International Conference on Cyber Security and Computer Science (ICONCS'18),
Oct 18-20, 2018 Safranbolu, Turkey

A hybrid intrusion detection system based on Sine Cosine Algorithm and Naïve Bayes algorithm

SALAAD MOHAMED SALAAD [1] and ERKAN ÜLKER[2]

[1] Selçuk University, Konya /Turkey, daauudmsalaad@hotmail.com
[2] Konya Technical University, Konya/Turkey, eulker@konyateknik.edu.tr

Abstract - **Due to improving technology and spreading internet the entire world, people adapted using it in an extensive manner. Our critical private data are encountering threads which are coming outside of the computer systems and network environments. In other word, intruders access folk's information without authentication and unauthorized mode. To overcome such kind of security vulnerability matter, a lot of scientific researchers have attracted their awareness the use of this new model called hybrid intrusion detection systems, which is an integration of two or more algorithms, then one of the algorithms is utilized as input while the functionality of other one is tasking or classifying. The new model has a very powerful and plays a significant role in cybersecurity. In recent years, the combination of machine learning methods with metaheuristic algorithms is hybridized to obtain an optimum solution. In this study, we present a new model using the Sine Cosine Algorithm for feature selection and the Naïve Bayes Classifier (NBC) algorithm for classification. Our main goal is to find a model that emphasizes a good performance for detecting and finding preferable accuracy. We compare our experimental results with other algorithms such as KNN, Decision Tree classifications and etc., to realize which one is performed an excellent in terms of accuracy and detection rate. İn addition to this, Sine Cosine Algorithm will be contrast to Particle swarm optimization (PSO), not only PSO but also genetic algorithm (GA) in terms of feature reduction and selection the quality ones, various datasets such as NSL-KDD, ISCX 2012 and etc., has been applied on the new presented method to examine its performance. Finally, the introduced method will prove that whether it has better performance and superior accuracy compared to the other algorithms.**

Keywords - Sine-Cosine Algorithm, KNN, Naïve Bayes Algorithm, and Particle swarm optimization, ISCX 2012, Hybrid intrusion detection system.

I. INTRODUCTION

The last few years cybersecurity is one of the most crucial subjects of concern the information security researchers. Since the internet is developed, many technology application devices are developed which are running on the internet and people's needs depended on the internet for instances online money transfers, transactions, online businesses & etc. Meanwhile, attackers are also increasing day by day and their aim is to violate the policy of the security issue by breaking--the low of the computer systems, such as taking an action that. Threats to the confidentiality, integrity, and availability of the System [1].To struggle with these malicious movements of cyber-attack, many search experts regularly creates a powerful tool to acquire an optimum solution which able to halt this threat circumstances. Creating a good intrusion detection system was one of the attempts. İntrusion detection system is an integral both hardware and software, which is designated to check all incoming packets on a host or a network by distinguishing them from normal or abnormal traffic.. Essentially an IDS implementing can be considered as a classification procedure, which is emphasizing to improve classification performance in order to prevent the intruders [2]. IDS is an intelligent indicator tool which monitors computer system activities and reports back if there is any unwanted movement in the security strategy. The main goal of IDS is to confirm the three main security fundamentals Confidentiality, availability, and integrity of system information [3]. Confidentiality – means the information can only reveal the authorized or certified ones. Integrity – is that the data cannot be change or demolish in an unauthorized fashion. Availability – the system must be handiness or accessible by the approved individuals any time. Keep in mind that availability is one of the most prime of the three secure system.

Moreover there are traditional guards which play a role to detect hosts and networks from malicious interest such as firewalls, user authentication, and encryption method, but unfortunately, occasionally this technique faces a lake of protection, they are not intelligence as IDS which can monitor and block the abnormal network activities that are why we put IDS after firewalls in the network layer. Additionally, during the analyzing process of packet checking, the IDS focuses on the various field of the packet such as IP addresses of the packet source, service, flags and etc. while firewall only investigates a few parts of the packet [4].

