International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

Ensemble Approach for DDoS Attack Detection in Cloud Computing Using Random Forest and GWO

Savita Devi,

Dept. of Computer Science, Jamia Millia Islamia University, New Delhi, India, Email: savita4992dahiya@gmail.com

Taran Singh Bharti,

Dept. of Computer Science, Jamia Millia Islamia University, New Delhi, India, Email: tbharti@jmi.ac.in

Abstract: When multiple technologies are added to a traditional network, it becomes increasingly difficult to meet newly imposed requirements, such as those regarding security. Since the widespread adoption of telecommunication technologies for the past decade, there have been an enhancement in the number of security threats that are more appealing. However, many new security concerns have arisen as a consequence of the introduction of the novel technology. One of the most significant of these is the potential for distributed denial of service attacks. Therefore, a DDoS detection method based on Random Forest Classifier and Grey Wolf Optimization algorithms in this work was developed to mitigate the DDoS threat. The results of the evaluation show that the Random Forest Classifier can achieve substantial performance improvements with respect to 99.96% accuracy. Comparison is also made to several state-of-the-art techniques for detecting of DDoS attacks for the real dataset.

Keywords: Distributed Denial of Service, Machine learning, Random Forest, Grey Wolf Optimization.

1. Introduction

The information technology (IT) industry is undergoing a period of transformation as a direct consequence of the fast growth of cloud computing over the past several years. The comparable ease of operability of cloud computing (CC), huge volumes of stored data, and apparent transparency make them easy and vulnerable targets for various different types of predatory attacks, the most significant of which are distributed denial of service (DDoS) attacks, for instance those against Spanhaus and Cloudflare, that are increasingly and alarmingly utilized for the exploitation of sample network management protocol (SNMP).

Reflection assaults, coercive parsing, large XML, port scanning, user to root, spoofing, flooding, and other similar methods are among the most common forms of DDoS attacks. It was stated that the possibilities of a DDoS attack happening are much higher than average. This is due to two factors: first, the fact that the tools needed to launch a DDoS are readily available, and second, the apparent absence of effective and timely defense mechanisms against DDoS attacks. The prevalence of assaults such as distributed denial of service puts CC's maximal gains and advantages at risk of being severely damaged or even nullified. There is a need for in-depth research on this area and its numerous repercussions over time, how DDoS has been evolving, and remedial activities triggered to offer permanent remedies and address them. These studies need to be supported by evidence and research that has been validated. CC might best handle these challenges and develop solutions are also needed. The description of the problem is that the actual degree and amount of damaging effects wrought by DDoS cannot be immediately quantified. The distributed denial of service attacks against CC has a significant potential to cause major damages and disadvantages, particularly if they are not regulated and no remediation is attempted.

2. Literature Review

The purpose of the research in [1,2] was to demonstrate that the use of cloud computing services is secure by constructing a cloud computing server and then subjecting it to a DOS (Denial of Service) assault. The term cloud computing [4], sometimes known as online computing, is referring to a kind of information technology computing service that offers application services, software, and hardware which could be attained through the use of the internet. The service should be customized to meet the requirements of the user, as well as fee for using the service is calculated based on the total amount of time or number of resources that are consumed on a monthly or per-minute basis [3].

A denial-of-service attack, often known as a DOS attack, is one in which the attacker makes a concerted effort to foil the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

efforts of authorized users in accessing network resources [8]. Attacks that use the denial-of-service technique seek mainly to render a network or computer inoperable. DoS attacks have lately been launched against numerous state of the art cloud-based organizations, such as Sony, Microsoft, Amazon EC2, and RackSpace [9].

DoS assaults could be carried out for a variety of reasons, including the desire to achieve sub-cultural status, obtain access, exact retribution, for political purposes, or for commercial reasons [10]. The purpose of a DoS attack is to prevent legitimate users from accessing servers. This can have a substantial impact on any activity that takes place online and a negative effect over the long term. The number of devices that make up targeted attack networks is significantly higher [11]. There are multiple forms of a DoS assault that can be launched against a cloud system, each of which has a unique purpose, set of criteria for the task, and scale [12]. DOS attacks could potentially lead to an increase in application demand, which would make it necessary to add more computing power to the added capacity [13].

During a DOS attack, thousands of data packets are sent to a single target in order to slow down all of the services that are being provided by the several computers that are attacking it at the same time [14]. Keeping all of the software and rules installed on a computer, on the other hand, is the best way to reduce the number of vulnerabilities and anomalies the system has [15].

