Network Intrusion Detection Using Distributed Genetic Random Forest Method

[1]Dr. N. Prakash,

Assistant Professor, Department of Computer Science, Gobi Arts & Science College, Gobichettipalayam. Second

[2]Mr. N. Prabhu,

Assistant Professor, Department of Computer Science, Gobi Arts & Science College, Gobichettipalayam.

Article Info
Page Number: 8843 - 8866
Publication Issue:
Vol 71 No. 4 (2022)

Abstract

The increasing use of internet networks has led to an increase in threats and new attacks every day. To detect an anomaly or misused detection, an Intrusion Detection System (IDS) has been proposed as an important component of a secure network. Because of the model-free properties that enable them to identify the network pattern and discover whether they are normal or malicious, the Machine-learning technique has been useful in the area of intrusion detection. Different types of machine learning models have been exploited in anomaly-based IDS. There is a growing demand for reliable and real-life attack data sets within the research community.To improve the detection rate of NIDSs and reduce the false alarm rate, a lot of works apply a variety of methods of machine learning on NIDS. The Proposed Distributed Genetic Random Forest Method (DGRFM) to build intrusion detection systems exhibits a superior performance than the traditional genetic methods. The proposed multiple-level hybrid classifier to build NIDS to enhance RF with unlabeled data and trained to detect previously "unseen" attacks. The huge amount of network data with the unbalanced distribution of normal and anomaly behaviors lead to the problems of low detection rate and high false alarm rate in most NIDS.

Article History
Article Received*: 25 March 2022*
Revised*: 30 April 2022*
Accepted*: 15 June 2022*
Publication*: 19 August 2022*

Keyword: Attacks, IDS, network, machine learning, false rate, genetic algorithm, random forest, and unbalanced data.

1. Introduction

Computer networks and their related applications have become a gorgeous source within the era of the data society. Similarly, the potential threat to the global information infrastructure has also increased greatly [1]. To guard against several cyber-attacks and computer viruses, numerous computer security approaches have been extensively researched in recent years. The major security techniques proposed are cryptography, firewalls, anomaly, intrusion detection, etc. Among the available standard techniques, intrusion detection techniques have been considered to be one of the most significant and competent techniques for protecting complex and dynamic intrusion attacks [2, 3].

An intrusion is a violation of the security policy of the system and thus intrusion detection mainly refers to the methods that detect violations of system security policy [4]. IDS is a program that analyses what happens or has happened during execution and tries to find out the indications that the computer has been misused. The cruelty of attacks in the network has increased radically then the Intrusion Detection System (IDS) has become an essential factor to the security infrastructure of several companies [5, 6]. It facilitates companies to defend their systems from various attacks that come with rising network connectivity and dependence on information systems. Several information security techniques are available today to guard information systems against unauthorized use, duplication, alteration, destruction, and virus attacks [7].

Interruption Detection is one in every one of the critical pieces of the organization the executives. IDS analyses a host or organization to recognize the expected interruptions or assaults. It is frequently either network-based or host-based, while network essentially based interruption recognition frameworks are normal [8-14]. Organization-based interruption identification frameworks inspect every one of the parcels coursings through the organization for indications of assaults. Host-based frameworks investigate client and technique movement on the nearby machine for the indications of interruptions [15-16].

Random Forest (RF) is one of the most powerful supervised methods is used for classification problems in machine learning. The algorithm is carried out in two different stages, the first one for the creation of the forest of the given dataset, and the second one deals with the prediction obtained in the first step. Considering nodes from a single location RF establishes forests and it does not allow nodes from multiple locations[17-21]. To overcome the problem Genetic Algorithm (GA) has been used for intrusion detection problems for different kinds of scenarios to derive the

classification rules. GAs is used to select required features and to determine the optimal and minimal parameters of the given functions with different machine learning methods used to derive rules for acquisition [22-23]. This approach includes both quantitative and categorical features of network data for deriving classification rules. However, the inclusion of quantitative features can increase the detection rate but no experimental results are available [24, 25].

Distributed genetic algorithm is a parallel genetic algorithm that has its independent running's on separate machines. It does not matter what one genetic rule is, as a result of distributed genetic algorithms having multiple machines running's severally to unravel constant task.Before using the proposed method, the dataset is pre-processed by resampling the original training set. The modified training sets are formed by classifiers are constructed using the training sets and then combined. The projected hybrid methodology selects the most effective from people of secondary machines are connected to the running computation [26, 27].

2. Distributed Genetic Random Forest Method

The DGRFM proposed hybrid data optimization data consists of sampling and feature selection. Data sampling is to delete outliers in the dataset to reduce the negative impact of unbalanced data distribution on intrusion detection. Feature selection is to reflect the best difference between anomalous behaviours and normal behaviours and to delete the useless features to enhance the detection performance of NIDS.

An effective NIDS based on data sampling is built by using the proposed method. By sampling method to extract representative training data and combining it with random forest method, the performance of NIDS can be improved effectively. Sampling technique to select the representative dataset to identify anomalous network data, proving that data sampling can improve the accuracy and speed of intrusion detection. The combination of the modified genetic method with machine learning methods is to build intrusion detection models that can discover similar structures between datasets with higher quality. The block diagram of the proposed method is shown in Figure 1.

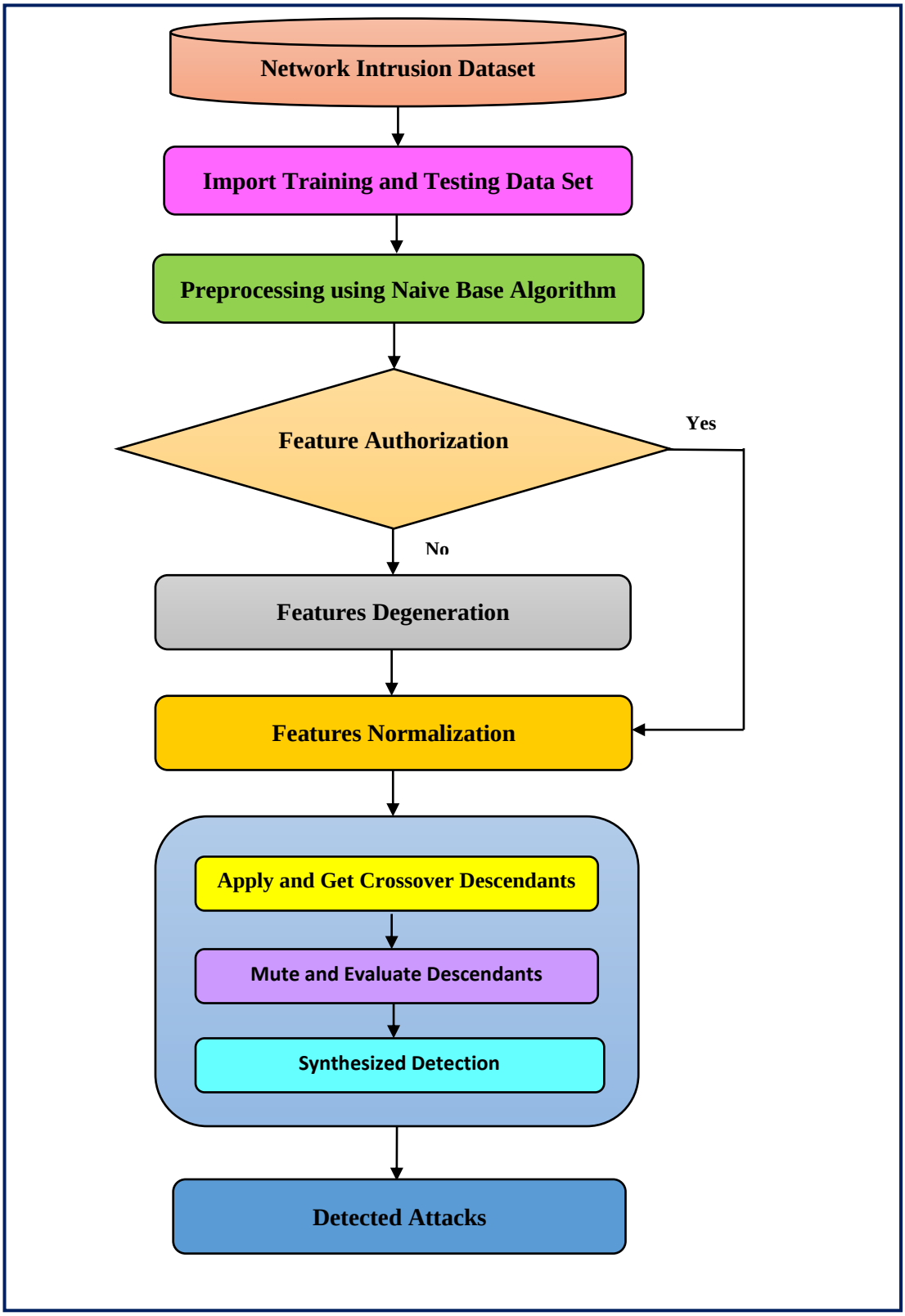

Figure 1. Block Diagram of Proposed DGRFM for Network Intrusion Detection

Integrating DGM with Random Forest to construct the classifier of intrusion detection can greatly reduce the processing time of the intrusion detection system.The hybrid optimal locations are predicted by using the selective process is addressed by using the rules mentioned in Equations 1 to 3 as,