In 2005 the FBI (Federal Bureau of Investigation) with the Computer Security Institute cooperatively organized security survey and annual computer crime, eventually, they specified that *$130 million* economic losses of companies and agencies caused by network hackers. Hence intrusion detection is a vital research difficulty in cybersecurity, by the way, 1980 the conception of intrusion detection presented by *Anderson*. IDS is a paramount tool that controls the entire network security activities and traffic exchanges packets, while a doubtful activity or malicious packet is prevented, it immediately makes alarm.

Intrusion detection approaches are categorized into two misused and *anomaly*. Misused (knowledge- or signature-based) all signatures and characters of attack signs are known and stored them in the database, this technique can only protect known attack signatures by comparing its pattern signatures and the newly arrived packet signatures, if they are same the packet is considered as malicious activity, otherwise; it is recognized as normal. In contrast anomaly (behavior based) IDS is based on statistical behavior which deals with user change activities in the network [5]. Network and host security systems are facing variety kind of attacks such as *flooding* (denial of a service, which makes a system too busy), *port scanning* (which scans the vulnerability of the system), *password guessing* (tempting to log in a system with unauthorized way) and finally *buffer overflow attacks* (attacker access the root system as one of the normal users, benefiting this opportunity he can steal or even change the system normal functionality). Because of these unwelcomed activities IDS is a very necessary task which follows some rules to supply a secure system by differentiating between permitted activities and unjustified use of the system [6].

Every network Organization uses their own suitable rules and especially data sets to keep their business and private movement from invaders. There are two ways to implement the detection rules one way is system-linked another way is third-party-integrated software system is implemented each and every system and network, for instance, antivirus, internet security services, firewall, data encryption and system of network detection services. IDS is divided into six types. Host (detects a single system) and network (protects the entire network), active (functions when any abnormal activity did not happen) and non-active (passive: take saving action if any malicious events are registered in the system security), misused and behavior based. But the main ones are misused and anomaly intrusion detection system [7]. Any system security has weakness parties which are hard to solve and even cost a lot of money to fix them by the manufacturers [8]. According to all above refer to, our goal is to find an outstanding IDS, to achieve that we have use a feature selection method to increase efficiency and to have good accuracy results. *Feature selection* is selecting and reducing a proper dataset features into small subsets, which is equivalent to the whole original features. There is three type of feature selection methods are types, *wrapper*, *filter* and *embedded* method [4, 10]. The filter mothed is not depended on the classification algorithm to pass the features, e.g. correlation coefficient, mutual information. The wrapper method is based on learning or classification algorithm in order to judge the optimum features which are evaluated as the best feature ones during the feature selection process. While embedded is defined as a process of established with the structured method, e.g. LASSO, regularization methods.

Nowadays many research workers have utilized metaheuristic algorithms for feature reduction, and they mentioned these optimization techniques are preferred as they have been obtained better results. *Optimization* is described as the procedure of minimizing or maximizing the optimal output values of a given variable to the system. An optimization problem is focused on as a *black box* according to a stochastic optimization problem, means no need any induction mathematical model since it considers the input system swaps and gives an output. The second benefit is the high flexibility, the stochastic algorithm is capable to accept problems in various sides when the problem is assumed as black boxes [9]. Comparing the stochastic algorithm to the conventional optimization algorithms, they naturally advanced from elevated (higher) local optima avoidance. As you can recognize from the name stochastic, the algorithm randomly selects the problem. There are *three species of researches*. Incorporate the variety of algorithms (hybridizer), presenter of new methods, and upgrader of the existing algorithms.

The metaheuristic creativity is *based on six main* parts

- *Evolutionary-based algorithms*: like Genetic Algorithms (GA), Differential Evolution (DE).
- *Swarm intelligence* algorithms: such as PSO, and Bee Colony algorithm.
- *Algorithms derived from Physics*: Black Hole (BH).
- Human-associated *algorithms*: some of them are Mine Blast Algorithm (MBA), and Teaching Learning-Based Optimization (TLBO).