In addition to that, [16] has suggested a denial-of-service defense system that is based on fog. The objective of an attack of this kind is to cause the Private Cloud Computing server to become overloaded with the task of fulfilling requests and, as a result, to either cease an activity or stop itself because it is unable to fulfil the demands that have been given to it. When carried out in this manner, an attack can sometimes cause harm to the system as a whole or even bring it to a complete halt.

3. Methodology

The goal of proposed research is to develop a cyber-attack detection system utilizing a combination of machine learning techniques. We will use the CICID2017 dataset, which contains network traffic data and labels indicating whether each connection is normal or an attack.

For enhancing the results of the machine learning methods, we will use several pre-processing techniques. These include feature selection and extraction, packet reconstruction, protocol encoding, dimension reduction, target value encoding, anomaly detection, clustering, and sampling techniques. By processing the raw data in these ways, we can extract the most relevant information and reduce the noise in the dataset.

We will use the Random Forest algorithm, which is a type of decision tree algorithm, and optimize its hyperparameters using the Gray Wolf Optimizer (GWO) algorithm. GWO is a type of metaheuristic technique that can find the optimal set of hyperparameters for the Random Forest model, which can enhance its results compared to traditional hyperparameter tuning methods.

In addition, we will use ensemble learning techniques to combine multiple instances of the Random Forest method to further enhance the accuracy and robustness of the system. This can be done using bagging, boosting, or stacking, depending on the performance of each method.

Finally, we will use feature importance analysis to identify the highly significant features for the Random Forest method. This can help us to further improve the performance of the model by focusing on the highly relevant features.

Overall, this project will develop a novel approach to cyber-attack detection using machine learning techniques and pre-processing methods. By combining Random Forest with GWO, ensemble learning, and feature importance analysis, we aim to achieve high accuracy and robustness in detecting cyber-attacks. Figure 3.1 gives the architecture of the proposed methodology.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

