International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-1, October, 2019

Effective Parameter Optimization & Classification using Bat-Inspired Algorithm with Improving NSSA

R.Mohanraj, S.Anbu,

Abstract ***Network Security is an important aspectin communication-related activities. In recent times, the advent of more sophisticated technologies changed the way the information is being sharedwith everyone in any part of the world.Concurrently, these advancements are mishandled to compromise the end-user devices intentionally to steal their personal information. The number of attacks made on targeted devices is increasing over time. Even though the security mechanisms used to defend the network is enhanced and kept updated periodically, new advanced methods are developed by the intruders to penetrate the system. In order to avoid these discrepancies, effective strategies must be applied to enhance the security measures in the network. In this paper, a machine learning-based approach is proposed to identify the pattern of different categories of attacks made in the past. KDD cup 1999 dataset is accessed to develop this predictive model. Bat optimization algorithm identifies the optimal parameter subset. Supervised machine learning algorithms were employed to train the model from the data to make predictions. The performance of the system is evaluated through evaluation metrics like accuracy, precision and so on. Four classification algorithms were used out of which, gradient boosting model outperformed the benchmarked algorithms and proved its importance on data classification based on the accuracy obtained from this model.***

Keywords: Bat Algorithm, Gradient Boosting, KDD Cup 1999, Machine Learning, Network Security, Optimization.

I. INTRODUCTION

Internet becomes an essential part of our daily life and it involved in several critical operations for various purposes. The network-dependent applications consume most part of internet everyday such as financial transactions, online purchases, social media network, etc. As the demand is increasing more and more, the threats over internet-related activities have been raised up together. The increase in consumer satisfaction highly relies on the trustworthiness of the service provider. So, more security should be provided to improve the reliability and strengthening the quality of the network. Intrusion in-network is a common problem addressed over decades after the explosion of internet worldwide. It focuses on compromising the network by violating the integrity, availability, and confidentiality of the network resources.

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
R.Mohanraj*, Research Scholar, Department of Computer Science and Engineering,St.Peters Institute of Higher Education and Research, Avadi, Chennai. Email : mohanraj254@gmail.com.
Dr.S.Anbu, Professor, Department of Computer Science and Engineering, Peri Institute of Technology,Avadi, Chennai.

The concept of intrusion detection is proposed in 1980 by Anderson [1]. It states that the intruders are different in their behavior when compared with legal users. Intrusion Detection System (IDS) aims to identify the attacks performed on the network and to track and monitor the activities of the network to enhance the security of the network from anonymous intruders. Network IDS is an effective tool that manages to restrict the illegal access to the resources. Also, it locates suspicious behavior in network traffic and handles malicious requests. Anomaly Detection and Misuse detection are the two types of IDS categorized based upon their functionalities [2]. Malicious behavioral pattern is captured by correlating with normal activities occurred over the network in the anomaly detection system, whereas misuse detection system identifies the intruders based on suspicious activities. These profiles in anomaly detection system can be identified through developing models that contain metrics, derived from the operations on the system. These metrics are computed from the system parameters such as number of processes performed by a user, number of connections made per minute, CPU load, etc. [3]. Based on the data-source, IDS can be further divided into two types, host-based, and network-based detection [4]. In the preceding method, audit data can be used to detect intrusions like logs, usage time of CPU, systems call, etc. Network-based IDS uses network transfer data, payloads and packet headers collected from Network Interface Card (NIC) [5]. These data can be analyzed with sophisticated intelligent learning algorithms and statistical models to track the pattern of malicious behavior over normal activities. In this paper, machine learning-based effective intrusion identification and classification framework are proposed. Heuristic optimization algorithms were employed to select optimal parameters for better classification. Supervised machine learning algorithms are applied on the feature-selected dataset and the best model is identified through performance evaluation methods. The rest of the paper is organized as follows. In section 2, the existing literature related to this work is briefed. Section 3 explains themethodologies incorporated to develop this system in detail view. The resultsare projected in section 4 with concise graphical representations. Section 5 concludes the study by highlighting the significance of the work.

II. BACKGROUND STUDY

Advancement in communication technology improved the way information is shared across any part of the world. At the same time, attacks over the internet are increased and new techniques being handled to intensify the process.