The above-mentioned algorithms reduce dimensionality and computational cost. In addition to this, it improves the classification accuracy as they come up with optimal and satisfying results in a short time.

A hybrid model with Sine-Cosine algorithm (SCA) is presented. This metaheuristic algorithm recommends by Seyedali Mirjalili in 2015. It is one of the population and optimization methods. The mathematical model of the algorithm was derived from Sine Cosine function. In this method the feature reduction and search strategy are done by SCA and classification evaluation is being used is done by Naive Bayes classifier (NBC). The classification implementation process the NBC classified the whole data into two labels normal class and attack class. We also used an evaluation function or metric that is responsible to test the performance of the classification method by computing the regular performance measures such as accuracy, recall, specificity, f-measure, sensitivity and g-mean.

In this study, NSL-KDD is used which is intrinsic (inherent) from the original KDD CUP99 [2, 4, 5, 7]. However, selecting the most significant features meaning improving classification and obtaining the best optimal results. The standard KDD CUP99 is presented by Stolfo and Lee. Moreover, the KDD CUP99 dataset is favored and it is the most suitable data for anomaly detection. But there are two main problems which the dataset has, the first one is influencing the system evaluation performance, and the second one is leading the model system to achieve unacceptable and low results. To fix these two complications, a new version of the dataset named NSL-KDD is produced from the KDD CUP99. Our presented model selects the best optimal subset features with little time and optimum classification performance.

The following parts of the paper is scheduled into sections. Related works regarding to IDS that was done before will momentarily be debate in section II. Section III covers the proposed system and methodology of SCA for IDS. Experiment and their results are referred to in Section IV. The final portion is the Conclusion, which is mentioned in Section V.

II. Related works

Many researchers believe that obtaining a higher and efficiency classification performance is dependent on having minimum features, which is good for detecting malicious activities. In order to achieve this main goal, again authors suggested using a hybrid algorithm which is combining variety algorithms, each one of this algorithm has its own function, which is one of this algorithm is used as feature selection while another one is responsible classification process. [10] Assembling algorithms are two ways by the filter method or the wrapper method. The wrapper is connected to the classification algorithm to predict accuracy and estimate the best feature. In contrast the filter method in an independent to the other method such classification ones. A system based on negative selection algorithm with enhancing incremental PSO is presented. According to their experimental results using NSL-KDD, they obtained an accuracy of 97.75 % [11]. A model of is random forest and Average One-Dependence Estimator (AODE) was suggested, the performance estimation of the method implemented Kyoto data, 90.51% accuracy and FAR (False Error rate) is 0.14 [12]. Another author used SVM for feature selection, k-Medoids is utilizing as training data creation, then Naïve Bayes classification for evaluation. To appraise the model the KDD CUP'99 dataset is used. The suggested system performs an accuracy of 91.5%, a detection rate of 90.1 % and a false alarm rate of 6.36% [13]. An author [2] presented a model of intelligent water drops (IWD) with support vector machine (SVM). Feature selection is used by IWD and SVM responsible classification performance. To check their system performance, they carry out the KDD CUP'99 dataset. Lastly, rate of 99.4075% was obtained, and 99.0915%,1.405, 99.108 accuracy, false alarm rate ,precision respectively. Binary Particle Swarm Optimization (BPSO) joining with decision tree (DT) is presented [3]. In 2016, a hybrid model is presented [7], which is based on Multi-Class-Classification Based MCLP (Multiple Criteria Linear Programming) with PSO. And their final result is defined as this detection rate is 99.34%, accuracy is 99.14%, False alarm rate is 1.765%.

III. Proposed hybrid method for IDS

The core goal of this study is to obtain an IDS which is efficiency and has optimum accuracy and ability to detect malicious activities that involve and violates the stability maintains of the security system. According to the researcher's statements in related work, the hybrids that are based on the metaheuristic algorithms is more valuable, more achievable and even has a good a curacy and detection rates compared with the hybrids which are not based on metaheuristic techniques [2]. When we say a hybrid model means that integrating two or more algorithms, while one of the algorithms used as feature selection so that the outcome of this algorithm will be utilized as the input of the classification algorithm. Therefore the presented model is based on the combination of SCA with NBC algorithm.