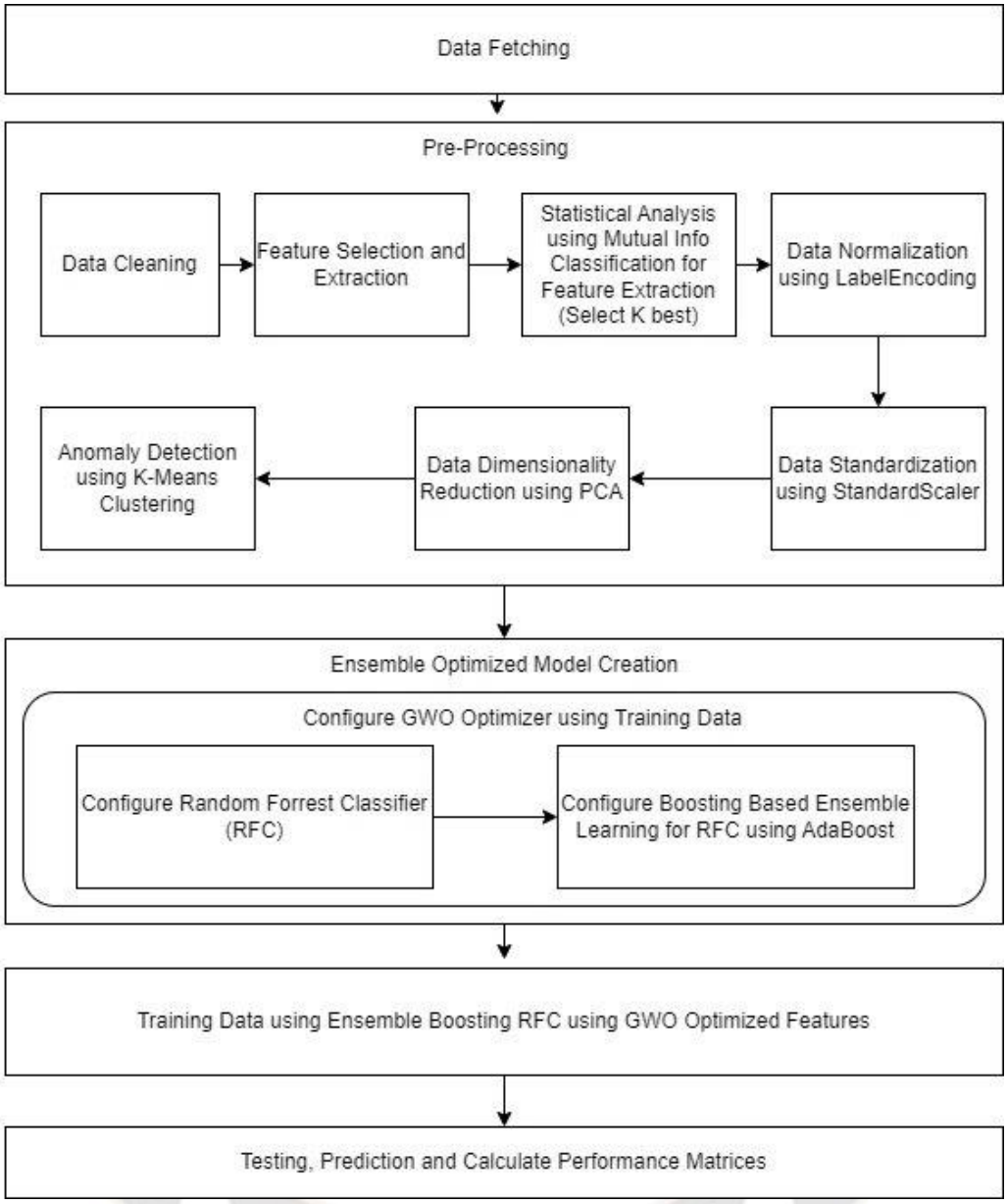

Figure 3.1: Architecture of the proposed methodology

3.1 Random Forest

Random Forest is a generic aspect of classifier combining which utilizes L tree-structure based classifier {h(X,Θn), N=1,2,3,…L}, where X is denoting the data at the input and {Θn} is a group of dependent and identical random distributed vectors. A random sample of the information included in the available data is used to construct each Decision Tree. For instance, a Random Forest for every Decision Tree (as in Random Subspaces) could be constructed by randomly picking a subset of features and/or by randomly sampling for every Decision Tree a training data subset (the idea of Bagging). Both of these methods are examples of random sampling.

A Random Forest is a classification method in which the features of each decision split are chosen at random. Randomly picking the features leads to a reduction in the correlation between the trees, which in turn enhances the prediction power and leads to a greater level of efficacy.

Since it could deal with missing values as well as binary, categorical, and continuous data, the Random Forest is an excellent option for modelling high-dimensional data sets. Random Forest is robust enough, thanks to bootstrapping and the ensemble scheme, to overcome the challenges posed by overfitting; as a result, there is no requirement to trim the trees.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

Random Forest is a classification method that, in addition to having a high prediction accuracy, is cost-effective, non-parametric, and interpretable for a wide variety of datasets [18]. The utilization of ensemble techniques in conjunction with random sampling allows for more accurate forecasts and more comprehensive generalizations.

3.2 Grey Wolf Optimization Algorithm

GWO is a novel optimization strategy for pack intelligence which is utilized extensively in a wide variety of important sectors. It does this by largely imitating the hierarchical structure and hunting behavior of grey wolf race packs, obtaining its optimal performance through the tracking, encircling, and pouncing activities of wolf packs. When compared to more traditional optimization strategies, such as PSO and GA, GWO's benefits include a reduced number of parameters, clearer principles, and easier application. It is well known that the results of the grey wolf technique in picking the greatest possible path can be affected by factors such as the proportional weighting technique, the convergence factor, and the initial wolf pack [19].

3.3 Performance metrics

For machine learning algorithms, the various performance metrices that are used to evaluate the performances are :

Precision: Precision is expressed as the quantity of correctly classified data in division by the overall quantity of accurately forecasted classified data in a given sample size.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall: The recall statistic is evaluated as the number of correctly classified data in division by the overall amount of correctly classified data.

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-Score: The F1-score is expressed as the harmonic mean of the precision as well as recall scores.

$$F1 - \text{Score} = \frac{2 * (\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}}$$

Accuracy: The accuracy of a dataset is evaluated by dividing the overall quantity of correct classifications on the dataset by the overall quantity of classifications in the dataset.

$$\text{Accuracy} = \frac{TP + FN}{TP + TN + FP + FN}$$

4. Results and discussion

4.1 Dataset

The CICID2017 dataset is used in the proposed research, which contains network traffic data and labels indicating whether each connection is normal or an attack. This dataset was chosen because it is a comprehensive and well-documented dataset that covers a wide range of cyber-attacks, including DoS, DDoS, port scans, and more. The dataset also includes a large number of features, which makes it suitable for feature selection and extraction techniques. Additionally, the dataset is frequently used as a benchmark for evaluating cyber-attack detection systems, which allows us for comparing our results with similar state-of-the-art techniques. Overall, the CICID2017 dataset provides a robust and realistic environment for evaluating and improving cyber-attack detection systems. The shape of the dataset is (55769, 79). The sample dataset is given in figure 4.1. The target labels are DDoS, Bot, Benign etc. The dataset type information of the 79 columns are Destination port, Flow duration, Total Forward Packets etc, Total Length of Forward packets.