$$Min = P(x_i) - Max \text{ ----------(1)}$$

$$inc = 2.0 \times \frac{Max - P(x_i)}{N_i} \text{ ----------(2)}$$

$$low = inc/2.0 \text{ ----------(3)}$$

Where *Min* is the lower bound value, *inc* difference between the fitness of the adjacent individuals and *low* is the expected number of trails selected. The tree-based outlier detection with linear time complexity and high precision is suitable for the high-dimensional and large amount of data.

The test dataset is divided into two parts: validation and test sets. The reason is that after every training, the genetic algorithm needs to practice over data to choose the best trees in each chromosome and evolve to reach the optimum solution. If missing values are present for predictors, they can be imputed by proximity measures of an RF, the procedure is as follows. Firstly, an RF is built with missing values that are replaced with the sample median of all the observations for considered random variables. In the next steps, the sample median is updated with a new imputed value by using the proximity measures. In the first RF, it is evaluated to which terminal nodes the missing values replaced by sample median were ended up and their proximity weights concerning other observations in that node were computed for missing values. Having the proximity weights the second RF is produced and the imputed values of these missing are obtained by the weighted average of those observations in the node. The fitness of the individual of the given distribution data sampling is obtained by the fitness function calculated directly as Equation 4,

$$f(x_i) = Min \ + \ 2(Max - 1)\frac{x_i - 1}{N_i - 1} \text{ ----------(4)}$$

Where xi is the position in the ordered population of individual i, selection of individuals for reproduction based on the fitness of individuals relative to one another. Because anomalies are less and different also more vulnerable to be isolated. In a data-orientated random tree, information is recursively reduced till all information is isolated. The proposed random partition makes outlier

records as a shorter path length because records with distinguishable attribute values are more likely to be separated in early partitions given by Equation 5.

$$x_{[t+1]} = x_{[t]} - A \times D \text{----------(5)}$$

Where and are the next and current position of the node isolation point, D be the difference vector depends on the separated attribute's location and is estimated in Equation 6.

$$D = Cx_p[t] - x[t] \text{----------(6)}$$

Also, A and C is the coefficient vector and calculated as where are the vectors randomly initialize with values between 0 and 1. Hence the implementation steps are as follows:

- Randomly select a fixed number of sample points from training data as sub-samples and put them in the root node of the tree.
- Specify an attribute to randomly generate a cutting point p in the current node data is generated between the maximum and minimum value of the specified attribute in the current node data.
- A hyperplane is generated and the data space of the current node is divided into two subspaces.
- The data less than p in the specified attribute is put into the left child of the current node.
- The data greater than or equal to p is put into the right child of the current node.
- Recursively execute steps 2 and 3, until the child node has only one record or the Tree has reached the defined height.

Comparing with the conventional clustering method such as K-means and C-Means, the ECM can obtain the number of clustering once the features of the dataset are determined. The design of classification models requires some crucial parameters for determining the structure. Locating the favorable areas of the solution space to initialize the search-guided techniques are generally aimed at achieving near-optimal solutions with the use of limited computational resources, rather than trying to find the globally most beneficial Pareto the front of a complicated problem. A subspace of functions is then randomly decided on at every node to develop branches of the decision tree. The trees are then mixed as an ensemble right into a forest. To optimize trees of a

random forest by two strategies. One straightforward strategy is to enhance the classification performance of individual trees by a feature weighting method for subspace sampling.

In this method, feature weights are computed for the correlations of features to the class feature and regarded as the probabilities of the feature to be selected in subspaces. The procedure makes the apparent need for the approaches that are capable of finding the nearest optimal solutions within the constraints of available computational resources and also aid in the effective application. To improve the accuracy of the algorithm for the high dimensional data, present a new dimension reduction method to reduce the number of dimensions according to the importance of the feature variables the dimensions of the dataset are reduced from M to m. The process of dimension-reduction is presented in Figure 2. DGRFM in the practical domain to explore the impact that integrating knowledge of performance has been applied to single objective RFM problems with exploring the impact on a multi-objective DGRFM intrusion detection system. To reduce the dimensions of the training dataset, calculate the importance of each feature variable according to the gain ratio of the variable.

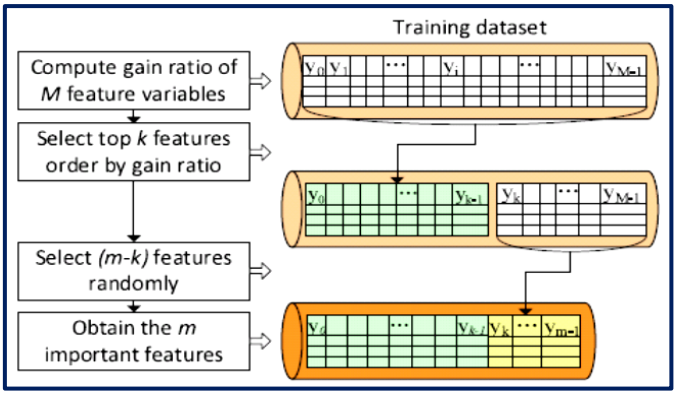

Figure 2. Dimension Reduction in Training Process

The accuracy of the RF algorithm was improved by applying the genetic algorithm on hybrid RF. By conducting an experiment and comparing its learning accuracy with other ensemble classifiers, it has been shown that HGRFM can be a good alternative to standard Random Forests.

3. Result and Discussion

The Proposed DGRFM is evaluated using three different benchmark datasets NSL-KDD, UNSW-NB15, and CIC-IDS 2017 Datasets.

4.1.1 Performance Evaluation of DGRFM for NSL-KDD Dataset

For the NSL-KDD dataset, the detection rate of attacks, false alarm rates, Detection capability, and Aptness accuracy is calculated using the defined equations 1-5.

a) Detection Rate Comparison

In this part, the detection rate of attacks of the proposed method DGRFM is compared with the traditional algorithms using equation 1.

$$\text{Detection Rate} = \frac{TP}{TP + TN}$$

Table 1 Comparison Table of DGRFM for Detection Rate on NSL-KDD Dataset

Attacks/ Methods	DoS	Probe	U2R	R2L
DGA	73.12	74.36	73.80	78.20
RF	76.88	77.70	75.90	78.90
DGRFM	85.24	90.34	89.90	91.00

From table 1, RF for the four attacks DoS, Probing, U2R, and R2L give values of 76.88, 77.70, 75.90, and 78.90 Where else DGRFM gives 85.24, 90.34, 89.80, and 91.00 which are better than other methods.