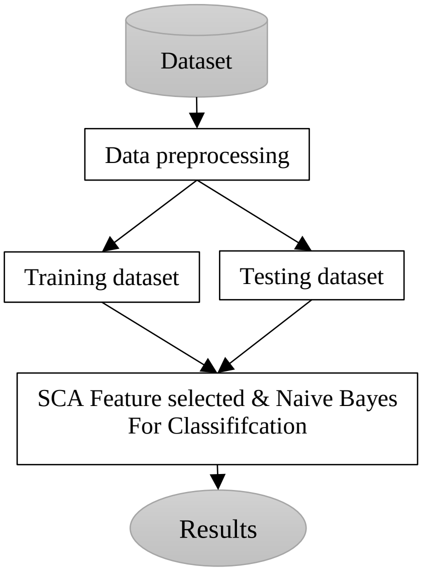


Figure. 1 Presented model for IDS.

3.1 Data Preprocessing

We going to discuss the way we have done the data preprocessing steps before the dataset implemented on dimensionally reduction and classification technique. The Dataset records consist of normal and intrusion. The instance dataset combines related records of extracted features. These attributes can be either symbolic or continuous, we converted the symbolic to numerical attributes. İn the process we also included Normalization. Normalization is required because of data scaling requirement before training and testing set in the same range, the values of each feature are normalized as in the range of [0,1], The normalization formula is defined as follows.

$$V = \frac{V-Min_A}{Max_A-Min_A}(newMax - newMin) + newMin \qquad (1)$$

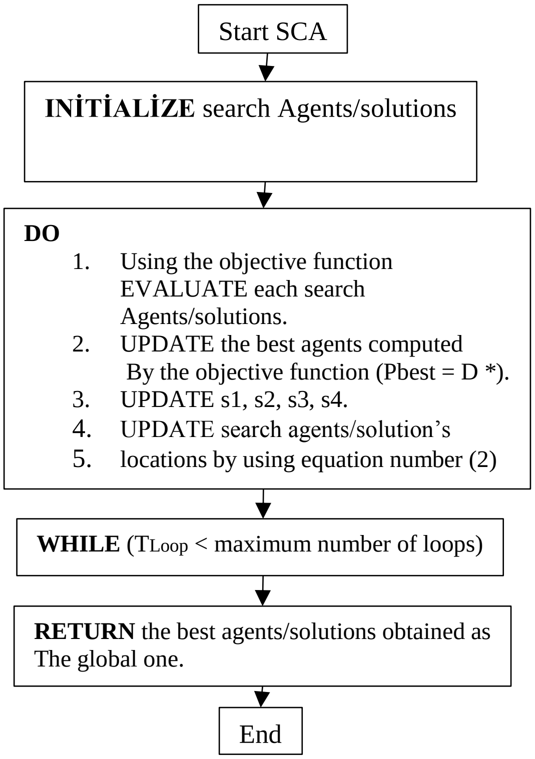
3.2 Feature selecting SCA Algorithm

3.2.1 The search mechanism of SCA

The sine-cosine algorithm is a stochastic population optimization algorithm, which is based on since and cosine function, it proposed by Seyedali Mirjalili in 2015. The algorithm is work as follows; first, it initializes random solutions or what we call search agents in the search space map problem with the random position. Each search agent stars looking its own way finding an optimum solution in the search space, the objective function is to a computer each agent's best-achieved solution so far, the best agent's location among them is denoted as P at each loop. To update this the position of the agents we use equation (2).

$$Lij^{(t+1)} = \begin{cases} Lij^{t} + s1 * sine(s2) * |s3\ Pj^{t} - Lij^{t}| & s4 < 0.5 \\ Lij^{t} + s1 * cos(s2) * |s3\ Pj^{t} - Lij^{t}| & s4 \geq 0.5 \end{cases} \quad (2)$$

Where $Lij^{(t+1)}$ is the update position of the agent i_{th} at iteration (t +1) at j dimension, Lij^{t} is the current location of the solution, for now *s1*, *s2*, *s3*, and *s4* are random parameters.