Destination Port	Flow Duration	Total Fwd Packets	Total Length of Fwd Packets	Fwd Packet Length Max	Fwd Packet Length Min
54865	3	2	12	6	6
55054	109	1	6	6	6
55055	52	1	6	6	6
46236	34	1	6	6	6
54863	3	2	12	6	6

Figure 4.1: Sample dataset

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

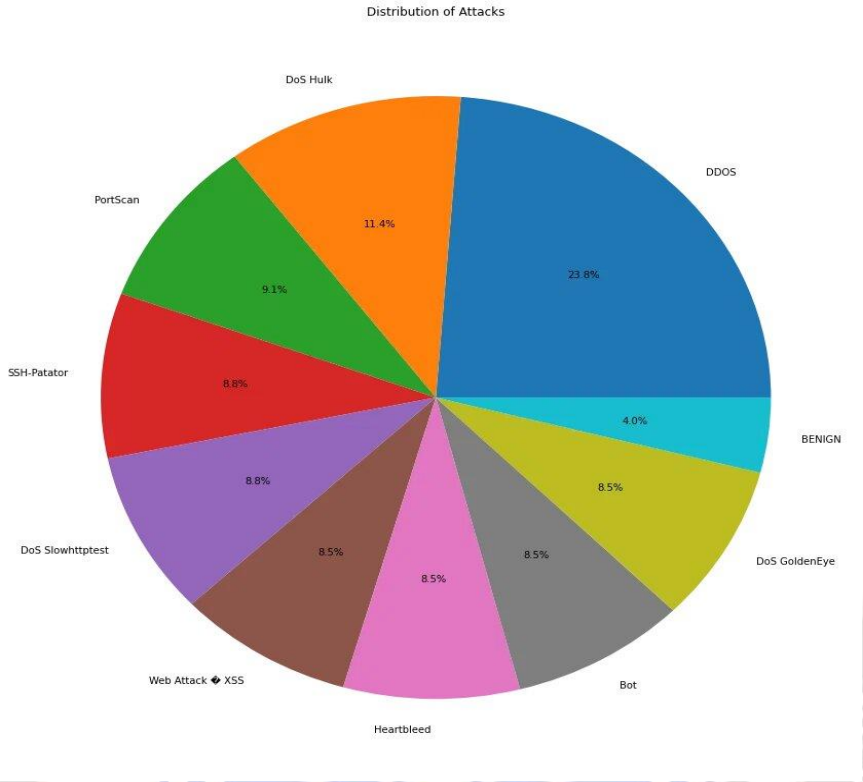

Figure 4.2: Attack Distribution Chart

4.2 Preprocessing

The duplicate values are later checked are removed. The number of duplicate values are 1217. After deleting the duplicate values, the size of the dataset is (55341, 79). Next the null values are checked and deleted. Shape of dataset After Deleting Null and NAN values is (55338, 79).

Next, labels after Dropping Low Sample Attack Label Data is considered. The labels are DDoS, DoS Hulk, SSH-Patator, DoS Slowhttptest, DoS GoldenEye, Bot, BENIGN.

The Attack Distribution Chart is given in figure 4.2. The DDoS attack is the highest with 23.8%, next is the DoS Hulk with 11.4% frequency followed by PortScan with 9.1% frequency and SSH-Patator with 8.8% frequency. The other attacks are minor ones.

Next the INF values are checked and dropped. Dataset after Checking and Dropping INF values is (55327, 79).

4.3 Feature Selection and Extraction

Feature selection and extraction is taking place using the variance threshold algorithm. A feature selection known as Variance threshold is applied to a dataset in order to eliminate any characteristics with a low variance that are not very helpful for modelling. Unsupervised learning is possible because it concentrates solely on the inputs, or features, rather than on the outcomes, or outputs. The value 0 is used for the Threshold setting by default. Next, Finding the Constraint and Non-Constraint Values is taking place. The number of non constant feature is 70. Next the non-constant features are dropped.