Identification of malicious behavior in a network channel could minimize the strength of the attack. Many methodologies have been developed to detect threats in the network. In recent times, machine learning-based automated traffic analysis systems are proposed by experts to prevent the network from the intruders. A hybrid deep belief network-support vector machines algorithm (DBN-SVM) is developed to identify intrusion in network. The dimension of the data is reduced with DBN and classification is performed with SVM [6]. In another attempt, hierarchical clustering algorithm is applied on ID data to select features and SVM model is trained to classify the malicious and non-malicious samples [7]. A detailed analysis is made on KDD cup 99 datasets and the performance is evaluated with machine learning algorithms. The outcome of the study revealed that random forest algorithm shows significant results over other algorithms such as J48, CART, NB and SVM [8]. Hierarchical neural networks are an ensemble of networks developed to identify the intrusion pattern from the data. The multi-level structure is adopted to design this classifier under different hierarchies with radial basis function [9]. Through statistical analysis, two issues that affect the model performance in KDD'99 dataset have been identified and a new dataset is proposed from the previous one as NSL-KDD [10].

An ensemble approach is proposed by fusing SVM and decision tree to improve the predictive performance as base classifiers. The results illustrate the performance is improved with hybrid model over individual learners [11]. A correlation-based hybrid model is proposed to select features from intrusion detection data. Two objectives are highlighted as the first insight is design a model from classification algorithm, soft computing methods and clustering techniques. Next objective targets filter and wrapper method to perform effective feature selection [12]. Self-organization maps (SOM) are constructed in a hierarchical design to investigate the performance on IDS. The number of features is reduced from 41 to 6 through this method. Also, the system justifies that it is capable of detective false positives (FP), which is an important criterion to evaluate model performance and reliability [13].

Genetic Algorithm bounded with SVM improved the overall performance of the system in intrusion identification and classification. Optimal feature subset is produced by GA is inputted to SVM, generates the best possible results [14-15]. In a novel attempt, a comprehensive view on misuse detection is depicted. This study reveals the reason for failure in misuse detection in intrusion detection data by the machine learning algorithms with proper evaluation results [16]. The selection of features for IDS is performed based on feature relevant analysis. In this experiment, preprocessing is made through discretization and through calculating information gain; the features are selected and evaluated with machine learning models [17]. Discriminative multinomial Bayes algorithm finds the best pattern on classifying attacks over normal samples with various filtering techniques and achieved less FP rate [18].
A combination of multiple classification algorithms enhanced the model accuracy with lower FP rate. Principal Component Analysis (PCA) reduced the dimensionality of the data and RBF-DT model discriminates the samples [19]. SVM IDS proved a better model for intrusion detection as the problem becomes a multiclass-classification scheme. SVM on RBF kernel works well under these conditions [20]. Long short term memory based intrusion detection system is proposed using recurrent neural network model to improve the classification performance of intrusion samples [21]. A novel extreme learning machine-SVM guided by modified k-means algorithm is developed to detect network intrusions [22]. Deep learning has become more popular due to its best performance in various applications. In the process of identifying intrusion in a network, RNN algorithm is deployed to train and evaluate the model. Moreover, J48, RF, SVM, and Artificial Neural Network (ANN) models are benchmarked with the performance of RNN. The result clarifies the efficacy of deep learning methodologies over traditional machine learning algorithms [23]. In another approach, deep neural networks (DNN) model projected on intrusion classification task with 6 features as the inputs for the model [24].

A deep belief network (DBN) and probabilistic neural network (PNN) technique are proposed to overcome the existing problems identified in the process of detecting intrusions. Such problems are said to be local optima, redundancy in information, high dimension of the data and longer training time. In initial phase, the essential features are sorted out with DBN from its non-linear property. The number of hidden nodes is selected from PSO and PNN classifies the lower dimensional data. The outcome reveals the proposed DBN-PNN model outperforms regular PCA-PNN, PNN, and DBN-PNN (non-optimized model) [25].

III. MATERIALS AND METHODS

1.1. Dataset Information

KDD 1999 cup dataset is accessed to perform the experimental analysis [26]. The dataset contains 41 attributes and a class variable. The class attribute holds 5 categories of labels such as DOS (Denial of Service), Probing (Surveillance), U2R (Unauthorized access to root privileged users), R2L (Unauthorized access from a local machine) and normal (No attack). The number of samples present in the dataset is not distributed evenly but still, with the available data, the model is prepared after processing the data. The information about the dataset is given in Table 1.