Start SCA
INİTİALİZE search Agents/solutions
DO
1. Using the objective function EVALUATE each search Agents/solutions.
2. UPDATE the best agents computed By the objective function (Pbest = D *).
3. UPDATE s1, s2, s3, s4.
4. UPDATE search agents/solution's
5. locations by using equation number (2)
WHILE (TLoop < maximum number of loops)
RETURN the best agents/solutions obtained as The global one.
End

Figure. 2 SCA steps

S1 is responsible for pointing the direction and the next location's region, the new position will be either the distance between the solution and destination when s1 <1 that means the movement goes toward the destination or outward of the destination if s1 >1. This is the main reason why *s1* parameter checks the balance of exploration and exploitation of the algorithm stages. For each loop *s1* linearly decreases from constant value of *c* to zero (0) [14, 15] .The balance equation is as follows:

$$s1 = c - t\frac{c}{Tmax} \quad (3)$$

Where *c* is constant, *t* is the current loop, *Tmax* is the maximum iteration. *S2* indicates how far away the movement is closing to the destination or going far from the destination. *S3* brings a random weight that is multiplied by the destination to stochastically emphasize when *s3* <1 or deemphasize when *s3* > 1. Parameter *s4* switches sine and cosine components in an equal way.

3.2.2 The Feature selection using SCA

The SCA is being utilized to reduce the features and improve the classification performance by applying the given training data and saving small features. Each agent's performance is computed by fitness function based on selected features and accuracy. Features of data set are equivalent to the utilized variables. Agents in the search space are demonstrated as a binary vector, which their length is identical to dataset features. In order to restraint whether of every single feature, the contestant is selected (1) or not selected (0) in the classification process, we computed the fitness of every single agent in equation (4) [17, 18].

$$fij = \begin{cases} 1, & if \quad Lij^{(t+1)} > \beta \\ 0, & otherwise \end{cases} \quad (4)$$

Where fij is a fitness of each agent, $Lij^{(t+1)}$ is the updated position of agent *i* with dimension *j* and *β* is the threshold, which decides whether the feature is select or not. Its random value is between the range of [0,1] .

Moreover, Naïve Bayes classification computes the probability accuracy of the class. Naïve Bayes theorem is defined in equation (5).

$$P(S|M) = \frac{P\ (M|S)P(M)}{P(S)} \quad (5)$$

Where P*(S)* is an independence probability of S, P*(S)* is an independence probability of M, *P (M* | S*)* coordinate a likelihood, *P (M) is prior probability and P (S* | M) equivalent to Posterior Probability.

It is phenomenal that our fitness function is associated with both classification accuracy and dimensionality reduction. Naïve Bayes classifier is a supervised learning algorithm which is based on the Bayesian Theorem. We trained and tested the new model, the task of the test is examined by using the data set to evaluate how clear our model is learned.

IV. EXPERİMENTAL STUDY AND RESULTS

The sine cosine is being used for feature minimization [18,19]. And Naïve Bayes classification is utilized to classify the accuracy performance of the reduced features. *1600* of NSL-KDD99 records (instances) were randomly selected, using random selection. The normal ones are *747* records while the malicious is *853*. The used NSL-KDD99 dataset, we split it into two subsections training set and testing set. The training set is 1440 instances, while the testing set is 160 records**.**

Table 1
The confusion matrix of two classes

		Predicted class	
		Abnormal	Nomaal
Actual class	Abnormal	TP	FN
	Nomal	FP	TN

Where TP is true positive, FN is false negative, FP is false that is classified as positive and TN is true negative.