Corelation Confusion Matrix Heatmap of Training & Testing Data is considered. Iis common practise to calculate correlation coefficients in order to investigate the degree to which certain quantitative variables are connected with one another. When one variable increases and the other also increases, we say that there is a positive correlation between the two variables. On the other hand, they are said to have a negative correlation when the high values of one variable are seen to go hand in hand with the low values of another variable. Number of correlation features is 41. The correlated values are next dropped. Shape After Dropping Correlated Values is (54284, 40).

Mutual Information Classification is considered after correlation coefficient. Mutual information (MI), which measures the interdependence of two random variables, has a positive value. Larger value represents stronger dependence, while 0 indicates that two random variables are independent.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

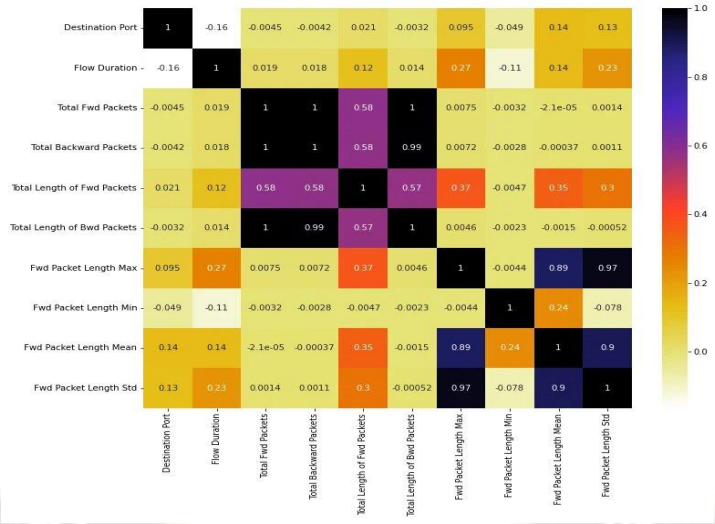

Figure 4.3: Correlation Coefficient Heatmap

Normalization of Target Values takes place after feature selection and extraction. The process of normalization is included in the procedures for data processing and cleaning. The primary objective of data normalization is to ensure that the data are consistent across all records and fields. It contributes to the creation of a linkage between the data entries, which, in turn, contributes to the cleaning and improvement of the data quality. Standardization of Actual Data is completed after normalization. Data standardization refers to the act of putting features that are not comparable on the same scale. Alternately stated, standardized data can be described as the attributes being rescaled in such a way that their mean becomes zero and their standard deviation becomes one.

Select K Best is also used for mutual classification. The features with best K values for Classification is given in figure 4.4.

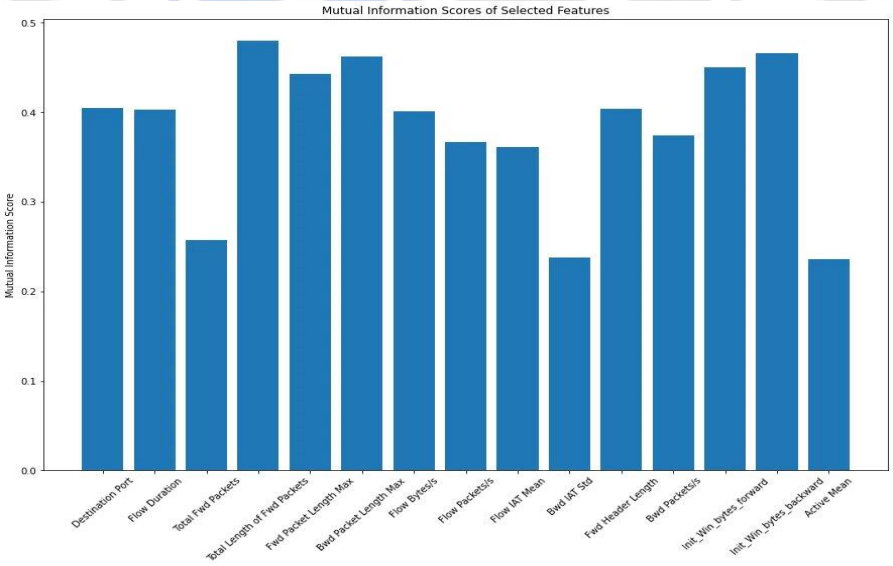

Figure 4.4: Features with best K values for Classification

PCA (Principal Component Analysis) is done on the dataset after standardization. PCA is a method for extraction of significant components (in the aspect of variables) from a vast variables set that are made useable in a data set. This is accomplished through the use of principal components. It does this by considering from a high-dimensional data set a projection of unimportant dimensions and then using that as a basis for extracting a low-dimensional set of features with

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

the goal of obtaining as much information as feasible. The visualization process also becomes significantly more relevant when fewer variables are obtained while the amount of information that is lost is minimized. When working with data that has three or more dimensions, PCA proves to be more beneficial.