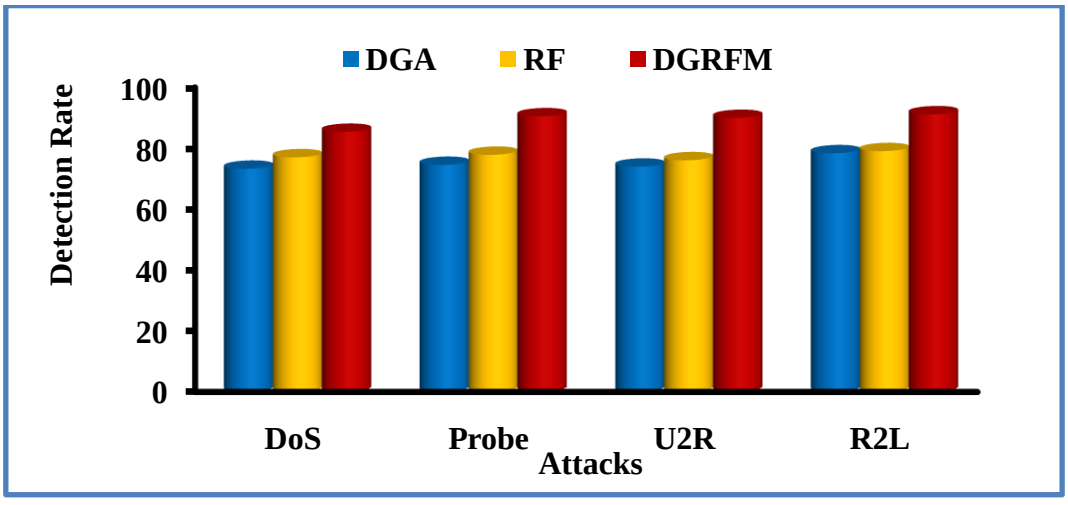

Figure 3 Comparison Graph of DGRFM for Detection Rate on NSL-KDD Dataset

It is clear that for the four attacks DoS, Probe, U2R, and R2L the detection rate of DGRFM gives a higher percentage of detection when compared with conventional methods.

b) False Alarm Rate Comparison

The proposed DGRFM is compared with the traditional algorithms using equation 2.

$$\text{False Alarm Rate} = \frac{FP}{FP + TN}$$

Table 2 Comparison Table of DGRFM for False Alarm Rate on NSL-KDD Dataset

Attacks/ Methods	DoS	Probe	U2R	R2L
DGA	3.64	2.92	2.42	2.50
RF	2.71	2.32	2.10	1.95
DGRFM	1.52	1.25	1.16	1.05

From table 2, The SVM and RF for four attacks give a higher alarm rate. For the proposed DGRFM, it produces a minimum rate of 1.52, 1.25, 1.16, and 1.05 which are better than the conventional methods.

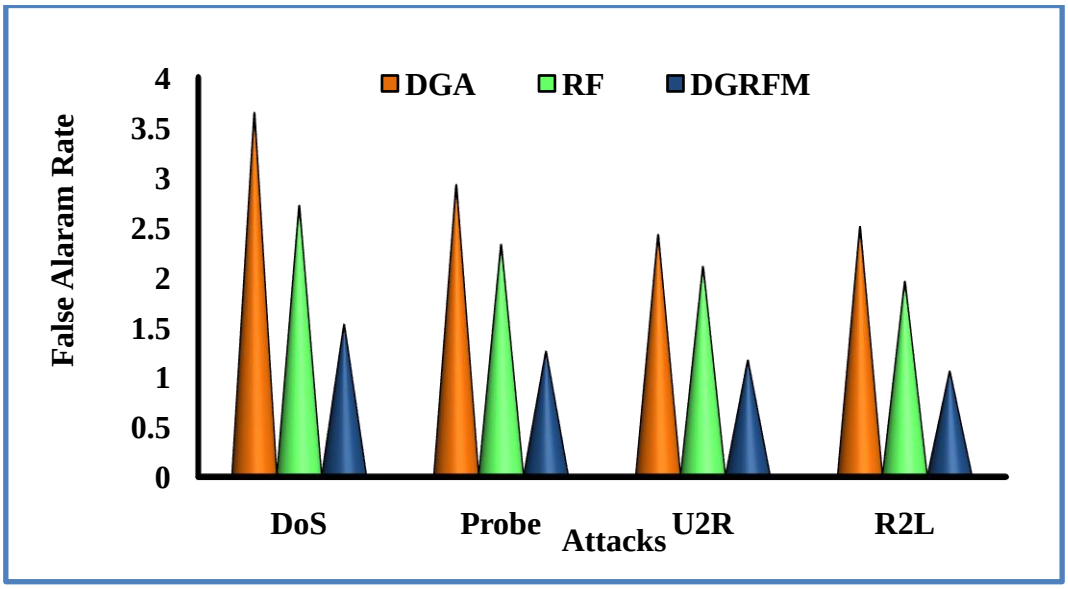

Figure 4 Comparison Graph of DGRFM for False Alarm Rate on NSL-KDD Dataset

From figure 4, the false alarm rate of DGRFM produces a minimum value when compared with the other conventional methods.

c) Intrusion Detection Capability

In this part, the Intrusion Detection Capability of the proposed method DGRFM is compared with the conventional algorithm using equation 3.

$$\text{Intrusion Detection Capability} = \frac{DR + TP}{\sqrt{2(FAR + TP)}}$$

Table 3 Comparison Table of DGRFM for Intrusion Detection Capability on NSL-KDD Dataset

Attacks/ Methods	DoS	Probe	U2R	R2L
DGA	73.30	68.60	70.30	73.00
RF	75.80	76.25	77.00	79.81
DGRFM	89.75	92.12	93.18	94.10

Table 3, For the proposed DGRFM, gives Detection capabilities of 89.75, 92.12, 93.18, and 94.10 which are better than other conventional methods.

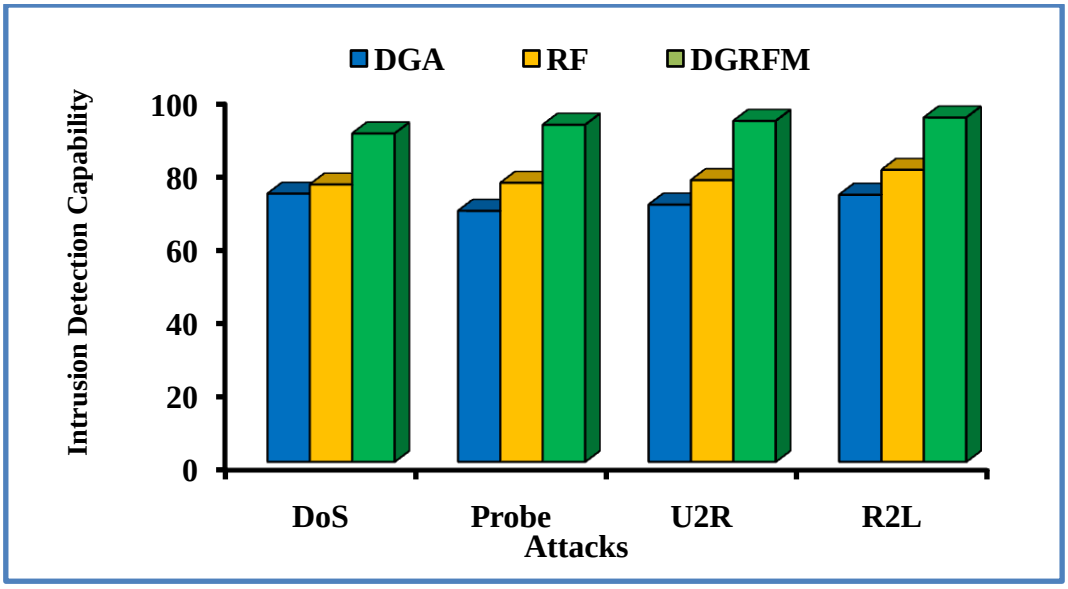

Figure 5 Comparison Graph of DGRFM for Intrusion Detection Capability on NSL-KDD Dataset

From figure 5, it can be seen that the Detection Capability of DGRFM produces higher when compared with the other conventional methods.

d) Aptness Accuracy

The Aptness Accuracy of the proposed method DGRFM is compared with the conventional RF Method algorithm using equation 4.

$$\text{Aptness Accuracy} = \frac{DR - FP}{DC}$$

Table 4.4 Comparison Table of DGRFM for Aptness Accuracy on NSL-KDD Dataset

Attacks/ Methods	DoS	Probe	U2R	R2L
DGA	76.00	77.10	75.23	78.00
RF	84.80	82.45	85.25	84.81
DGRFM	89.78	92.50	90.18	90.45

From table 4, the Aptness Accuracy of DGRFM produces the highest accuracy for all four attacks when compared with conventional methods.

