A Hybrid Approach for Feature Selection Analysis on The Intrusion Detection System Using Navi Bayes And Improved BAT Algorithm

Srinivasa Rao Pokuri[1], Dr. Nagaraju Devarakonda[2]

[1]Research scholar, Department of CSE, Acharya Nagarjuna University, Nagarjuna Nagar, AP, India
[2]Associate Professor, School of Computer science and engineering, VIT-AP University, Amaravati, AP, India
[1]srinivas.pokuri2@gmail.com,[2]dnagaraj_dnr@yahoo.co.in

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

Abstract: In recent days, millions of people in many institutions communicate with each other on the Internet. The past two decades have witnessed unprecedented levels of Internet use by people around the world. Almost alongside these rapid developments in the internet space, an ever-increasing incidence of attacks carried out on the internet has been consistently reported every minute. In such a difficult environment, Anomaly Detection Systems (ADS) play an important role in monitoring and analyzing daily internet activities for security breaches and threats. However, the analytical data routinely generated from computer networks are usually of enormous size and of little use. This creates a major challenge for ADSs, who must examine all the functionality of a certain dataset to identify intrusive patterns. Article collection remains an imperative factor in the modeling of anomaly-based intrusion detection system. Irrelevant characteristics may lead to over fitting, which in turn affects the modeling ability of the classification algorithm. The purpose of this research is to analyze and select the most distinguishing input features to construct an efficient and computationally efficient ADS solution. In the first step, based on the concept of entropy, by selecting the optimal subset, a heuristic algorithm NAIBA is proposed for dimensionality reduction. Then, the relevant and meaningful features are selected, before implementing Number of Classifiers which includes: (1) An irrelevant feature can lead to over fitting which in turn negatively affects the modeling power of the classification algorithms. Experiment was done on CICIDS-2017 dataset by applying (1) Random Forest (RF), (2) Bayes Network (BN), (3) J48 and (4) Random Tree (RT) with results showing better detection precision and faster execution time. The proposed heuristic algorithm outperforms the existing ones as it is more accurate in detection as well as faster. However, Random Forest algorithm emerges as the best classifier for feature selection technique and scores over others by virtue of its accuracy in optimal selection of features

Keywords: Functional selection of intrusion detection systems (IDS), Navie Bayes, Improved BAT classifier algorithm

1. Introduction

Millions of people in various organizations on the African continent communicate with each other over the Internet. In the past two decades, the number of people using the Internet has grown exponentially. Currently, nearly 4 billion users worldwide use the Internet [3]. An intrusion discovery scheme (IDS) displays system circulation towards identify malicious events or violations of privacy, and sends alerts to monitoring stations, or takes preventive measures against detected threats. IDS can be alienated keen on two groupings: one is grounded happening the location where it is installed in the system, or the uncovering method exposed in Figure 1.

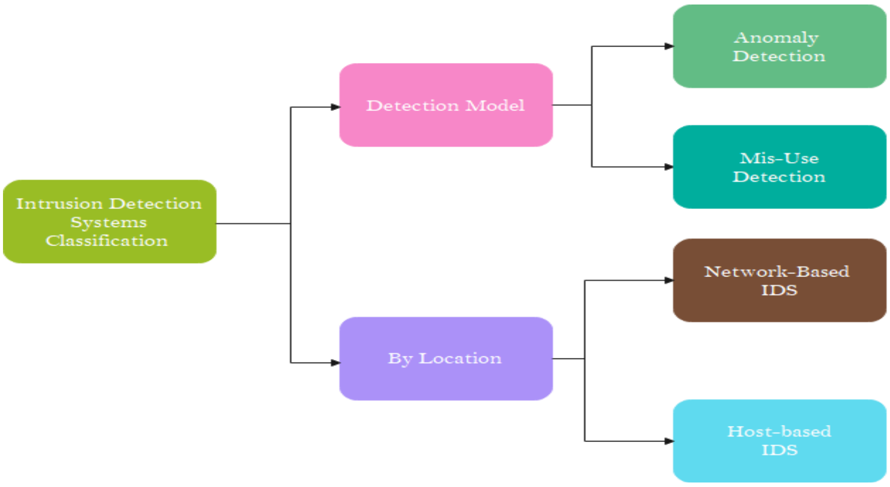


Fig. 1. Classification of Interruption Discovery System (IDS)

Host-based IDS: It is non-client specific and logs files, as well as checking and running processes and connections on the PC. If files are modified, a notification will be provided to the administrator so that he can intervene appropriately. [1].