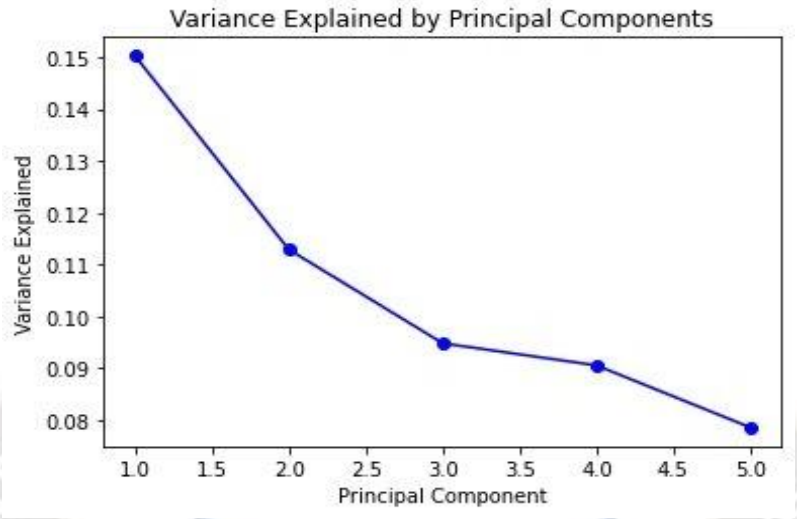

Figure 4.5: PCA Analysis and Transformation

Clustering using K Means for Anomalies Detection is the next step used. The K-means algorithm for grouping the data will continue to compute centroids until it identifies the one that is the most effective for clustering the data. It is reasonable to suppose that we are already familiar with the overall number of clusters. This specific method also goes by the name of the flat clustering model. The quantity of clusters that were uncovered from the data by utilising the methodology is represented by the letter 'K' in the term 'K-means,' which is also the name of the methodology.

The data points are then assigned to the clusters in a certain manner that the total of the squared distances among the centroid and the data points is as near to zero as is reasonably possible according to this approach. This is accomplished by assigning the data points are represented as the cumulation of the squared distances among the centroid as well as the data points is exactly zero. It is essential to keep in mind that a reduction in the diversity of elements included inside clusters results in an increase in the number of data points that are identical to one another while still being contained within the same cluster.

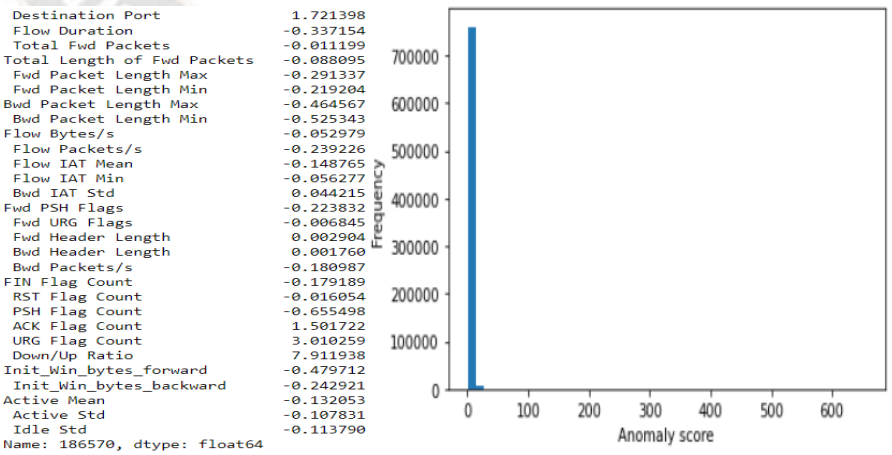

Figure 4.6: Anomalies Based Features and Its Frequency Graph

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

4.4 Training and Testing of the dataset

The dataset after K Means clustering is split into train and test dataset. The Train Test is Split in the 80-20 Ratio. The rows and columns that are split are ((9740, 33), (2436, 33), (9740,), (2436,)).