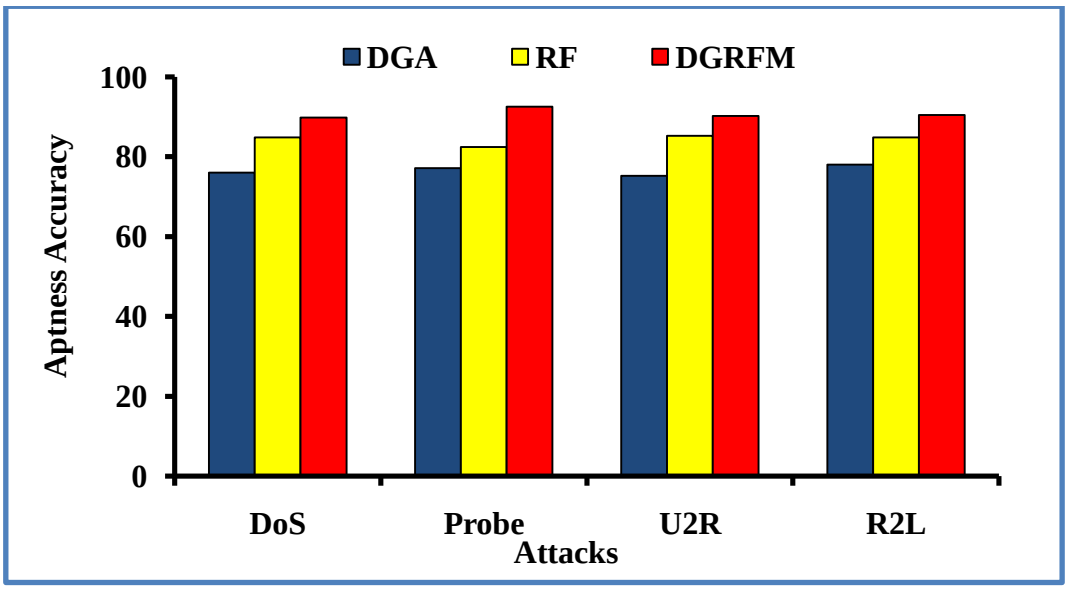

Figure 6 Comparison Graph of DGRFM for Aptness Accuracy on NSL-KDD Dataset

From figure 6, it can be seen that the Detection Capability of DGRFM produces higher when compared with the conventional methods.

e) Weighted Average

The Weighted Average of the proposed method DGRFM is compared with the traditional RF Method algorithm using equation 5.

$$\text{Weighted Average} = \frac{\sum_{i=1}^{C} PM_i \times n_i}{N}$$

Table 5 Comparison Table of DGRFM for Weighted Average on NSL-KDD Dataset

Attacks/ Methods	DoS	Probe	U2R	R2L
DGA	56.52	55.75	55.44	57.93
RF	60.05	59.68	60.06	61.37
DGRFM	66.57	69.05	68.58	69.15

From table 5, the Weighted Average Value of the proposed DGRFM produces a high value compared with the traditional RF and DGA.

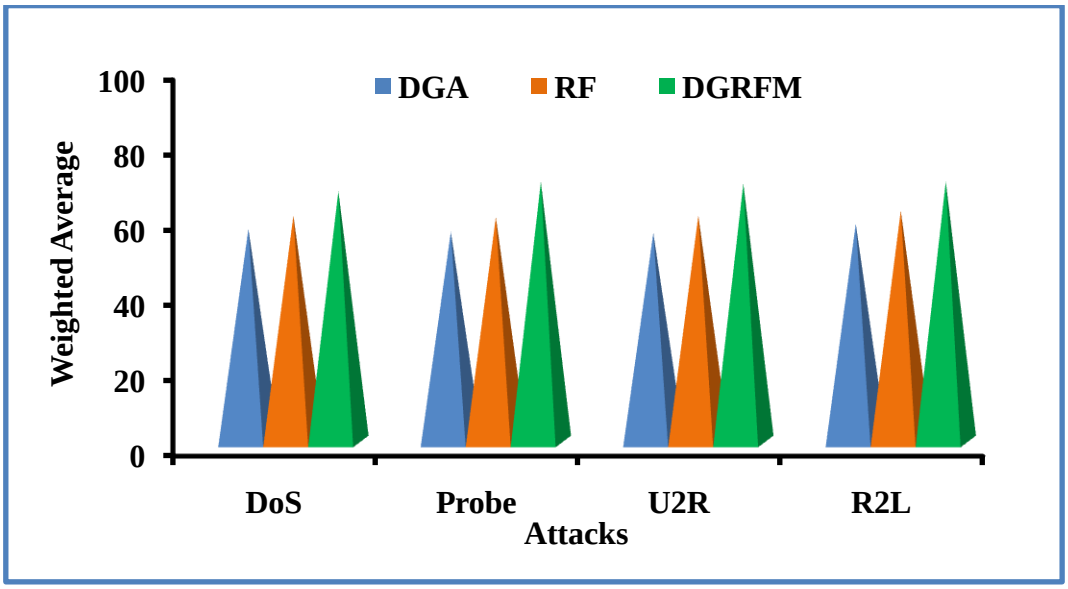

Figure 7 Comparison Graph of DGRFM for Weighted Average on NSL-KDD Dataset

From figure 7, it can be seen that the Detection Capability of DGRFM produces higher when compared with the other conventional methods.

4.1.2 Performance Evaluation of DGRFM for UNSW-NB15 Dataset

For the UNSW-NB15 Dataset, the detection rate of attacks, false alarm rates detection capability, and aptness accuracy is calculated using the defined equations 3.1, 3.2, 3.3, 3.4, and 3.5.

a) Detection Rate Comparison

In this part, the detection rate of attacks of the proposed method DGRFM is compared with the conventional algorithms.

Table 6 Comparison Table of DGRFM for Detection Rate on UNSW-NB15 Dataset

Attacks/Methods	Backdoor	DoS	Exploits	Fuzzers	Generic	Worms
DGA	77.56	74.39	76.91	71.26	79.22	67.72
RF	82.77	81.46	83.18	76.81	84.95	82.11
DGRFM	89.14	90.65	88.13	87.28	88.97	90.86

From table 6, for detecting intrusion attack detection DGRFM gives a better detection rate.

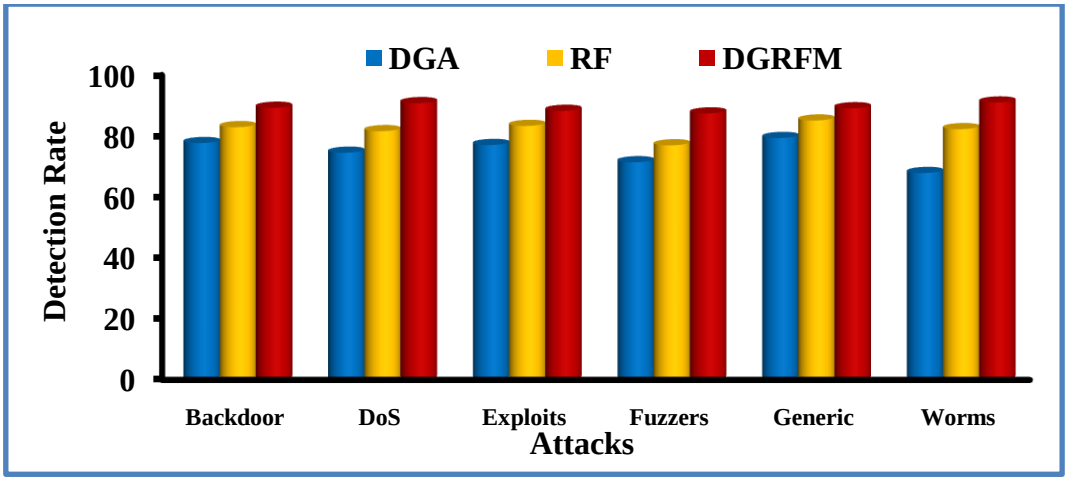

Figure 8 Comparison Graph of DGRFM for Detection Rate on UNSW-NB15 Dataset

From figure 8, the detection rate of DGRFM gives a higher percentage of detection when compared with other conventional methods.

b) False Alarm Rate Comparison

In this part, the false alarm rate of attacks of the proposed method DGRFM is compared with the traditional DGA and RF methods**.**

Table 7 Comparison Table of DGRFM for False Alarm Rate on UNSW-NB15 Dataset

Attacks / Methods	Backdoor	DoS	Exploits	Fuzzers	Generic	Worms
DGA	4.07	3.92	3.45	3.64	4.07	3.01
RF	3.45	2.68	3.51	2.71	2.87	2.27
DGRFM	2.64	1.32	1.94	1.27	1.01	1.12

From table 7, the DGA for the attacks gives a higher alarm rate. The proposed DGRFM produces minimum rates that are better than DGA, RF rates.