Network-based IDS: The system monitors and analyses data packets to identify network-based attacks There are two ways to determine if your system has been exposed to an IDS: An IDS can be broken down into two types: abuse detection and violation detection. A known attack signature can be used to identify customer abuse. If it is correct, it will match and block the new connection.This type has a high accuracy rate in detecting known attacks. Anomaly detection identifies outages by following anomalous behavior of network traffic that may specify an attack. Abnormal behavior can be defined as a violation of the edge of the repeated event identifier in the connection, or it can be defined as a violation of the actual configuration file generated by the client for normal behavior. This method can be described as a method based on statistics, data mining and learning [4]. Anomaly-based IDS has the ability to identify known attacks and new attacks [6]. However, the anomaly-based method analyzes the data based on the general properties of the data (such as size, connection time, and number of packets). Therefore, there is no need to view the content of the message. It can also analyze encryption protocols. Due to all these advantages, anomaly detection methods are widely used to detect and prevent network attacks. Anomaly-based IDS has the ability to identify known attacks and new attacks [6]. Therefore, he does not need to view the content of the message. It can also analyze encryption protocols. Due to all these advantages, anomaly detection methods are widely used to detect and prevent network attacks.

Previous works [9] - [13] have focused on the application of feature selection techniques in making more accurate identification of anomalies. Previous researchers have always relied upon Information gain for analysis of significant and relevant characteristics. In this study, a version of CICIDS-2017 dataset having critical features has been applied as it demonstrates highly dense traffic and possesses the capabilities to employ huge number of methods at detecting anomalies. As mentioned in [5], the learning model is affected by application of data having multiple features leading to overfitthat result in decreased performance, more memory and high computation expenses. But wherever there is involvement of complex functionalities with fewer values, information gain tend to be supportive. Here, a new mechanism has been introduced to select ensemble features, before slotting them in categories as per their weight values. Then the five classification algorithms, namely, J48 classifier, Bayes Net (BNC) classifier, Random Tree (RTC) classifier and Random Forest (RFC) classifier are assigned filters by each group of entities for detecting anomalies as well as fending off attacks on the dataset. Most relevant and significant features are extracted into different entity groups that are validated after doing comparison of detection results. With more accuracy in detection results, the perception and choice about the important and relevant the feature groups is made. The weighted features which are used in information gain versus anomaly / attack detection method are used to check the relevant and significant features of the selected entity groups. The better precision results shows the features groups which are more relevant and significant. Such features are applied to various classifiers like J48 classifier, Bayes Net (BNC) classifier, Random Tree (RTC) classifier and Random Forest (RFC) classifier on the given data set. Finally, the results are validated for relevant and significant features. The ones with better accuracy in detection results tend to be looked up as more meaningful and relevant the feature groups.

2. Related Works

Recently, most applications depended on the network or computer system and their behavior is to be analyzed and threaten by the known technique called Intrusion detection. Moreover, such technique also interrupts the features of the network or computer system which includes integrity accessibility, and confidentiality of concerned data [5]. The study the characteristics related to the network traffic and also identified number of mechanisms to handle introduction mostly they were filtered, wrapper, and combination of both algorithms [8].However, feature extraction with ensemble of fitter and wrapper assign weight for every feature and maximum ranked features applied to clustering approach [15]. In some work, most popular resampled method named artificial underground oversampling method [14] is applied to remove class imbalance problem. Later combined two techniques one is the Selection of Ensemble Characteristics (EFS) and the Principal Component Analysis (PCA) and then practical to the AdaBoost-based IDS to improve the performance of classification. One of the most popular wrapper methods used by the most of researchers known as information gain (IG) used as a feature selection mechanism and is worked to find the minimum ranking score for each feature as a result set. Next, the ranking weights are used to determine optimalfeatures and are to be considered as final class label. Number of researchers use weight score >0.4, > 0.001 and > 0.8 respectively [16] [14].

3. Feature Selection

The mechanism used to extract important and relevant information is known as feature selection. Generally, such kind of technique is used to discriminate the class label into relevant and irrelevant functionality .The relevant functionalities had information which is optimal to class and whereas in non-informative functionalities the class gained very little information about class [1]. The main objective of feature selection is to filter non-informative features and identify informative features and to pass maximum information related to class output. To achieve this, number of feature selection method are available but generally which is classified into filter, wrapper and combined or ensemble approaches [17][19]. The Filtering method is one used to access and extract relevant features from the given data using statistical approach. However, in case of the wrapper method selection of the relevant subset of features can be done by using the classification criteria. But the wrapper method is computationally very expensive. The next, method is ensemble or integrated method used to apply feature selection with learning criteria to extract optimal features to the given data. Such kind of ensemble feature selection methods are less expensive compare to the wrapper method.