Table 1: Dataset Description

Dataset Details	
Data Source	Kdd 1999 (Corrected)
Number Of Instances	494021
Number Of Features	41
Classes	Multiclass Classification (5 Types)

3.2 Data Pre-Processing

Discretization is a technique applied to transform the data into granules that helps to manage and evaluate the numbers with ease. This process converts the continuous data into a short-range of values that normalizes the data and distributes values within limited bound [27].

Table 2:Original Data

0	http	13787	1	0	0	2	2	0	177	0.01	NORMAL
0	http	3542	1	0	0	12	12	0	187	0.01	NORMAL
0	http	753	1	0	0	21	22	0.09	196	0.01	NORMAL
0	http	9235	1	0	0	5	5	0	58	0.05	NORMAL
0	http	185	1	0	0	3	3	0	255	0	NORMAL

Table 3: Data after Discretization

0-0.5	http	2453.5-7299	0.5-1	0-0.5	0-0.5	5.5-14.5	7.5-9.5	0-0.005	8.5-14.5	0-0.005	NORMAL
0-0.5	http	197.5-1183.5	0.5-1	0-0.5	0-0.5	5.5-14.5	7.5-9.5	0-0.005	14.5-52.5	0-0.005	NORMAL
0-0.5	http	1208.5-2186.5	0.5-1	0-0.5	0-0.5	5.5-14.5	7.5-9.5	0-0.005	14.5-52.5	0-0.005	NORMAL
0-0.5	http	1208.5-2186.5	0.5-1	0-0.5	0-0.5	5.5-14.5	4.5-7.5	0-0.005	14.5-52.5	0-0.005	NORMAL
0-0.5	http	1208.5-2186.5	0.5-1	0-0.5	0-0.5	5.5-14.5	4.5-7.5	0-0.005	14.5-52.5	0-0.005	NORMAL

The variations observed in Table 3 denote the features are directly affected by the discretization process;it boosts-up the feature selection process, due to the normalized form. In previous case, the data becomes too vague and sometimes it turns out to be an impossible task to select best features. Heuristic optimization algorithms are computationally expensive when compared with statistical methods. So, the normalized model speeds up the feature selection process.

3.3 Feature Selection

Feature selection is an important phase in the machine learning process. The selection of right parameters could impact the outcome of the prediction results. Filter, embedded, hybrid, wrapper techniques are commonly used feature selection techniques [28]. In filter method, the features are selected from statistical model and metrics. A learning model finds the best subset through random selection of features is the process followed in wrapper method. The fusion of wrapper and filter method brings the concept of hybrid models. Regularization adopted techniques are called as embedded techniques. Through penalization, best performing features are sorted out to make better predictions [29].

IV. OPTIMIZATION ALGORITHMS

Heuristic optimization techniques are developed after inspired by the natural processes [30]. These methods are effective in identifying the best parameters for the learning models. It follows a way to identify optimal solution in search space that incorporates finite bound solutions. In most cases, these techniques are profound and it adapts the unique behavior of animals and insects. Some examples are Ant Colony Optimization, Bee Search Optimization, Cuckoo Search Optimization, Firefly Optimization, etc. [31].Each of these algorithms has its own strength and weakness based on the procedure it follows from the behavior of the natural process. In this paper, to select the best features from the intrusion dataset, three meta-heuristic search optimization algorithms are used.

4.1 Particle Swarm Optimization

It is a computational technique, which optimizes a given problem by trying to improve the candidate solution in iterations with respect to a quality measure fixed to evaluate the betterment of the identified solution in search space.It is a population-based, stochastic algorithm inspired by the behavior of flock of birds or fish schooling [32].

Algorithm :Particle Swarm Optimization

1. Initialization of Parameters:
   a. Find the initial population;
   b. Determine the fitness function(f);
   c. Set initial velocity;
   d. Set Pb=f, Gb=min(f); (Gb is Gbest and Pb is Pbest)
2. Repeat:
   until the stopping criteria is met:
   a. Update Gb and Pb:
      For i=1:PS (Pop size)
         i. If f_i<Pb(i)
            Pb(i)= f_i
            End
         End
         ii. b=min(f);
         iii. If b<Gb
            Gb=b;
         End
   b. Generate the next set of the population from (2) and (3)
   c. Checking the feasibility of the solutionstrategy.
   d. Calculate the objective function

Obtain the best solution

4.2 Bat Algorithm

Bat Algorithm is a meta-heuristic search optimization algorithm, developed based on the echolocation behaviors of the bats. This technique is applied to the intrusion detection dataset to sort out the best-performing feature subsets [34]. The steps involved in BA are described in next section.