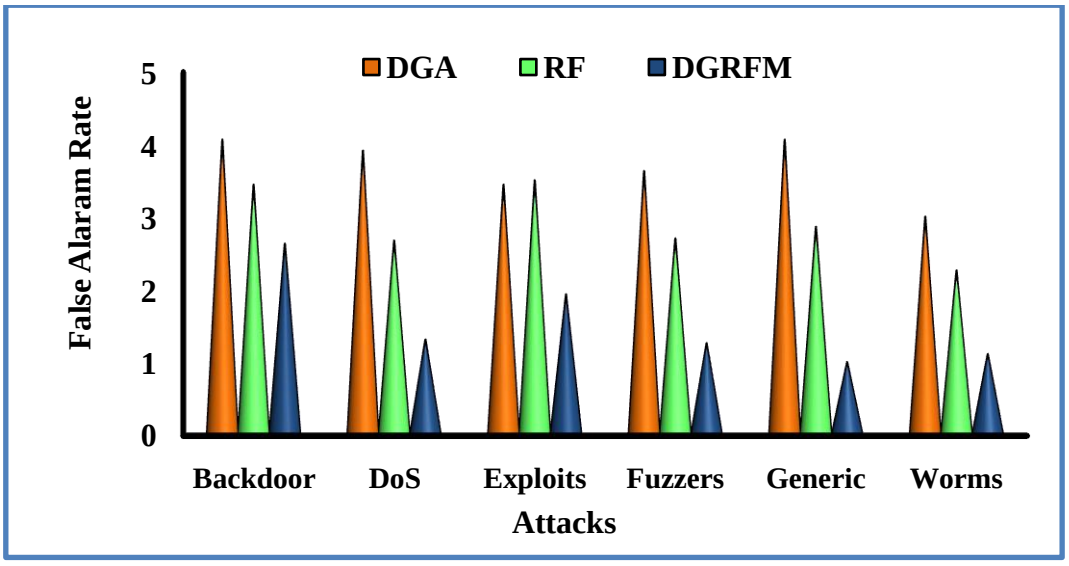

Figure 9 Comparison Graph of DGRFM for False Alarm Rate on UNSW-NB15 Dataset

From figure 9, it can be seen that the attacks for the false alarm rate of DGRFM produce a minimum value when compared with the traditional RF Method and DGA.

c) Intrusion Detection Capability

In this part, the intrusion detection capability of the proposed method DGRFM is compared with the traditional RF and also DGA algorithms.

Table 8 Comparison Table of DGRFM Intrusion Detection Capability on UNSW-NB15 Dataset

Attacks/Methods	Backdoor	DoS	Exploits	Fuzzers	Generic	Worms
DGA	80.07	81.92	71.45	63.64	74.07	62.01
RF	83.45	80.68	77.51	88.71	85.87	71.27
DGRFM	86.64	91.32	87.94	91.27	90.01	90.95

From table 8, the intrusion detection capability of DGRFM produces the highest for all attacks. So, the proposed DGRFM detection capabilities are better than RF capability.

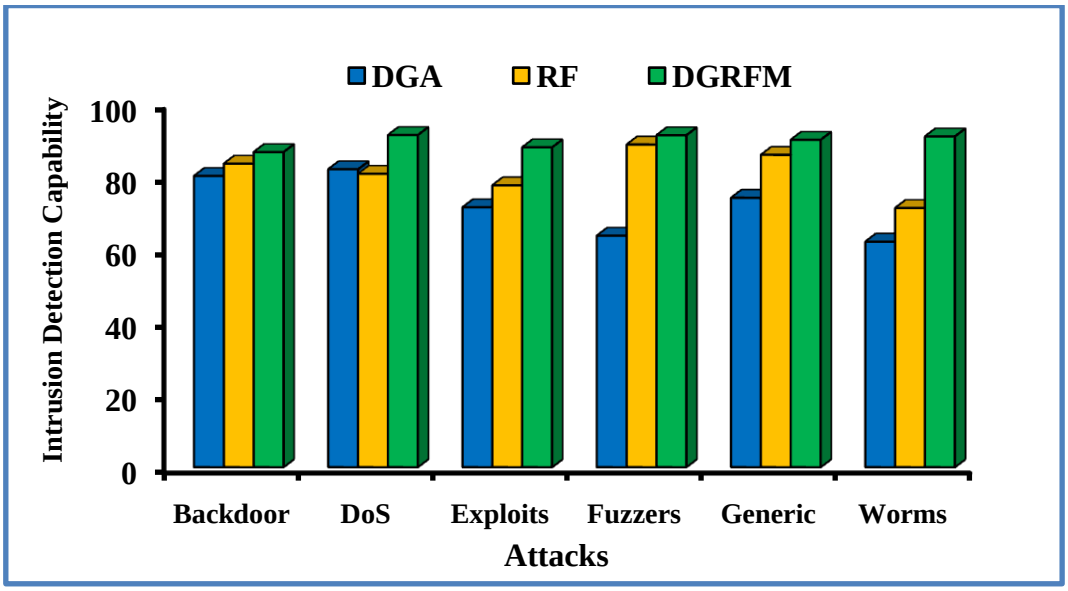

Figure 10 Comparison Graph of DGRFM for Intrusion Detection Capabilityon UNSW-NB15 Dataset

From figure 10, it can be seen that for the attacks DGRFM produces higher accuracy when compared with conventional RF.

d) Aptness Accuracy

In this part of the Aptness Accuracy of the proposed method DGRFM is compared with the traditional RF algorithm.

Table 9 Comparison Table of DGRFM for Aptness Accuracy on UNSW-NB15 Dataset

Attacks/Methods	Backdoor	DoS	Exploits	Fuzzers	Generic	Worms
DGA	78.90	77.21	80.38	78.71	82.45	82.00
RF	84.21	85.45	86.11	82.45	84.12	89.45
DGRFM	89.64	91.32	92.94	91.27	90.01	92.95

From table 9, the Aptness Accuracy of DGRFM produces the highest for all attacks when compared with the conventional RF Method.

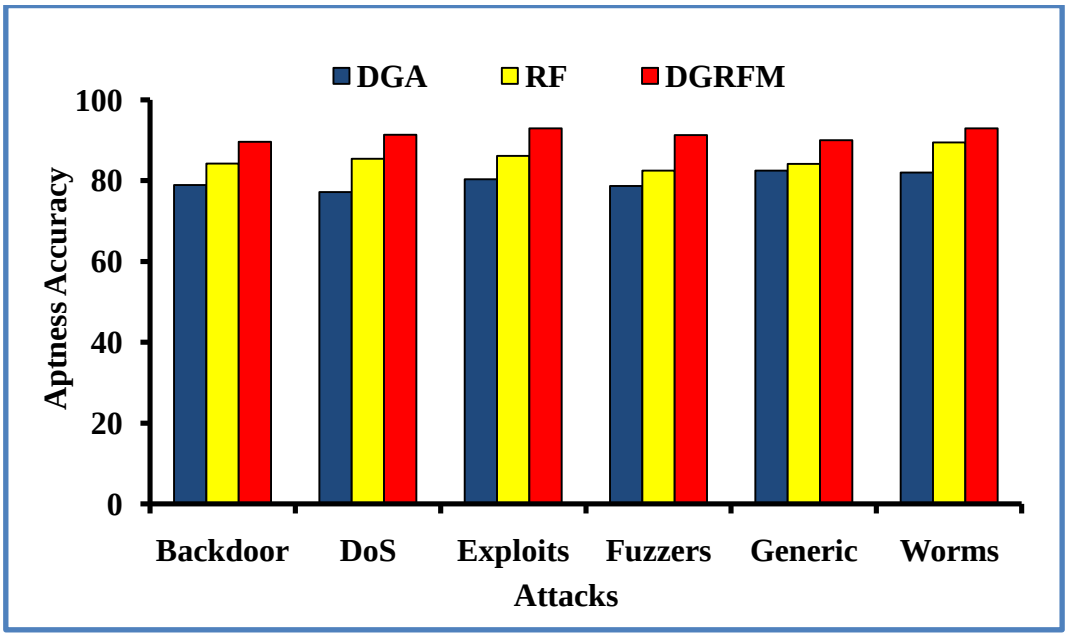

Figure 11 Comparison Graph of DGRFM for Aptness Accuracy on UNSW-NB15 Dataset

From Figure 11 The Aptness Accuracy of DGRFM provides good results than the conventional RF and DGA methods.