3.1. Naive Bayes (NB)

Naive Bayes classifier works as surveys: let X be a vector of random variables denoting the observed attribute valuesin the training set $X=[x_1,x_2,...,x_n]$ to certain classlabel c in the training set. The probability of both class assumed the vector of experiential values for the prognostic attributes canbe computed using the subsequent formula:

$$P(Y_j\,|\,X)=\frac{P(Y_j)P(X\,|\,Y_j)}{\sum_{i=1}^{c}P(Y_i)P(X\,|\,Y_i)},\qquad j=1,...,c$$

where $P(Y_i)$ is the prior probability of class Y_i and $P(Y_j\,|\,X)$ is the class conditional probability density functions. Basically put, the conditional independence assumption assumesthat each variable in the dataset is conditionally independentof the other. This is simple to compute for test cases and toestimate from training data as follows:

$$P(X\,|\,Y_j)=\prod_{i=1}^{n}P(X_i\,|\,Y_j),\qquad j=1,...,c$$

where X_i is the value of the i^{th} attribute in X and n is the number of attributes. Let k be the number of classes, and c_i is the i^{th} class; the probability distribution over the set of features is calculated using the following equation:

$$P(x)=\prod_{i=1}^{k}p(c_i)p(X\,|\,c_i)$$

Thus, NB implementation is domain-independent and parameter-free. There are only two disadvantages of NBs: they assume the features are independent, and then they can't be. Despite these drawbacks, NB is a well-regarded benchmark because of this You can find an excellent survey of NB creative adaptations in the literature.

3.2. Bat Algorithm (BA)

A swarm-based method, created by Yang in 2010, was based on echolocation (or echo-location) behaviour of bats nocturnal animals that possess echolocation, like dolphins, whales, and many other creatures use it to communicate, such as shrews do. There are generally uncreativerated signals above the limits of human perception (about 20kHz). Bats use echolocate to distinguish moving in the darkness, like radar, to exchange information among themselves, as well as detect the location of their food. He's gone from describing bats' echolocation behaviour to formulating an algorithm which adheres to some of it.The bats use echolocation to sense distance, and theyalso know the difference between prey and backgroundbarriers in some magical way.

- A Creative position: Bats are generally fly at random with frequency at wavelength of intensity to seek prey by using the proximity sensor, they can increase or decrease the length of the pulses to a pre-determined rate as well as increase or decrease the wavelength as they move closer to their target. Although the loudness can vary in many ways, it is assumed that the loudness varies from a large (positive) A_0 to a minimum constant value A_{min} .
- Generally, frequency f in the range of $[f_{min},f_{max}]$ correspond to the wavelength in the range of $[min, max]$. For instance, a frequency range of $[20kHz, 500kHz]$ corresponds to a range of wavelengths from

0.7mm to 17mmin the air. Simply, f is assumed to be in the range of $[0, f_{max}]$ High frequencies are known to have short wavelength and short-distance travels. The typical range for bats is a few meters. Signal rate is inthe range of [0,1]; 0 means no signal while 1 means maximum signal emission rate. The positions x_i and velocities v_i of bats and new solutions x_i^t and velocities v_i^t in time t in a d-dimensional search space are shown as follows.

$$f_i = f_{min} + (f_{max} - f_{min})\beta$$

$$v_i^t = v_i^{t-1} + (x_i^t - x_*)f_i$$

$$x_i^t = x_i^{t-1} + v_i^t$$

$\beta \in [0,1]$ is a random vector indomitable by usual distribution. x_* is the best present explanation for the minute with the comparison of all explanations between the bats. Velocity increment $f_i\lambda_i$ can be used to adjust the velocity variations while other factors f_i is fixed, depending on the nature of the problem. At the beginning, each bat is assigned a random frequency value in the range of $[f_{min}, f_{max}]$ .Aimed at the local exploration part; a answer is designated among the current best answers, a new solution for each bat isgenerated locally using random walk.

$$x_{new} = x_{old} + \varepsilon A^t$$

where $\varepsilon \in [-1,1]$ is a random number, while $A^t =< A_i^t >$ is the average loudness of all the bats at this time step.

4. Proposed Method

Machine Learning (ML) based methods are become popular now and are used in this study to improve performance of the Anomaly Detection System (ADS) and also worked for solution to prevent attack from the providers. Ensemble optimization ML based feature selection method applied first and extracted optimal features and then set of classifiers used to detect the attack type. The approach is used a10-fold cross-validation (CV) during the experiment and to validate the model performance. Finally, model is to classify attack especially benign traffic attack. The proposed method framework shown in Figure 2, and overall work is divided into major four parts and are given below:

1. **Preprocessing:** The step in which original or raw data is to be converted into desired formats which are helps for further analysis.
2. **Feature Selection :**The second step, applied proposed the NA-IBA based feature selection approach used to retrieve the subset of date sets and retrieved most relevant or suitable features related to each type of the attack class.
3. **Classification:** The last step of the proposed work is deal classification which is helps to improve overall performance of theIDS. The number of classifiers used in this work which includes : (i) Random Forest(RF) (ii) Random Tree (iii) naïve Bayes (iv) Bayesian Network and (v) J48.