Algorithm : Bat Algorithm

1. Objective Function definition
   $f(z)$, $z = (z1, ..., zd)^T$
2. Initializing the population of bats
   z_i $(i = 1, 2, ..., n)$ and v_i
3. Define the pulse frequency of f_i at z_i
4. Loudness A_i, Pulse rates r_i initialization.
5. while (t <Maximum no. of iterations)
6. Update velocities, solutions/locations and adjust frequency to generate new solutions [Steps 2 to 4]
   i. if (random > r_i)
   ii. From the best solutions, select one solution.
   iii. Among the selected best solution, generate a local solution.
   iv. end if
   v. Generate new possible solution from the random flying
   vi. if (random < A_i& $f(z_i) < f(z_*)$)
   vii. Accept new solutions
   viii. Reduce A_i and Increase r_i
   ix. end if
7. Rank selected bats and find the current best z_*
8. end while

Visualization and Post process the results

Table 4: Feature subset selected in common by Optimization Algorithms (11 from 41 features)

Parameters	Data Types
duration	Continuous
service	Symbolic
dst_bytes	Continuous
logged_in	Symbolic
su_attempted	Continuous
num_shells	Continuous
count	Continuous
srv_count	Continuous
srv_diff_host_rate	Continuous
dst_host_count	Continuous
dst_host_srv_diff_host_rate	Continuous
Class	DOS, U2R, T2L, PROBE, NORMAL (Categorical)

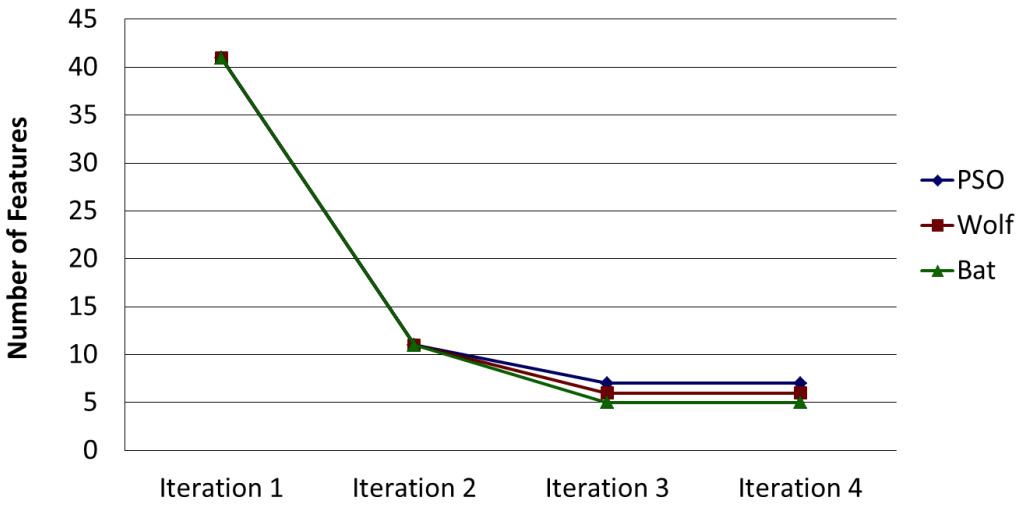
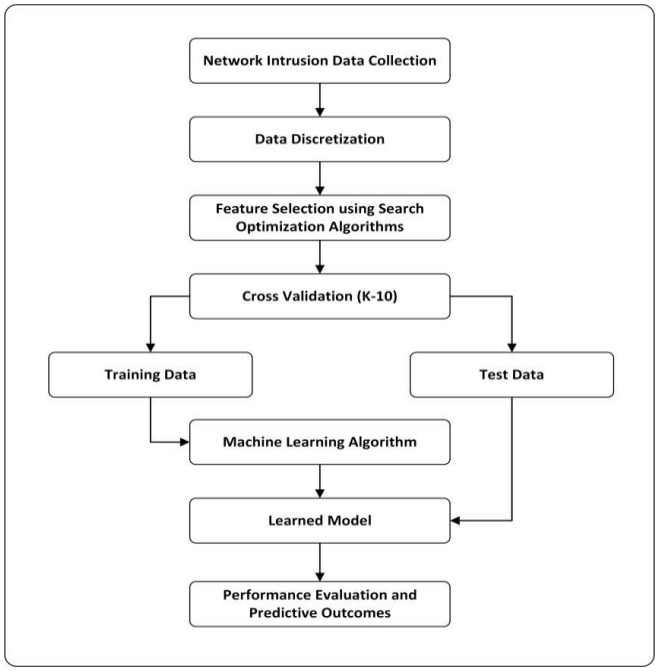