e) Weighted Average

The Weighted Average of the proposed method DGRFM is compared with the traditional RF Method algorithm using equation 5.

Table 10 Comparison Table of DGRFM for Weighted Average on UNSW-NB15 Dataset

Attacks/Methods	Backdoor	DoS	Exploits	Fuzzers	Generic	Worms
DGA	60.15	59.36	58.05	54.31	59.95	53.69
RF	63.47	62.57	62.58	62.67	64.45	61.28
DGRFM	67.02	68.65	67.74	67.77	67.50	68.97

From table 10, the Weighted Average Value of the proposed DGRFM produces a high value compared with the traditional RF and DGA methods.

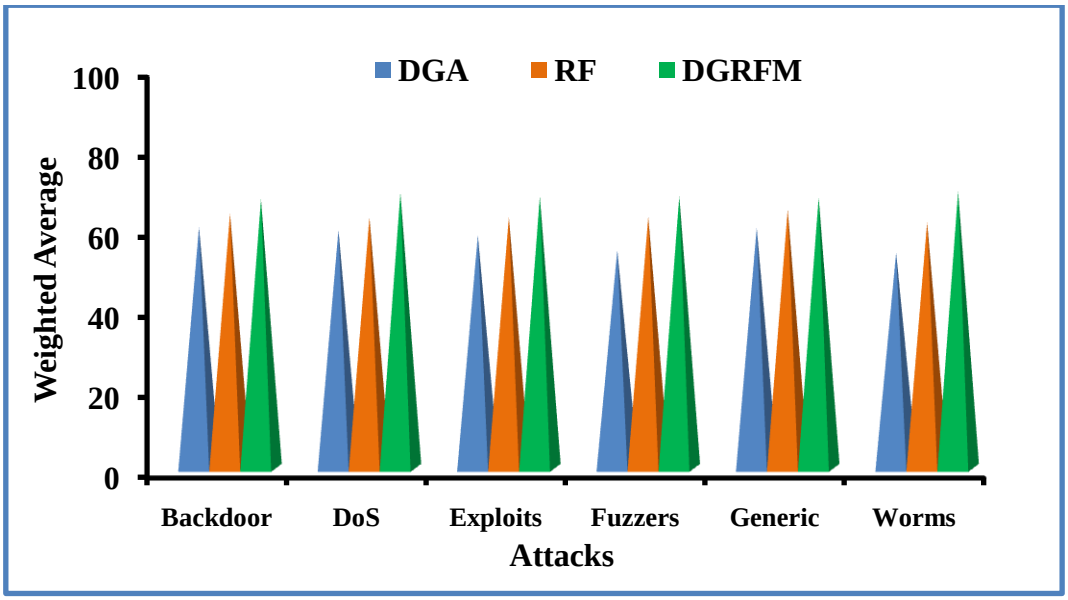

Figure 12 Comparison Graph of DGRFM for Weighted Average on UNSW-NB15 Dataset

From figure 12 the weighted average of the proposed DGRFM is better to compare with the conventional DGA and RF methods.

4.1.3 Performance Evaluation of DGRFM for CIC-IDS 2017 Dataset

For the CIC-IDS2017 dataset, the selected features are labeled under the seven categories for the detection rate of attacks, false alarm rates Detection Capability, and Aptness accuracy are calculated using the defined equations 3.1, 3.2, 3.3, 3.4, and 3.5.

a) Detection Rate Comparison

In this part, the detection rate of attacks of the proposed method DGRFM is compared with the conventional algorithms.

Table 11 Comparison Table of DGRFM for Detection Rate on CIC-IDS 2017 Dataset

Attacks/ Methods	Brute Force	Heart bleed	Web	Infiltration	Botnet	Port Scan	DDoS
SVM	77.40	76.00	75.10	74.94	74.39	74.91	76.80
RF	83.30	85.12	84.70	82.15	83.46	85.61	87.80
DGRFM	90.60	91.00	92.20	89.50	90.89	92.20	93.00

From table 11, for attacks, DGRFM gives the values which are better than the RF detection rate.

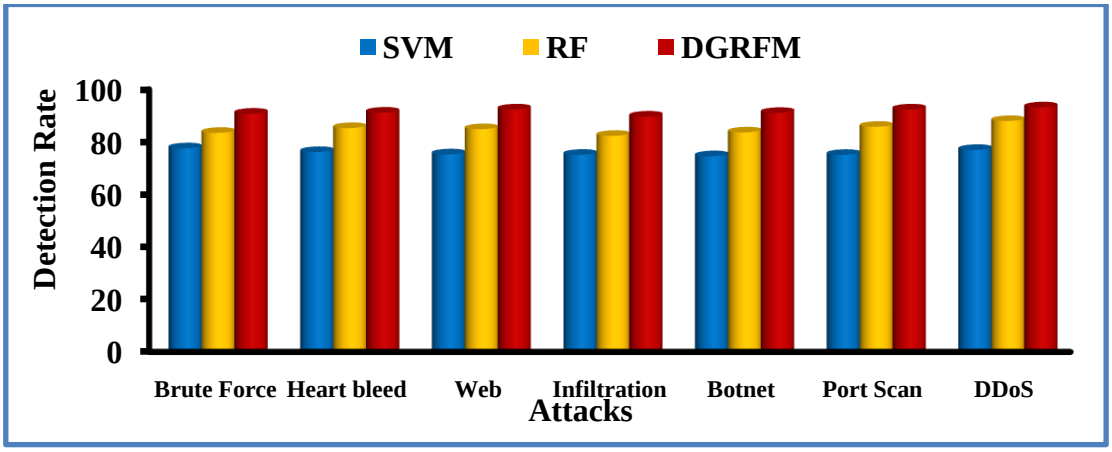

Figure 13 Comparison Graph of DGRFM for Detection Rate on CIC-IDS2017 Dataset

In figure 13, it is clear that the detection rate of DGRFM gives a higher percentage of detection when compared with the conventional RF algorithm.

b) False Alarm Rate Comparison

In this part, the false alarm rate of attacks of the proposed method DGRFM is compared with the traditional DGA and RF algorithms**.**

Table 12 Comparison Table of DGRFM for False Alarm Rate on CIC-IDS2017 Dataset

Attacks /	Brute	Heart bleed	Web	Infilt	Botnet	Port	DDoS

Methods	Force			ration		Scan	
SVM	3.67	4.32	3.28	3.27	2.46	3.27	2.09
RF	2.81	3.15	2.75	1.73	2.18	1.83	1.97
DGRFM	1.71	2.18	1.34	1.52	1.73	1.44	1.56

From table 12, the DGA for the attacks gives a higher alarm rate. For the proposed DGRFM, it produces a minimum rate that is better than DGA, RF Method False Alarm Rate.