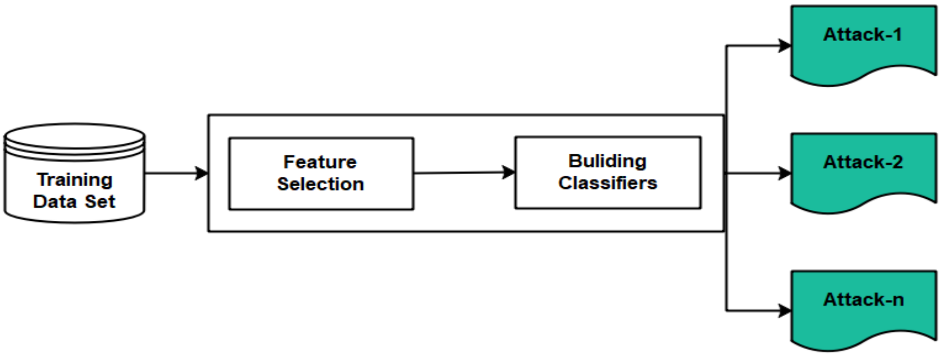


Fig 2: Proposed method framework for the classification

4.1. Improved BAT guided by Naive Bayes Classifier (NB-IBA)

The complete algorithm for the proposed Improved BAT guided by Naive Bayes classifier (NBIBA) is shown below.

Algorithm: Feature Selection using Proposed NB-IBA Method

```
Initialize parameters: $A, A_{min}, A_{max}, r, f_{min}, f_{max}, P_{max}, I_{max}, V_{max}, V_{min}, \Phi, \delta, \gamma, \alpha$
Generate a swarm with $P_{max}$ bats
Calculate cost function for all bats
Find the current best bat ( $x_*$ )
While stop condition not met Do
  for $i = 1$ to $P_{max}$ Do
Frequency $f_i = f_{min} + (f_{max} - f_{min})\beta$
Velocity $v_i^t = v_i^{t-1} + (x_i^t - x_i)f_i$
If $(V_i > V_{max})$ then
$(V_i = V_{max})$
End-If
If $(V_i < V_{min})$ then
$(V_i = V_{min})$
End-If
Locations $x_i^t = x_i^{t-1} + v_i^t$
If $(Rand > r_i)$Then
calculate $\varepsilon A^t$
If $(\varepsilon A^t > A_{max})$Then
$\varepsilon A^t = A_{max}$
End-If
Generate a local solution around the best solution $(x_*)[x_{gb} = x_{old} + \varepsilon A^t]$
End-If
Calculate $\varepsilon A^t$
If $(\varepsilon A^t < A_{min})$Then
$\varepsilon A^t = A_{min}$
End-If
Generate a new solution around the currentSolution $(x_i^t)[x_i = x_{old} + \varepsilon A^t]$
if $(x_l \geq x_{gb})$
$f_x = x_l$
      else
$f_x = x_{gb}$
if $(Rand < A_i)\,\&\,(f(f_x) < f(x_*))$
      Accept New Solutions
Increase $r_i$ $r_i^{i+1} = r_i^0[1 - \exp(-\gamma t)]$
Decrease $A_i[A_i^{t+1} = \alpha A_i^t]$
End-If
End-for
  Find the Best Solution ( $x_*$ )
End-While
```

4.2. Classification Algorithms

Although several previous works have supported many diverse algorithms, in this work, number of classifiers used which includes:(i) Random Forest (RF) (ii) Random Trees (RT) (iii) Bayesian Network (iv) J48.

4.2.1. Bayes Network (BN)

The model in which among variables there exist encoding probabilistic relation which is called the Bayesian Network (BN). On the general assumption of the behavior of the target system model, the precision of the method is determined, with any notable departure from it is likely to reduce precision in detection. Bayesian networks have been applied in a few anomaly detectionsstudies[22][25].

4.2.2. Random Forest (RF)

Random Forest, one of the classification methods, a classifier in a collection of number of decision tree. Next the word, Forest represented as a collection of classifiers. The decision tree is different from one to other depends on random selection of the desired attributes corresponds to each node. Number of works has been done related to anomaly detection using random forest [22][24].