Fig 1: Number of features selected by heuristic algorithms oneachiteration

From the above graph, bat algorithm generates optimal subset with 5 parameters which is the least than Wolf and PSO that identifies 6 and 7 features respectively. As the number of features is minimal in bat identified subset, the performance of the model significantly increased due to the less complexity.



Fig; 1: Architecture of the proposed model

V. CLASSIFICATION

5.1 Machine Learning

Machine Learning is an integral part of Artificial intelligence. The ability to learn through experience by a machine without explicitly programmed with various sophisticated mathematical functions, logical models and massive real-time data [35]. The fields adopting machine learning are heterogeneous. Few domains are finance, healthcare, sales, government and so on.Each application has its own complex problems to be solved and machine learning provides support for most of the cases with optimal solutions. In predictive analytics, ML models are widely used to find the patterns from loan assessment, fraud detection, and purchasing schemes, etc. Most of the recommendation systems, social media analytics, and sentiment classification models are deployed through standard ML algorithms. Futuristic applications were strongly adapted to machine learning models to make their system to perform intelligent actions over traditional systems.

5.2 Supervised Learning

The data that contains independent feature variables and dependent target labels are collectively known as labeled data. The algorithms trained by the labeled information are called supervised learning. Classification and regression are two major types of supervised learning. The inputs are directly mapped to some discrete outcome through functions in classification problems. In case of regression, the input is mapped into continuous outcome [36].

5.3 Unsupervised Learning

Most of the real-time data collected from different sources are unstructured. To identify the pattern from huge data, where the samples do not belong to a specific group, unsupervised learning comes into play. The algorithms under this technique find some specific pattern in the data that deviates from other samples in the data [37]. It groups them together and forms the clusters, where each of the clusters holds the samples of similar pattern.

5.4 Semi-Supervised Learning

This method combines the properties of both supervised and unsupervised learning methods. In some cases, few data are labeled and other left unlabeled. So, initially, the unlabeled data is grouped together with reference to the labeled one. Together, the data will be labeled and trained through some supervised algorithm [38].

5.5 Reinforcement Learning

An algorithm or an agent that learns by itself through interaction with the surroundings and its factors is called reinforcement learning. It involves performing a few actions, gains rewards, and punishments based on its response perceived from the scenario. The goal of the model is to increase the performance to maximum level by learning from the environment and acting upon it [39].

5.6 Gradient Boosting Algorithm

Gradient Boosting Model, an ensemble learning technique is used for gene expression data classification [40]. An ensemble of decision trees together forms GBM in a stage-wise manner. Any loss function, which can be differentiable, is optimized using GBM. To minimize the loss, gradient descent is introduced, which is also called as functional gradient descent. An intuitive explanation of GBM algorithm is given below

Algorithm : Gradient Boosting Algorithm

1. Initialize $f_0(x) = \arg\min_\gamma \sum_{i=1}^{N} L(y_i, \gamma)$.
2. For m = 1 to M:
   a. For I = 1,2,...,N compute
   $$r_{im} = -\left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right] f = f_{m-1}$$
   b. Fit the regression trees to the targets r_{im} giving terminals regions
   $R_{jm}, j = 1, 2, \ldots J_m$.
   c. For $j = 1, 2, \ldots J_m$ compute
   $$\gamma_{jm} = \arg\min_\gamma \sum_{x_i \in R_{jm}} L(y_i, f_{m-1}, (x_i) + \gamma).$$
   d. Update $f_m(x) = f_{m-1}(x) + \sum_{j=1}^{Jm} \gamma_{jm} I(x \in R_{jm})$.

Output $f_x = f_m(x)$.