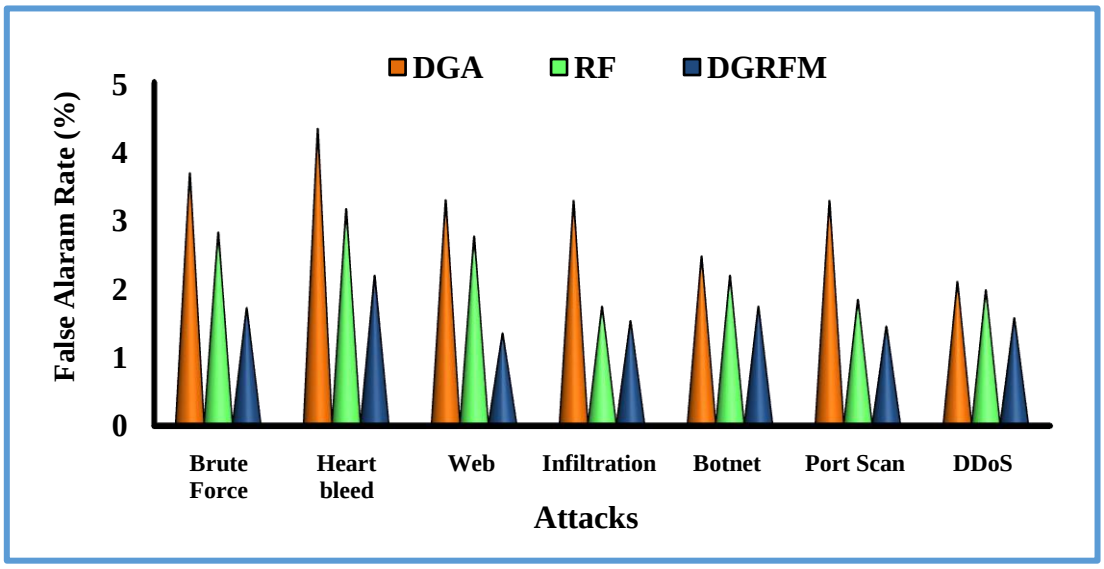

Figure 14 Comparison Graph of DGRFM for False Alarm Rate on CIC-IDS2017 Dataset

From figure 14, it can be seen that the attacks for the false alarm rate of DGRFM produce a minimum value when compared with traditional RF and DGA.

c) Intrusion Detection Capability

In this part, the Intrusion detection capability of the proposed method DGRFM is compared with the conventional RF Method and also DGA algorithms.

Table 13 Comparison Table of DGRFM for IntrusionDetection Capability on CIC-IDS2017 Dataset

Attacks/ Methods	Brute Force	Heart bleed	Web	Infilt ration	Botnet	Port Scan	DDoS
SVM	72.80	68.19	77.40	78.80	88.00	87.10	84.50

RF	75.45	80.34	82.00	84.60	89.50	79.70	87.00
DGRFM	81.39	83.00	92.45	89.00	91.40	90.90	94.00

From table 13, the detection capability of DGRFM produces comparatively higher detection capability with the conventional RF method.

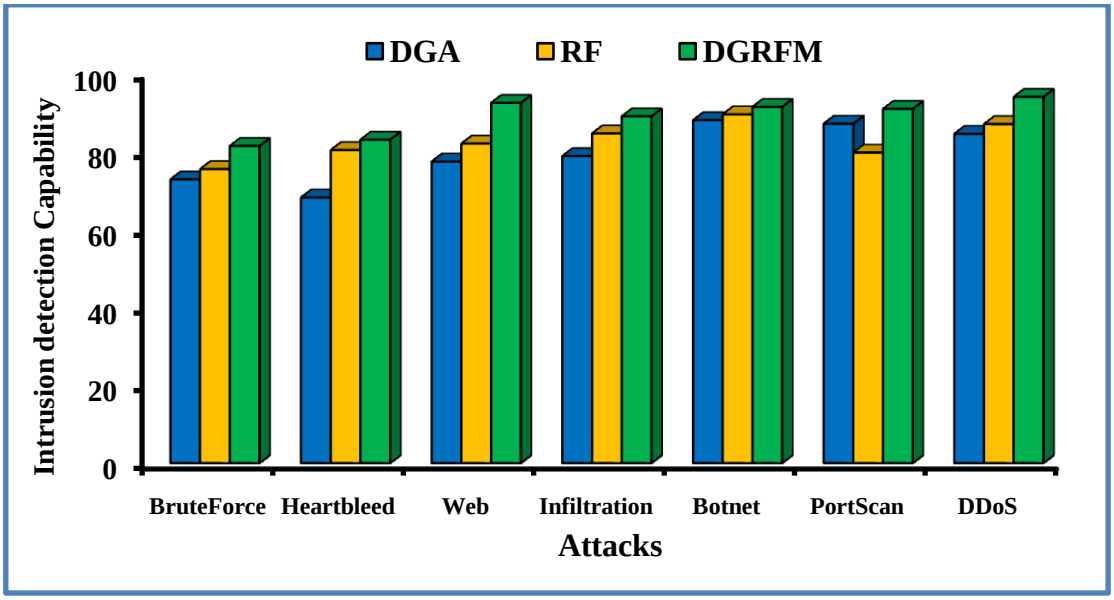

Figure 15 Comparison Graph of DGRFM for intrusion Detection Capability on CIC-IDS 2017 Dataset

In figure 15, seen that for the attacks DGRFM produces higher Intrusion detection capability when compared with the conventional RF Method.

d) Aptness Accuracy

In this part of the Aptness Accuracy of the proposed method DGRFM is compared with the conventional RF algorithm.

Table 14 Comparison Table of DGRFM for Aptness Accuracy on CIC-IDS 2017 Dataset

Attacks/ Methods	Brute Force	Heart bleed	Web	Infilt ration	Botnet	Port Scan	DDoS
SVM	78.39	73.12	70.50	79.55	75.40	72.00	76.36
RF	83.45	77.00	84.12	83.72	79.12	85.41	78.52

DGRFM	91.39	83.00	92.45	89.11	81.40	90.32	84.00

From table 14, the Aptness Accuracy of DGRFM produces the highest for all attacks when compared with the conventional RF Method.

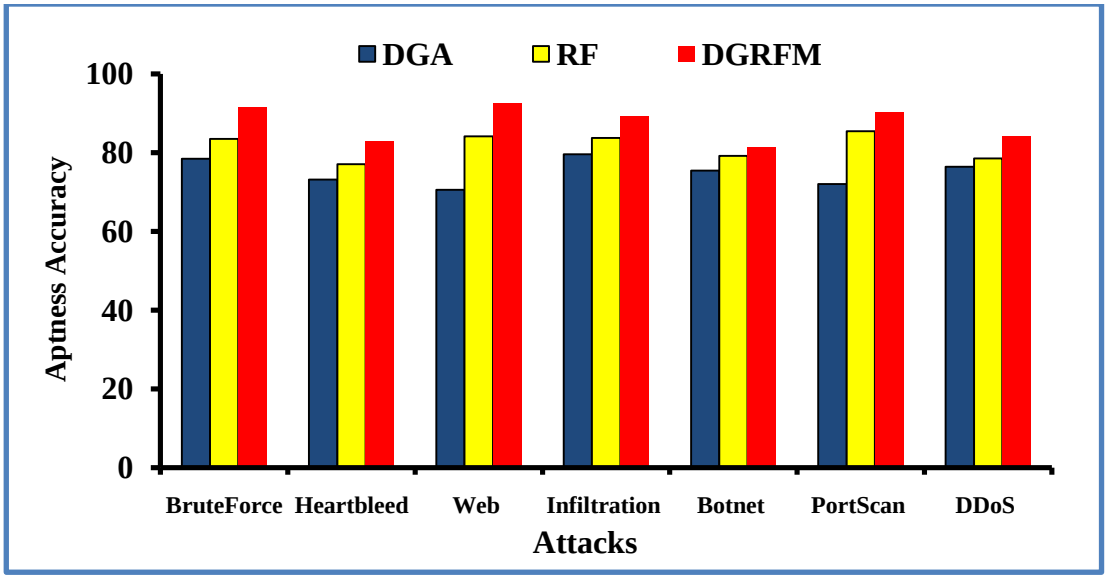

Figure 16 Comparison Graph of DGRFM for Aptness Accuracy on CIC-IDS 2017 Dataset

In figure 16, seen that for the attacks DGRFM produces higher Aptness Accuracy when compared with the conventional RF Method.

e) Weighted Average

The Weighted Average of the proposed method DGRFM is compared with the conventional RF Method algorithm using equation 5.