4.2.3. Random Tree (RT)

The decision tree which is a collection of random attributes called Random Tree and complete tree is built with the combination of two elements nodes and branches. However, node to be considered as a test attribute and branch to be the results. Decision sheets depict the final decision reached following making calculation of all attributes as class labels. This method has been included in certain anomaly detection studies [28] [30].

4.2.4. J48

A machine learning algorithm corresponds to family of decision tree i.e., J48 or C4.5, make use of training data to a decision tree usingentropy [23]. Unlike IDE3, this method used to create a decision tree keeping the ability togeneratesequence of attributes. The J48 algorithm applied to anomaly detection included in many research work[29].

4. Experimental setup

5.1. CICIDS2017 dataset

The dataset [5], is introduced in 2018 at the Canadian Institute for Cybersecurity and is used to detect DDoS attacks. However, data set is present benign and attack processconsidering real world network traffic data. Also, data set includes 79 features which is comprise of class labels and are used to specify major attacks mentioned: (i) Brute Force SSH (ii) Brute Force FTP (iii) Infiltration (iv) Heartbleed (v) Web Attack (vi) DoS (vii) Botnet and (viii) DDoS and the complete attacks information shown in Table 3. Total 225,746 records related to DDoS and Benign attacks included in CICIDS2017 and each record comprised with total 80 features like (i) protocol (ii) stream ID (iii) source IP (iv) destination IP (v) source port, and etc. The complete records and features is included in Table 1.

Table 1. The records in data set CICIDS2017

Source IP	Source port	Destination port		Duration of flow	Total number of Fwd packages	Total back packets
192.xxx.xx.20	41938	334	...	143346	46	70
192.xxx.xx.20	42978	80		40907	1	1
192.xxx.xx.20	41955	445		143896	47	69
192.xxx.xx.21	12887	54		314	2	2
192.xxx.xx.20	41946	444		142609	44	59
192.xxx.xx.21	33065	55		255	2	2
192.xxx.xx.20	41942	443		142488	47	57
192.xxx.xx.20	41939	444		23838	28	32

5.2. Experimental setup

As an initial model fitting, the complete original data is split into two subsets one is training data (80%) and other is test data (20%). Next, applied proposed IG-BA feature selection method and extracted optimal set of feature set. The algorithm which helps to avoid irrelevant features from the data set and also improved the performance of classification.

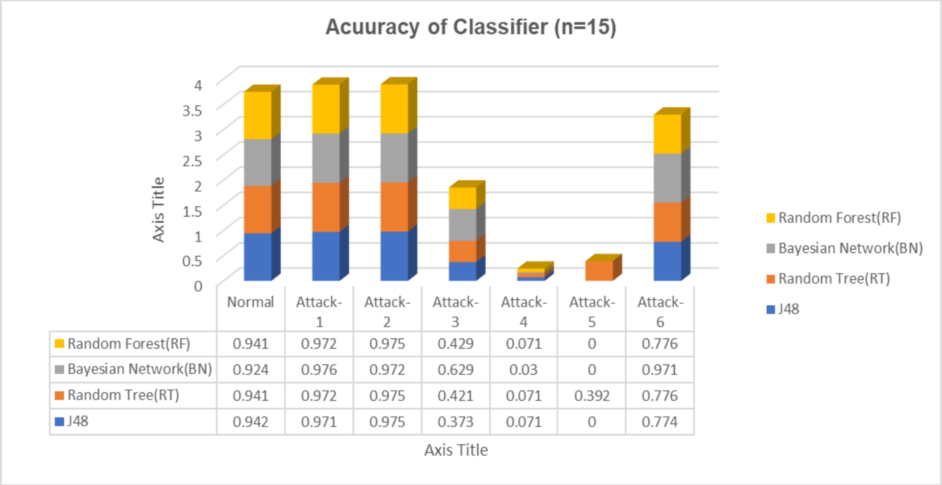
Table 2: Training and testing of the CICIDS2017 Dataset

Attack class	**No. Records**	**Train set (80%)**	**Test set (20%)**
Benign	61562	49250	12312
Bot	1966	1573	393
Brute force	1507	1206	301
DoS / DDoS	58134	46507	11627
Golden Eye back	10293	8234	2059
Back Hulk	10486	8389	2097
Slowhttptest back	5499	4399	1100
Slowloris back	5796	4637	1159
FTP-Patator	7938	6350	1588
Heartbleed	11	9	2
Infiltration	36	29	7
PortScan	60294	48235	12059
SQL	21	17	4
SSH-Patator	5897	4718	1179
XSS	652	522	130
Total	**230092**	**184074**	**46018**

After performing the feature selection using hybrid proposed method the result subset is applied to different classifiers which are (i) Random Forest(RF) (ii) Random Tree (iii) naïve Bayes (iv) Bayesian Network (v) J48.