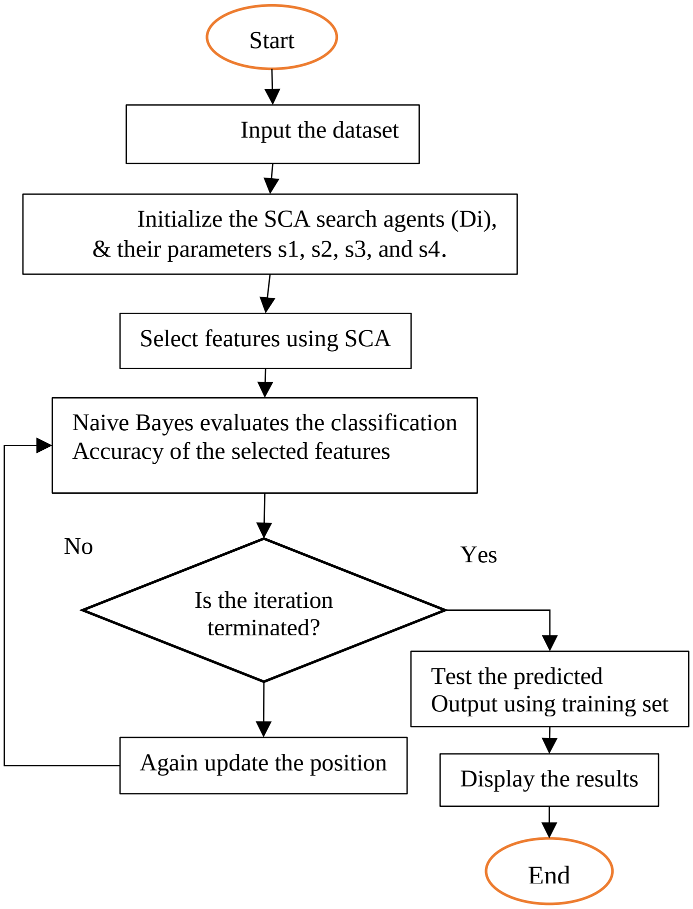

Start
Input the dataset
Initialize the SCA search agents (Di),
& their parameters s1, s2, s3, and s4.
Select features using SCA
Naive Bayes evaluates the classification
Accuracy of the selected features
No
Yes
Is the iteration
terminated?
Test the predicted
Output using training set
Again update the position
Display the results
End

Figure. 3 The presented model of SCA-NBC

$$\text{Accuracy(AC)} \;=\frac{tp+tn}{tp+tn+fp+fn}$$
(6)

The number of data that is correctly classified divided the number of complete data. To compute the accuracy assessment, testing data is utilized. The best accuracy is 1, while the worst is 0.

$$sensitivity(SE)=\frac{tp}{tp+fn}$$
(7)

Sensitivity is the positive simples divided all positive simples.

$$Specificity(SP)\;=\frac{\boldsymbol{tn}}{\boldsymbol{tn+fp}}$$
(8)

Specificity is the simples which the model took as a negative over true negative with false positive.

$$\text{Precision(PR)}\;=\frac{tp}{tp+fp} \qquad (10)$$

Precision is number of data that classified as positive divided by positive data.

F-score is a combination of precision and sensitivity, which is named harmonic.

$$F-measure(FM)=2*\frac{Precision*Sensitivity}{Precision+Sensitivity}$$
(11)

F- Measure means precision multiply by sensitivity over precision with sensitivity

Parameters	**Values**
SCA identifications	
Population size	41
Feature numbers	41
Instances	1600
Maximum iteration	100
μ (equalization factor) or constant	0.999
c (Constant value linearly dropping *from c to 0)*	2
PSO identifications	41
Population size	41
Feature numbers	1600
Instances	100
Number of iteration	C1 = 2
Coefficient1 (C1)	C2 = 2
Coefficient (C2)	100
Maximum iteration	0.9
Inertia weight	

Table 2. The of SCA & PSO parameters

The experimental of the proposed method is tested 30 times, and the maximum iteration number is 100. **Table 3** demonstrates the SCA-Naïve Bayes measurement performance such as, accuracy (AC), specificity (SP), precision(PR), sensitivity (SE), false positive rate (FPR), error rate (ER) and f-score (FS).

