mining*, Vol. 6, No. 1, pp. 157-162, 2018.

[9] N. Acharya and S. Singh, "An IWD-Based Feature Selection Method for Intrusion Detection System.", *Soft Computing*, Vol. 22, No. 13, pp. 407-416, 2018.

[10] A.S. Eesa, Z. Orman and A.M.A. Brifcani, "A New Feature Selection Model Based on ID3 and Bees Algorithm for Intrusion Detection System", *Turkish Journal of Electrical Engineering and Computer Sciences*, Vol. 23, No. 2, pp. 615-622, 2015.

[11] E. Zorarpaci and S.A. Ozel, "A Hybrid Approach of Differential Evolution and Artificial Bee Colony for Feature Selection", *Expert Systems with Applications*, Vol. 62, pp. 91-103, 2016.

[12] Barnali Sahu,Satchidananda Dehuri and Alok Jagadev, "A Study on the Relevance of Feature Selection Methods in Microarray Data", *The Open Bioinformatics Journal*, Vol. 11, No. 2, pp. 117-139, 2018.

[13] Swagatam Das, Arijit Biswas, Sambarta Dasgupta and Ajith Abraham, "*Bacterial Foraging Optimization Algorithm: Theoretical Foundations, Analysis, and Applications*", Foundations of Computational Intelligence, Vol. 3, pp. 23-55, 2009.

[14] L.P. Dias and J.J.F. Cerqueira, "Using Artificial Neural Network in Intrusion Detection Systems to Computer Networks", *Proceedings of 9th IEEE International Conference on Computer Science and Electronic Engineering*, pp. 1-8, 2017.

[15] J. Martens and I. Sutskever, "Learning Recurrent Neural Networks with Hessian-Free Optimization", *Proceedings of 28th IEEE International Conference on Machine Learning*, pp. 1-6, 2011.

[16] Xuegong Zhang, Xin Lu and Qian Shi, "Recursive SVM Feature Selection and Sample Classification for Mass-Spectrometry and Microarray Data", *BMC Bioinformatics*, Vol. 7, No. 19, pp. 1-18, 2006.

[17] V.R. Shewale and H.D. Patil, "Performance Evaluation of Attack Detection Algorithms using Improved Hybrid IDS with Online Captured Data", *International Journal of Computer Applications*, Vol. 146, No. 8, pp. 1-12, 2016.

[18] S. Kalaivani and Gopinath Ganapath, "Bio-Inspired Modified Bees Colony Feature Selection based Intrusion Detection System for Cloud Computing Application", *International Journal of Advanced Science and Technology*, Vol. 29, No. 3, pp. 1-12, 2020.

[19] S. Kalaivani and Gopinath Ganapath, "Bacterial Foraging Optimization for Enhancing the Security in Intrusion Detection System", *International Journal of Scientific and Technology Research*, Vol. 9, No. 2, pp. 1-8, 2020.