Table 3. Attacks worked on this job

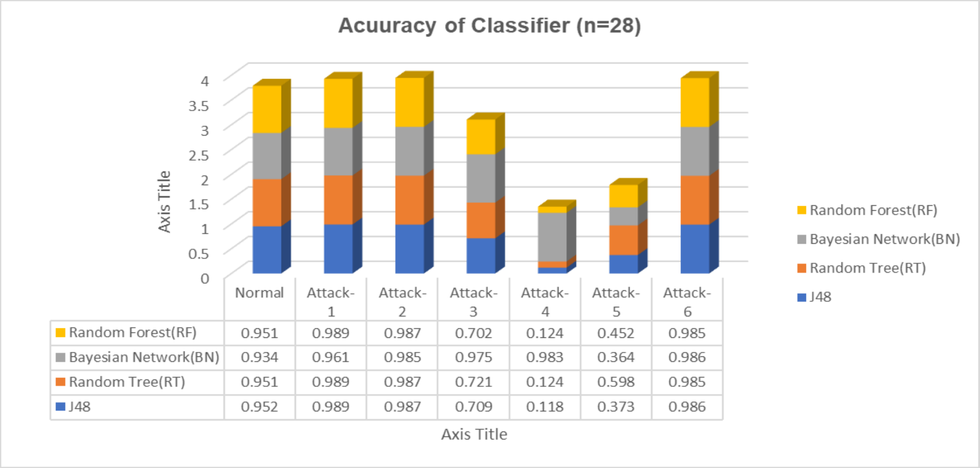
Attack number	**Attack name**
Attack-1	DoS / DDoS attack
Attack-2	Port scan attack
Attack-3	Bot attack
Attack-4	Web attack
Attack-5	Infiltration
Attack-6	Brute force



	Normal	Attack-1	Attack-2	Attack-3	Attack-4	Attack-5	Attack-6
Random Forest(RF)	0.941	0.972	0.975	0.429	0.071	0	0.776
Bayesian Network(BN)	0.924	0.976	0.972	0.629	0.03	0	0.971
Random Tree(RT)	0.941	0.972	0.975	0.421	0.071	0.392	0.776
J48	0.942	0.971	0.975	0.373	0.071	0	0.774

Fig.3.Performance of classification algorithms considering feature set of size 15

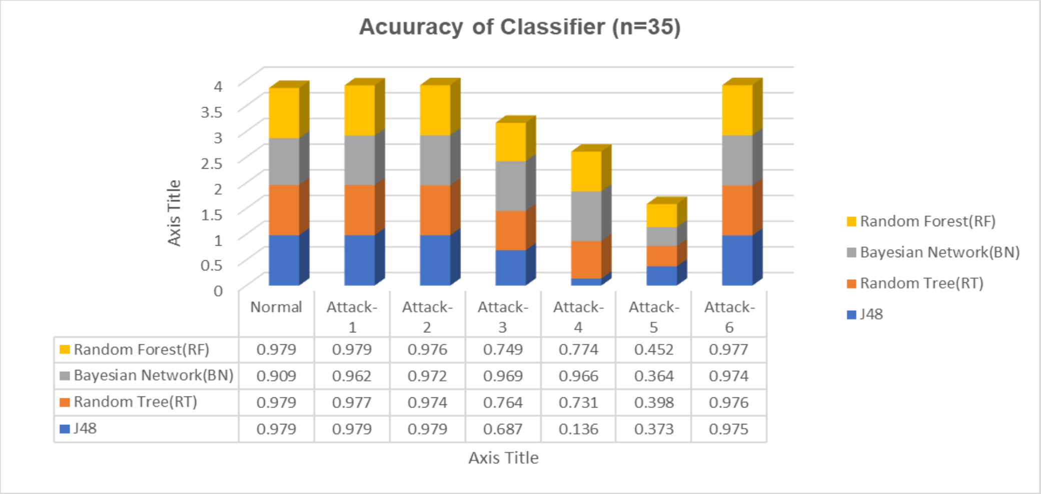
The performance of classification algorithms by applying feature set of size15 is shown in fig 3. Random Forest (RF) produced almost 97% accuracy when compared other classification methods. The experimental results with the given classification algorithms RandomForest (RF), Random Tree (RT), and J48 are promising while detecting at Normal, Attack1 to 3. However, classification algorithms results difficulties in detecting Attack 3 and Attack 5 traffic.