Table 3. Experimental of SCA with NBC using NSL-KDD dataset

SCA with NBC	
AC	99.64%
SE	99.50%
SP	96.%
PR	99.8%
FPR	0.0101%
FS	98.80%
ER	0.016%
Attributes	16

In table 4 we compared the performance measurement of our new model (SCA-Naïve Bayes) and PSO-Naive Bayes [20]. by using the same NSL-KDD dataset, 41 attributes, and 100 iterations. As illustrated in table 4, the new method performces well contrasting with PSO-Naive Bayes.

Table 4. Comparing SCA- NBC and BPSO- NBC using same NSL-KDD dataset

	SCA- NBC	PSO-NBC
NSL-KDD	NSL-KDD	NSL-KDD
AC	99.64%	98.12%
SE	99.50%	99.18%

SP	96.%	94.73%
PR	99.8%	98.37%
FPR	0.0101%	0.0125%
FS	98.80%	98.75%
ER	0.016%	0.018%
Attributes	16	18

In **table 5**, decision tree and Naïve Bayes classifiers was compared to see their detection performance by using the NSL-KDD, as clarified in table 5, each one of these classifiers are good for a specific part of performance measures, such as Naïve Bayes is better than the decision tree in Sensitivity performance, while Decision tree has superior performance of all performance measures except in Sensitivity [20].

Table 5. Naïve Bayes is compared to Decision tree using NSL-KDD dataset

Decision tree	Naive Bayes	
	NSL-KDD	NSL-KDD
AC	96.69%	89.10%
SE	89.19%	92.15%
SP	99.49%	85.52%
PR	98.51%	88.18%
FPR	0.038%	0.15%
FS	93.62%	90.12%
ER	0.0331%	0.11%

V. Conclusion

The objective of this study was to present a new model for IDS which is Sine Cosine (SCA) algorithm with Naive Bayes Classification. Where the feature reduction of a given dataset is being used by the SCA, and Naive Bayes is evaluated classification accuracy performance. The design of the fitness function is included both increasing accuracy performance and choosing minimum features. We have done a comparison between the SCA feature selection consequence (result) and the PSO algorithm, eventually, we substantiated that the SCA is better than PSO in both classification performance and feature minimization as shown in the experimental results. We mentioned to use not only NSL-KDD99 but also ISCX2012 for performance evaluation and KNN to be a part of the comparison, for limitation of deadline submission. In this study, the dataset implemented on the new model is NSL-KDD99, the test and results of ISCX2012 and KNN are in progress to compare our new system .

References

[1] Shailendra and Silakari, "*A survey of Cyber Attack Detection Systems*," IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.5, May 2009

[2] Neha Acharya & Shailendra Singh, "An IWD-based feature selection method for intrusion detection System," *Springer-Verlag Berlin Heidelberg,* DOI 10.1007/s00500-017-2635-2, 2017

[3] Arif Jamal Malik1 & Farrukh Aslam Khan, "A hybrid technique using binary particle swarm optimization and decision tree pruning for network intrusion detection," Springer Science+ Business Media, LLC. DOI 10.1007/s10586-017-0971-8, 2017

[4] Vajiheh Hajisalem & Shahram, "A hybrid intrusion detection system based on ABC-AFS algorithm for misuse and anomaly detection," Department of Computer Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran. Elsevier, 2018

[5] Jasmin Kevric1 Samed Jukic1 & Abdulhamit Subasi, , "*An effective combining classifier approach using tree algorithms for network intrusion detection.* ", Neural Comput & Applic, DOI 10.1007/s00521-016-2418-1, 2016