The Random Forrest Classifier is first initialized. Training of Data using Base Random Forrest takes place. The cross validation mean score is 0.989. The model accuracy is 0.995. The confusion matrix of Random Forrest Classifier is represented in figure 4.7. The classification report is represented in figure 4.8.

```
Confusion matrix:
 [[302   0   0   2   0   0   0   0]
 [  1 302   0   1   0   0   0   0]
 [  0   0 304   0   0   0   0   0]
 [  0   0   0 305   0   0   0   0]
 [  0   0   0   0 302   3   0   0]
 [  2   0   0   0   1 302   0   0]
 [  0   0   0   0   0   0 304   0]
 [  1   0   0   0   0   0   0 304]]
```

Figure 4.7 : Confusion Matrix of Random Forrest Classifier

```
Classification report:
                  precision    recall  f1-score   support

          BENIGN       0.99      0.99      0.99       304
            DDoS       1.00      0.99      1.00       304
   DoS GoldenEye       1.00      1.00      1.00       304
        DoS Hulk       0.99      1.00      1.00       305
DoS Slowhttptest       1.00      0.99      0.99       305
   DoS slowloris       0.99      0.99      0.99       305
     FTP-Patator       1.00      1.00      1.00       304
        PortScan       1.00      1.00      1.00       305

        accuracy                           1.00      2436
       macro avg       1.00      1.00      1.00      2436
    weighted avg       1.00      1.00      1.00      2436
```

Figure 4.8 : Classification report of Random Forrest Classifier

AdaBoost classifier is used for Ensemble Learning and training for Random Forrest Segment. An AdaBoost classifier is a meta-estimator which operates by initially fitting a classifier on the first set of data, and then fitting extra copies of the classifier on the identical dataset, but employing the weights of incorrectly classified instances to adjust according to a way which after classifiers focus more on difficult cases. This process is repeated until the desired accuracy has been achieved. Classifiers built with AdaBoost are able to produce results that are more accurate than those generated by ordinary meta-estimators. This process is repeated a great number of times until the most accurate classifier is discovered. The cross validation mean score is 0.991 and the model accuracy is 0.995.

4.6 Optimization using Grey Wolf Optimization

The data is optimized using GWO and then training using Ensemble Random Forrest Model takes place. The Best Fitness Graph is represented in figure 4.9. The term convergence refers to the steady point that is reached at the conclusion of a progression of solutions reached via the use of an iterative optimization technique. A stable point that is identified too soon, sometimes in close proximity to the initial point of the search, and with a lower evaluation than was anticipated is referred to as having premature convergence.

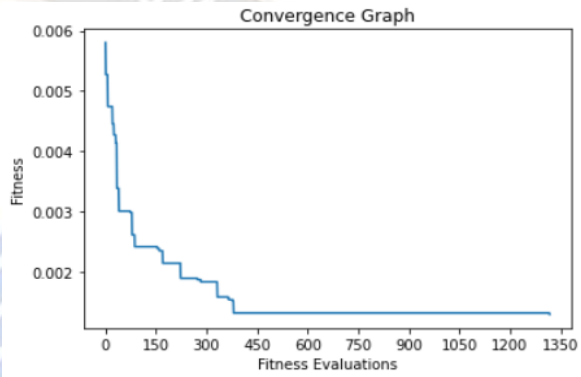

Figure 4.9 : Best Fitness Graph

Global/Local Best Objective Graph is represented in figure 4.10. A viable solution that possesses an objective value that is either equal to or superior to that of all other feasible solutions to the model is referred to as a globally optimal solution.

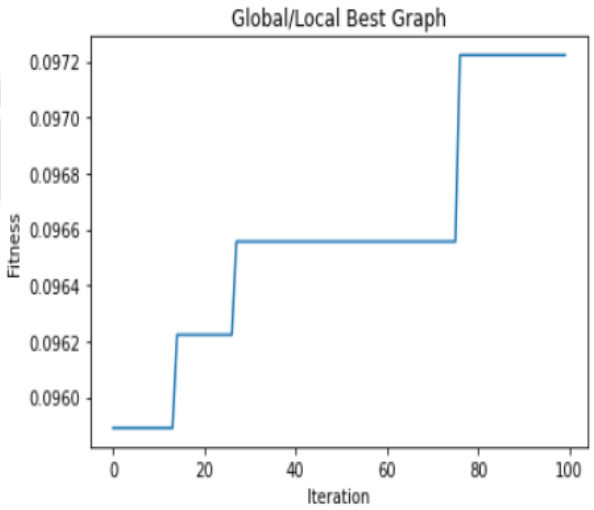

Figure 4.10 : Global Best Local Best Objective Graph

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

Figure 4.11 shows the execution time taken for finding best optimum point per iteration, and the second graph shows the Partical Exploration for finding best optimum points and its validity. Number of best Optimized Features is 8.