Table 15 Comparison Table of DGRFM for Weighted Average on CIC-IDS 2017 Dataset

Attacks/ Methods	Brute Force	Heart bleed	Web	Infilt ration	Botnet	Port Scan	DDoS
SVM	58.07	55.41	56.57	59.14	60.07	59.32	60.00
RF	61.26	61.41	63.40	63.05	63.14	63.14	63.83
DGRFM	66.28	64.80	69.61	67.29	66.36	68.72	68.14

From table 15, the Weighted Average of DGRFM produces the highest for all attacks when compared with the conventional RF Method. In Figure 4.17 the Weighted Average of DGRFM is better than RF and DGA.

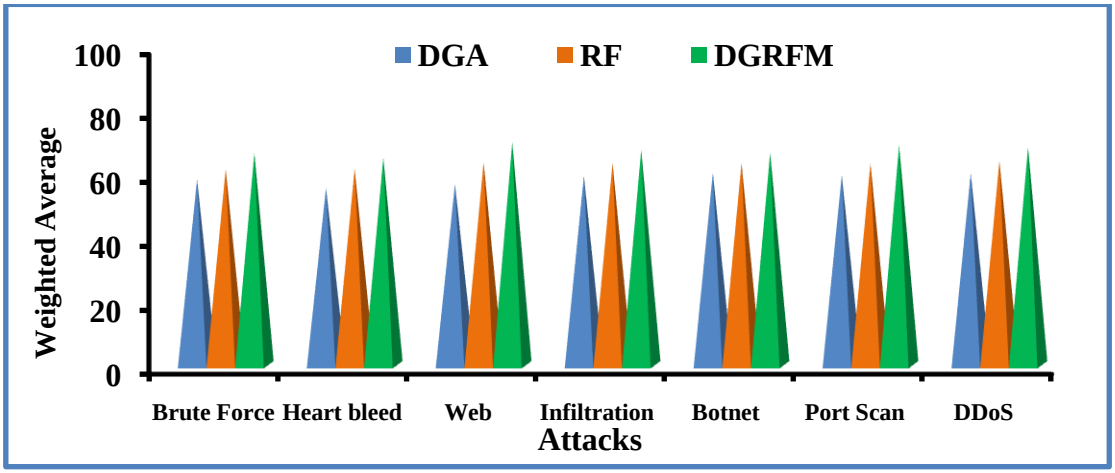

Figure 17 Comparison Graph of DGRFM for Weighted Average on CIC-IDS 2017 Dataset

In figure 17 shows that the weighted average of the proposed DGRFM is comparatively higher than the conventional DGA and RF Methods.

4. Conclusion

This paper focuses on Intrusion Detection systems using ML classification of attacks. It mainly deals with the DGRFM for Intrusion detection and acts as the classifier for the datasets. The experimental results of the table and graph obtained for the proposed DGRFM Method for Intrusion Detection are also given along with the discussions.

Reference

[1] Loganathan, G. Scholarship @ Western Real-time Intrusion Detection using Multidimensional Sequence- to-Sequence Machine Learning and Adaptive Stream Processing,2018.

[2] Luckert, M., & Schaffer-Kehnert, M. Using Machine Learning Methods for Evaluating the Quality of Technical Documents, 2015.

[3] Ma, L., & Fan, S. CURE-SMOTE algorithm and hybrid algorithm for feature selection and parameter optimization based on random forests. 1–18, 2017.

[4] Mahfouz, A., Abuhussein, A., Venugopal, D., & Shiva, S. Ensemble classifiers for network intrusion detection using a novel network attack dataset. Future Internet, 12(11), 1–19, 2020.

[5] Moghaddam, M. H., & Calix, R. A. (2015). Network Intrusion Detection Using a Hardware-Based Restricted Coulomb Energy Algorithm on a Cognitive Processor. 246–251,2015.

[6] Mourabit, Y. E. L., Bouirden, A., Toumanari, A., & Moussaid, N. E. L. (2015). Intrusion Detection Techniques in Wireless Sensor Network using Data Mining Algorithms : Comparative Evaluation Based on Attacks Detection. 6(9), 164–172,2015.

[7] Naganhalli, N. S., & Terdal, S. (2019). Network Intrusion Detection Using Supervised Machine Learning Technique. 8(09),2019.

[8] Narvaez, P. ., Hernández, E. V., & López-Lezama, J. M. (2013). Hybrid Genetic Algorithm for the Optimal Location of Distributed Generation in Distribution Systems. Simposio Internacional Sobre La Calidad de La Energía Eléctrica-SICEL, 7(52), 1–5, 2013.

[9] Othman, S. M., Alsohybe, N. T., Ba-alwi, F. M., & Zahary, A. T. (2018). Survey on Intrusion Detection System Types. December, 2018.

[10] Ahmad, I., Basheri, M., Iqbal, M. J., & Rahim, A. "Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection", IEEE Access, 6, Pp.33789–33795, 2018.

[11] Alqahtani, M., Gumaei, A., Mathkour, H., Maher, M., & Ismail, B. "A Genetic-Based Extreme Gradient Boosting Model". 2019.

[12] Alsmadi, I. "The integration of access control levels based on SDN", International Journal of High Performance Computing and Networking, 9(4), 281–290, 2016.

[13] Ambikavathi, C., & Srivatsa, S. K. "Predictor Selection and Attack Classification using Random Forest for Intrusion Detection. 79, (May), 365–368, 2020.

[14] Amjad Mohamed Al Tobi E., B. Anomaly-Based Network Intrusion Detection Enhancement by Prediction Threshold Adaptation of Binary Classification Models, 2018.

[15] Bang, J. ho, Cho, Y. J., & Kang, K. Anomaly detection of network-initiated LTE signaling traffic in wireless sensor and actuator networks based on a Hidden semi-Markov Model. Computers and Security, 65, 108–120, 2017.

[16] Belouch M, El Hadaj S, Idhammad M. Performance evaluation of intrusion detection based on machine learning using Apache Spark. Procedia Computer Science. Vol.127, pp:1–6, 2018.

[17] Baraneetharan, E. Role of Machine Learning Algorithms Intrusion Detection in WSNs : A Survey. 02(03), 161–173, 2020.

[18] Berman, M., Chase, J. S., Landweber, L., Nakao, A., Ott, M., Raychaudhuri, D., Ricci, R., & Seskar, I. GENI: A federated testbed for innovative network experiments. Computer Networks, 61, 5–23, 2014.

[19] Calders, T., Esposito, F., Hüllermeier, E., & Meo, R. (Eds.). Machine Learning and Knowledge Discovery in Databases (Vol. 8726). Springer Berlin Heidelberg, 2014.

[20] Chakraborty, A., Alam, M., Dey, V., Chattopadhyay, A., & Mukhopadhyay, D. Adversarial Attacks and Defences: A Survey. x(x), 2018.

[21] Chaudhari, R. R., & Patil, S. P. Intrusion Detection System : Classification Techniques And Datasets To Implement. 1860–1866, 2017.

[22] Haudhary, A., & Shrimal, G. Intrusion Detection System Based on Genetic Algorithm for Detection of Distribution Denial of Service Attacks in MANETs. 370–377, 2019.

[23] Chellam, A., Ramanathan, L., & Ramani, S. ScienceDirect ScienceDirect Intrusion Intrusion Detection Detection in in Computer Computer Networks Networks using using Lazy Lazy Learning Learning Algorithm Algorithm. Procedia Computer Science, 132, 928–936, 2018.

[24] Cheng, J., Huang, W., Cao, S., Yang, R., Yang, W., Yun, Z., Wang, Z., & Feng, Q. Enhanced performance of brain tumor classification via tumor region augmentation and partition. PLoS ONE, 10(10), 2015.

[25] Communications, B. Knowledge-Defined Networking : A Machine Learning based approach for network and traffic modeling, 2017.

[26] Dang, Q. Studying machine learning techniques for intrusion detection systems To cite this version : HAL Id : hal-02306521, 2019.

[27] Das, A. K., Nayak, J., Naik, B., Pati, S. K., & Pelusi, D. Computational Intelligence in Pattern Recognition, 2019.