[6] Salma Elhag & et al, "*A multi-objective evolutionary fuzzy system to obtain a broad and accurate set of solutions in intrusion detection systems.*" Springer-Verlag GmbH Germany 2017, DOI 10.1007/s00500-017-2856-4

[7] A M VISWA BHARATHY & A MAHABUB BASHA, "*A multi-class classification MCLP model with particle swarm optimization for network intrusion detection.*" Department of Computer Science and Engineering, K.S.R. College of Engineering, Tiruchengode 637215, Sadhana Indian Academy of Sciences, DOI 10.1007/s12046-017-0626-8

[8] Arvinder Kaur & et al, "*Hybridization of K-Means and Firefly Algorithm for intrusion detection system.*" Received: 28 February 2017, the Society for Reliability Engineering, Quality and Operations Management (SREQOM), Int J Syst Assur Eng Manag, DOI 10.1007/s13198-017-0683-8

[9] Seyedali Mirjalili, "*SCA: A Sine Cosine Algorithm for solving optimization problems.*" a School of Information and Communication Technology, Griffith University, Nathan Campus, Brisbane, QLD 4111, Australia b Griffith College, Mt Gravatt, Brisbane, QLD 4122, Australia. 2015 Elsevier B.V. All rights reserved

[10] Khalil El-Khatib, Member, "Impact of Feature Reduction on the Efficiency of Wireless Intrusion Detection Systems", IEEE, VOL. 21, NO. 8, AUGUST 2010

[11] Manikandan & G. Bhuvaneswari, "An intelligent intrusion detection system for secure wireless communication using IPSO and negative selection classifier for secure wireless communication using IPSO and negative selection classifier" , Springer Science +Business Media, LLC, part of Springer Nature 2018, https://doi.org/10.1007/s10586-017-1643-4

[12] M A Jabbar & et al. "RFAODE: A Novel Ensemble Intrusion Detection System". 2017 The Authors. Published by *Elsevier* B.V. 7th International Conference on Advances in Computing & Communications, ICACC-2017, 22-24 August 2017, Cochin, India

[13] L. Khalvati & et al, "Intrusion Detection based on a Novel Hybrid Learning Approach", Shiraz, Iran. Received 27 August 2016; Revised 01 February 2017; Accepted 03 June 2017.

[14] Hafez& et al, "Sine Cosine Optimization Algorithm for Feature Selection", Scientific Research Group in Egypt (SRGE), http://www.egyptscience.net, 2016 IEEE.

[15] Huiwen Wang & et al, "An effective intrusion detection framework based on SVM with feature augmentation", journal homepage: www.elsevier.com/locate/knosys, 2017 Elsevier B.V. All rights reserved

[16] Rana Aamir Raza Ashfaq & et al, "Toward an efficient fuzziness based instance selection methodology for intrusion detection system", Int. J. Mach. Learn. & Cyber. (2017) 8:1767–1776,DOI 10.1007/s13042-016-0557-4, Springer-Verlag Berlin Heidelberg 2016

[17] Mohamed Issa & et al, "ASCA-PSO: Adaptive sine cosine optimization algorithm integrated with particle swarm for pairwise local sequence alignment", journal homepage: www.elsevier.com/locate/eswa, 2018 Elsevier Ltd. All rights reserved

[18] R. Sindhu & et al, "Sine–cosine algorithm for feature selection with elitism strategy and new updating mechanism", Received: 14 October 2016/Accepted: 2 January 2017, The Natural Computing Applications Forum 2017, DOI 10.1007/s00521-017-2837-7

[19] Mohamed E. Abd Elaziz & et al, "A Hybrid Method of Sine Cosine Algorithm and Differential Evolution for Feature Selection", conference Paper · October 2017,DOI: 10.1007/978-3-319-70139-4_15

[20] Abdullahi Hussein ABDULLAHI, "AN INTRUSION DETECTION APPROACH BASED ON BINARY PARTICLE SWARM OPTIMIZATION AND NAIVE BAYES", Master, University of Selçuk, Turkey