	Normal	Attack-1	Attack-2	Attack-3	Attack-4	Attack-5	Attack-6
Random Forest(RF)	0.951	0.989	0.987	0.702	0.124	0.452	0.985
Bayesian Network(BN)	0.934	0.961	0.985	0.975	0.983	0.364	0.986
Random Tree(RT)	0.951	0.989	0.987	0.721	0.124	0.598	0.985
J48	0.952	0.989	0.987	0.709	0.118	0.373	0.986

Fig.4.Performance of classification algorithms considering feature set of size 28

The performance of classification algorithms by applying feature set of size 28 is shown in Table 7. Random Forest (RF) produced almost 97% accuracy, recall i.e., 0.989. However, this classification algorithms results difficulties in detecting Attack 5 traffic.The experimental results with the given classification algorithms Random Forest (RF), Random Tree (RT), and J48 are promising while detecting at Attack1 to 3 and produced better FRP. Finally, it is observed that Naïve Bayes(NB) produce low FRP.



	Normal	Attack-1	Attack-2	Attack-3	Attack-4	Attack-5	Attack-6
Random Forest(RF)	0.979	0.979	0.976	0.749	0.774	0.452	0.977
Bayesian Network(BN)	0.909	0.962	0.972	0.969	0.966	0.364	0.974
Random Tree(RT)	0.979	0.977	0.974	0.764	0.731	0.398	0.976
J48	0.979	0.979	0.979	0.687	0.136	0.373	0.975

Fig.5.Performance of classification algorithms considering feature set of size 35

Similarly, while considering 35 features Random Forest (RF) produced accuracy of 97.9%compared to other classification algorithms

5. Conclusions

The proposed method validates that feature selection improves the performance of feature selection on anomaly detection data. The proposed feature selection produces the ranking of features based on their weight values using IG algorithm, resulting in a subset of features to rank. Later, individual subset applied to BA algorithms and then processed which results optimal features for the further classification. From the overall Random Forest performs promising using all sizes of feature sets from 15, 28,and 35, Also noticed that J48 results better in case of feature sets of 28 and 35. All the traffics detects properly using feature subsets of 28, and 35. However, the Bayes Naïve (BN) results low accuracy compared other classifiers. Also notice in this classification subset of features impact on reduction of FPR. In the future, work plan to conduct study on multi classification.