VI. MODEL EVALUATION AND PERFORMANCE EVALUATION METRICS

For model evaluation, many techniques are available such as train-test split, k-Fold cross-validation, Leave-One-Out Cross-Validation (LOOCV), stratified k-Fold cross-validation, etc. In this system, k-Fold cross-validation techniques are adopted to evaluate the model. In k-fold, 10 folds were applied on the dataset.

The performance of a ML model is evaluated based on proper evaluation metrics. Confusion Matrix, an important metric used to evaluate the performance of classifiers. It can be calculated using four important factors such as true positive (TP), false positive (FP), true negative (TN) and false-negative (FN).

It identifies the correctly and incorrectly classified instances from the samples given to test the model. Accuracy, precision, sensitivity, recall, specificity and f-measure were calculated for proposed model to evaluate its performance [41].

6.1 Results and Discussion

This system is implemented in Windows 10 Operating System. Machine learning libraries were accessed through Anaconda Distribution and Pycharm Integrated Development Environment supports the development process as modules in different phases. Discretization technique is applied to the raw dataset to transform into short range of values. In next phase, three heuristic optimization algorithms were employed to select best feature subset. Bat search algorithm identified well-performing features over PSO and Wolf search process. The performance of classifiers on optimal features selected by bat algorithm is depicted as graph in Fig 3.

Table 5: Performance of the algorithms after feature selection with BA

Algorithms	ACCURACY (%)	SENSITIVITY (%)	SPECIFICITY (%)
LR	95.73	96.32	94.25
NB	94.64	95.93	93.36
SVM	96.29	97.08	94.86
GB	98.61	99.21	96.64

K-Fold Cross Validation is adopted with 10 folds each as 9 for training and 1 for testing during every cycle. The reduced subset is inputted into different learning algorithms. Overall, gradient boosting shows better performance than the benchmarked models SVM, RF, LR, and NB. 98.61% accuracy is obtained through BA-GB model. Other metrics such as sensitivity, specificity and ROC proves the same. In fig 4, the accuracy calculated for all the three heuristic algorithms with ML classifiers is projected.

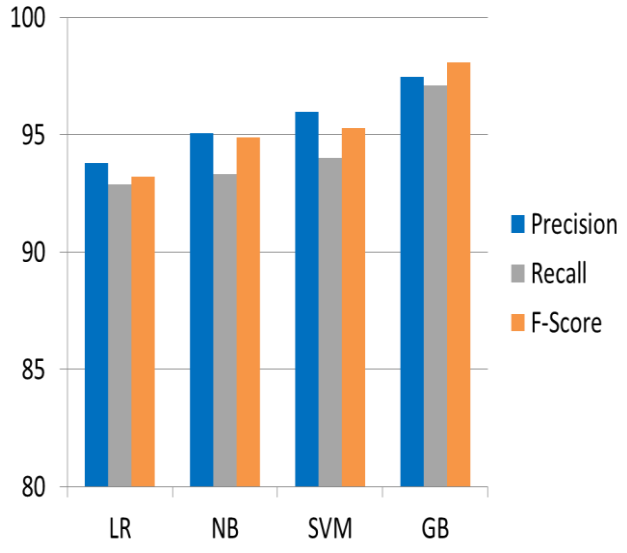


Fig 3: Performance of classifiers with optimal parameters of BA

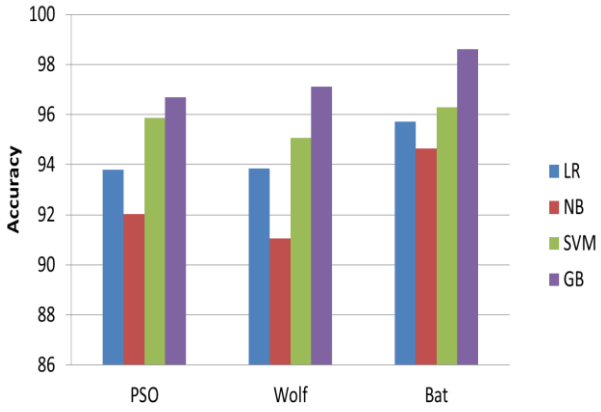


Fig 4: Accuracy obtained by the classifiers on different heuristic optimization algorithms

The role of machine learning in critical applications is highly significant that could manage to identify the hidden patterns from the data generated from the sources [42]. This improves the understandability of the complex processes performed behind the simplified representation projected to the high-end users. Moreover, these models are mathematically proven, reliable and can be evaluated with well-defined validation schemes. Over the next decade, the data produced from automated machines would be huge and are being effectively handled by high-performance intelligent algorithms.