References

1. K. Scarfone, P. Mell, Special Publication 800-94: Guide to Intrusion Detection and Prevention Systems (IDPS), National Institute of Standards and Technology (NIST) (2007).
2. Bace, Rebecca Gurley: Intrusion detection. Copyright 2000 by Macmillan Technical Publishing, ISBN 1-57870-185-6.
3. J. Zhang, H. Li, Q. Gao, H. Wang and Y. Luo, “Detecting Anomalies from Large Network Traffic Data Using an Adaptive Detection Approach,” Inf. Sci., Vol. 318, pp. 91-110, October 2015.
4. García-Teodoro, P .; Díaz-Verdejo, J .; Maciá-Fernández, G .; Vázquez, E. Anomaly-based network intrusion detection: techniques, systems, and challenges. Comput. Secur. 2009, 28, 18–28. [CrossRef]
5. R. Panigrahi and S. Borah, `` A Detailed Analysis of the CICIDS2017 Dataset for the Design of Intrusion Detection Systems '', Int. J. Eng. Technol., Vol. 7, no. 24, pp. 479–482, 2018
6. Alazab, A .; Hobbs, M .; Abawajy, J .; Alazab, M. Using the feature selection for the intrusion detection system. In Proceedings of the 2012 International Symposium on Communications and Information Technologies (ISCIT), Gold Cost, Australia, October 2-5, 2012; 296-301.
7. MK Kundu, DP Mohapatra, A. Konar, and A. Chakraborty, `` Decision Tree Techniques Applied to NSL-KDD Data and Its Comparison with Various Feature Selection Techniques, '' Smart Innov. Syst. Technol., Vol. 27, no. 1, pp. 205-211, 2014.
8. W. Wang, Y. He, J. Liu, and S. Gombault, “Building Important Features from Massive Network Traffic for Light Intrusion Detection,” IET Inf. Secur., Vol. 9, no. 6, pp. 374–379, November 2015
9. I. Ahmad, M. Hussain, A. Alghamdi and A. Alelaiwi, “Improving SVM Performance in Intrusion Detection Using the Selection of Optimal Feature Subsets Based on Key Genetic Components”, Neural Comput. Appl., Vol. 24, nos. 7–8, pp. 1671-1682, June 2014.
10. S.-H. Kang and KJ Kim, “A feature selection approach to find optimal feature subsets for the network intrusion detection system”, Cluster Comput., Vol. 19, no. 1, pp. 325–333, March 2016.
11. AI Madbouly, SA King Abdulaziz University Jeddah, AM Gody and TM Barakat, `` Relevant feature selection model using data mining for intrusion detection system '', Int. J. Eng. Trends Technol., Vol. 9, no. 10, p. 501-512, March 2014.
12. E. Popoola and A. Adewumi, “Efficient feature selection technique for a network intrusion detection system using discrete differential evolution and a decision tree”, Int. J. Netw. Secur., Vol. 19, no. 5, pp. 660–669, 2017.
13. BA Tama and KH Rhee, “A Combination of PSO-Based Feature Selection and Tree-Based Classifier Set for Intrusion Detection Systems”, Adv. Comput. Sci. Ubiquitous Comput., Vol. 373, pp. 489–495, February 2015.
14. A. Yulianto, P. Sukarno and N. Suwastika, `` Improving the performance of the AdaBoost-based intrusion detection system (IDS) on the CIC IDS 2017 dataset '', J. Phys., Conf. Ser., Vol. 1192, March 2019, art. no. 012018, doi: 10.1088 / 1742-6596 / 1192/1/012018.
15. S. Bhattacharya and S. Selvakumar, “Multi-measure multi-weight ranking approach for the identification of network characteristics for the detection of DoS and probe attacks”, Comput. J., vol. 59, no. 6, p. 923–943, June 2016.
16. TA Alhaj, MM Siraj, A. Zainal, HT Elshoush and F. Elhaj, “Feature Selection Using Information Gain for Better Structure-Based Alert Correlation,” PLoS ONE, vol. 11, no. 11, 2016, art. no. e0166017.
17. Guyon, I., Elisseeff, A. (2003). An introduction to selecting variables and features. Journal of machine learning research 3 (March) 1157-1182.
18. Saeys, Y., Inza, I., Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. bioinformatics 23 (19) 2507-2517.
19. Yang, XS. (2010). A new algorithm inspired by metaheuristic bats. In: Nature Inspired Cooperative Strategies for Optimization (NICSO 2010) (Eds.Cruz C., Gonzalez J., Krasnogor N. and Terraza G.), Springer, SCI 284, pp 65-74.
20. Yang, XS. (2008). Metaheuristic algorithms inspired by nature. Luniver Press.
21. Xin-She Yang and Amir H. Gandomi. 2012. Bat Algorithm: A New Approach for Global Engineering Optimization. Engineering Computations, 29 (5), 464-483.
22. M. Reazul, A. Rahman and T. Samad, “A Network Intrusion Detection Framework Based on a Bayesian Network Using a Wrapper Approach”, Int. J. Comput. Appl., Vol. 166, no. 4, p. 13-17, May 2017.
23. J. Jiang, Q. Wang, Z. Shi, B. Lv and B. Qi, `` RST-RF: A hybrid model based on approximate set theory and a random forest for network intrusion detection '' , in Proc. ACM Int. Conf. Process., 2018, p. 77-81.
24. RK Singh, S. Dalal, VK Chauhan and D. Kumar, “Optimizing FAR in an Intrusion Detection System Using a Random Forest Algorithm”, SSRN Electron. J., vol. 5, pp. 3-6, March 2019.
25. N. Ding, H. Gao, H. Bu and H. Ma, `` RADM: Real-time Anomaly Detection in Multivariate Time Series Based on a Bayesian Network '', in Proc. IEEE Int. Conf. Smart Internet Things, August 2018, p. 129-134.

26. K. Goeschel, `` Reducing False Positives in Intrusion Detection Systems Using Data Mining Techniques Using Supporting Vector Machines, Decision Trees, and Naive Bayes for Offline Analysis '', in Proc. SoutheastCon, March 2016, p. 1–6.
27. S. Shakya and S. Sigdel, `` An Approach to Developing a Hybrid Algorithm Based on a Support Vector Machine and Naive Bayes for Anomaly Detection '', in Proc. Int. Conf. Comput. Common. Autom. (ICCCA), January 2017, pp. 323–327.
28. R. Chitrakar and H. Chuanhe, `` Anomaly Detection Using Classification of Support Vector Machines with Clustering k-Medoids '', in Proc. 3rd Asian Himalayas Int. Conf. Internet, November 2012, p. 1–5, doi: 10.1109 / AHICI. 2012.6408446.
29. AP Muniyandi, R. Rajeswari and R. Rajaram, “Network Anomaly Detection by Cascading K-means Clustering and C4.5 Decision Tree Algorithm”, Procedia Eng., Vol. 30, pp. 174-182, February 2012.
30. [41] S. Thaseen, intrusion detection model using the fusion of PCA and optimized SVM. Boca Raton, FL, USA: CRC Press, 2014, pp. 879–884.