VII. CONCLUSION

In this paper, a novel intrusion classification system is proposed to classify malicious samples over normal records. The dataset is discretized to transform continuous data with minimal loss into a finite set of intervals. Optimal parameters for effective classification are made through heuristic algorithms. Bat Algorithm generated best subset of features with 5 features out from 11 in the complete set. As the number of features is less and the model accuracy is more, the performance of the system becomes optimal.Fourmachine-learning classification algorithms were employed for model preparation and training. Gradient boosting model shows better performance over other models in terms of various evaluation metrics. Together, BA-GB algorithm obtained 98.61% accuracy on classifying the samples. The outcome of this work reveals the importance of machine learning in identifying the intrusion on the network. In future, more data with additional parameters can be collected and tested with complex computational models to improve the detection of malicious activities that could enhance the security and reliability of a network.

REFERENCES

1. Sundaram, A. (1996). An introduction to intrusion detection. Crossroads, 2(4), 3-7.
2. Mukkamala, S., Janoski, G., & Sung, A. (2002, May). Intrusion detection using neural networks and support vector machines. In Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290) (Vol. 2, pp. 1702-1707). IEEE.

3. Denning, D. E. (1987). An intrusion-detection model. IEEE Transactions on software engineering, (2), 222-232.
4. Proctor, P. E. (2000). Practical intrusion detection handbook. Prentice Hall PTR.
5. Northcutt, S., & Novak, J. (2002). Network intrusion detection. Sams Publishing.
6. Salama, M. A., Eid, H. F., Ramadan, R. A., Darwish, A., & Hassanien, A. E. (2011). Hybrid intelligent intrusion detection scheme. *Soft computing in industrial applications* (pp. 293-303). Springer, Berlin, Heidelberg.
7. Horng, S. J., Su, M. Y., Chen, Y. H., Kao, T. W., Chen, R. J., Lai, J. L., & Perkasa, C. D. (2011). A novel intrusion detection system based on hierarchical clustering and support vector machines. Expert systems with Applications, 38(1), 306-313.
8. Revathi, S., & Malathi, A. (2013). A detailed analysis of NSL-KDD dataset using various machine learning techniques for intrusion detection. International Journal of Engineering Research & Technology (IJERT), 2(12), 1848-1853.
9. Zhang, C., Jiang, J., & Kamel, M. (2005). Intrusion detection using hierarchical neural networks. Pattern Recognition Letters, 26(6), 779-791.
10. Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009, July). A detailed analysis of the KDD CUP 99 data set. In 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications (pp. 1-6). IEEE.
11. Peddabachigari, S., Abraham, A., Grosan, C., & Thomas, J. (2007). Modeling intrusion detection system using hybrid intelligent systems. Journal of network and computer applications, 30(1), 114-132.
12. Park, J. S., Shazzad, K. M., & Kim, D. S. (2005, December). Toward modeling lightweight intrusion detection system through correlation-based hybrid feature selection. In International Conference on Information Security and Cryptology (pp. 279-289). Springer, Berlin, Heidelberg.
13. Kayacik, H. G., Zincir-Heywood, A. N., & Heywood, M. I. (2003, July). On the capability of an SOM based intrusion detection system. In Proceedings of the International Joint Conference on Neural Networks, 2003. (Vol. 3, pp. 1808-1813). IEEE.
14. Kim, D. S., Nguyen, H. N., & Park, J. S. (2005, March). Genetic algorithm to improve SVM based network intrusion detection system. In 19th International Conference on Advanced Information Networking and Applications (AINA'05) Volume 1 (AINA papers) (Vol. 2, pp. 155-158). IEEE.
15. Kim, D. S., Nguyen, H. N., Ohn, S. Y., & Park, J. S. (2005, May). Fusions of GA and SVM for anomaly detection in intrusion detection system. In International Symposium on Neural Networks (pp. 415-420). Springer, Berlin, Heidelberg.
16. Sabhnani, M., & Serpen, G. (2004). Why machine learning algorithms fail in misuse detection on KDD intrusion detection data set. Intelligent data analysis, 8(4), 403-415.
17. Kayacik, H. G., Zincir-Heywood, A. N., & Heywood, M. I. (2005, October). Selecting features for intrusion detection: A feature relevance analysis on KDD 99 intrusion detection datasets. In Proceedings of the third annual conference on privacy, security and trust (Vol. 94, pp. 1723-1722).
18. Panda, M., Abraham, A., & Patra, M. R. (2010, August). Discriminative multinomial naive bayes for network intrusion detection. In 2010 Sixth International Conference on Information Assurance and Security (pp. 5-10). IEEE.
19. Panda, M., Abraham, A., & Patra, M. R. (2012). A hybrid intelligent approach for network intrusion detection. Procedia Engineering, 30, 1-9.
20. Kim, D. S., & Park, J. S. (2003, February). Network-based intrusion detection with support vector machines. In International Conference on Information Networking (pp. 747-756). Springer, Berlin, Heidelberg.
21. Kim, J., Kim, J., Thu, H. L. T., & Kim, H. (2016, February). Long short term memory recurrent neural network classifier for intrusion detection. In 2016 International Conference on Platform Technology and Service (PlatCon) (pp. 1-5). IEEE.
22. Al-Yaseen, W. L., Othman, Z. A., & Nazri, M. Z. A. (2017). Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system. Expert Systems with Applications, 67, 296-303.
23. Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. Ieee Access, 5, 21954-21961.
24. Tang, T. A., Mhamdi, L., McLernon, D., Zaidi, S. A. R., & Ghogho, M. (2016, October). Deep learning approach for network intrusion detection in software defined networking. In 2016 International Conference on Wireless Networks and Mobile Communications (WINCOM) (pp. 258-263). IEEE.
25. Zhao, G., Zhang, C., & Zheng, L. (2017, July). Intrusion detection using deep belief network and probabilistic neural network. In 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC) (Vol. 1, pp. 639-642). IEEE.
26. kdd.ics.uci.edu/databases/kddcup99/
27. Dougherty, J., Kohavi, R., & Sahami, M. (1995). Supervised and unsupervised discretization of continuous features. In Machine Learning Proceedings 1995 (pp. 194-202). Morgan Kaufmann.
28. Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. Computers & Electrical Engineering, 40(1), 16-28.
29. Liu, H., Motoda, H., Setiono, R., & Zhao, Z. (2010, May). Feature selection: An ever evolving frontier in data mining. In Feature Selection in Data Mining (pp. 4-13).
30. Yang, X. S. (2010). Nature-inspired metaheuristic algorithms. Luniver press.
31. Binitha, S., & Sathya, S. S. (2012). A survey of bio inspired optimization algorithms. International journal of soft computing and engineering, 2(2), 137-151.
32. Kennedy, J. (2010). Particle swarm optimization. Encyclopedia of machine learning, 760-766.
33. Tang, R., Fong, S., Yang, X. S., & Deb, S. (2012, August). Wolf search algorithm with ephemeral memory. In Seventh International Conference on Digital Information Management (ICDIM 2012) (pp. 165-172). IEEE.
34. Yang, X. S. (2010). A new metaheuristic bat-inspired algorithm. In Nature inspired cooperative strategies for optimization (NICSO 2010) (pp. 65-74). Springer, Berlin, Heidelberg.
35. Bishop, C. M. (2006). Pattern recognition and machine learning. springer.
36. Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160, 3-2
37. Zanero, S., & Savaresi, S. M. (2004, March). Unsupervised learning techniques for an intrusion detection system. In Proceedings of the 2004 ACM symposium on Applied computing (pp. 412-419). ACM.
38. Chapelle, O., Scholkopf, B., & Zien, A. (2009). Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]. IEEE Transactions on Neural Networks, 20(3), 542-542.
39. Sutton, R. S., & Barto, A. G. (1998). Introduction to reinforcement learning (Vol. 2, No. 4). Cambridge: MIT press.
40. Friedman, J. H. (2002). Stochastic gradient boosting. Computational statistics & data analysis, 38(4), 367-378.
41. Karthik, S., Perumal, R. S., & Mouli, P. C. (2018). Breast Cancer Classification Using Deep Neural Networks. In Knowledge Computing and Its Applications (pp. 227-241). Springer, Singapore.
42. Mitchell, T. M. (2006). The discipline of machine learning (Vol. 9). Pittsburgh, PA: Carnegie Mellon University, School of Computer Science, Machine Learning Department.