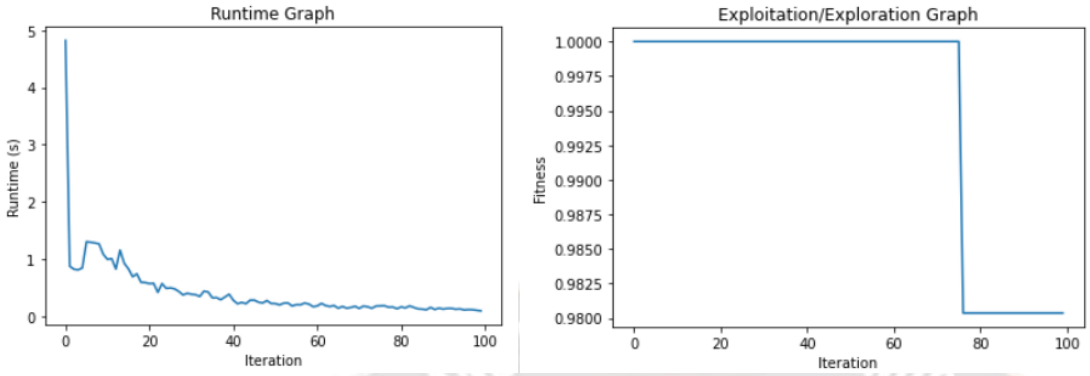

Figure 4.11 : Runtime graph and Exploitation/Exploration graph

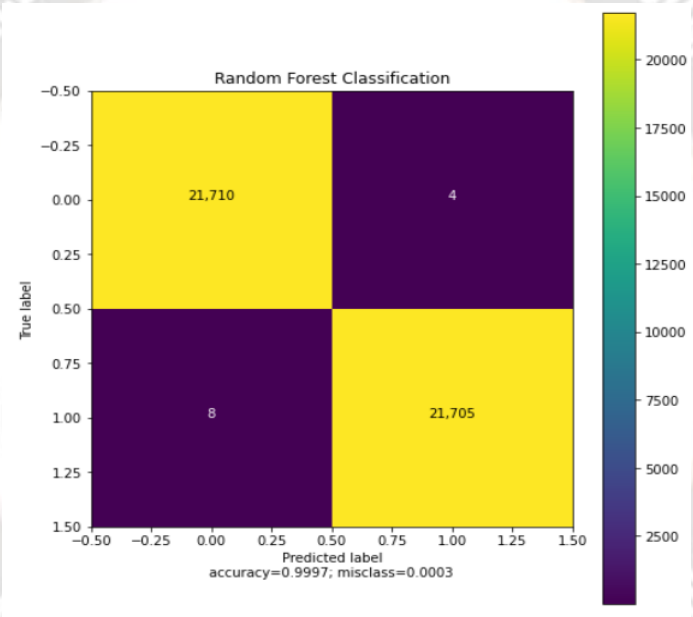

Figure 4.12 : Confusion matrix of the proposed model

The training Accuracy is 0.999 and Testing and Prediction Based Final Performance Scores are: cross validation mean score value is 0.9996 and the model accuracy score is 0.999.

The confusion matrix of the combined model with optimization is given in figure 4.12. The classification report of the presented method is represented in figure 4.13.

```
Classification report:
              precision    recall  f1-score   support

      BENIGN       1.00      1.00      1.00     21714
        DDoS       1.00      1.00      1.00     21713

    accuracy                           1.00     43427
   macro avg       1.00      1.00      1.00     43427
weighted avg       1.00      1.00      1.00     43427
```

Figure 4.13 : Classification report of the proposed model

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

4.7 Comparison of the proposed method with similar state of the art techniques

The proposed technique is compared with Random Forest, Boosting Ensemble RFC, Logistic Regression, Decision Tree and K NeighorsClassifiers. It could be understood from figure 4.14 that the scores of the presented technique is higher compared to similar methods.

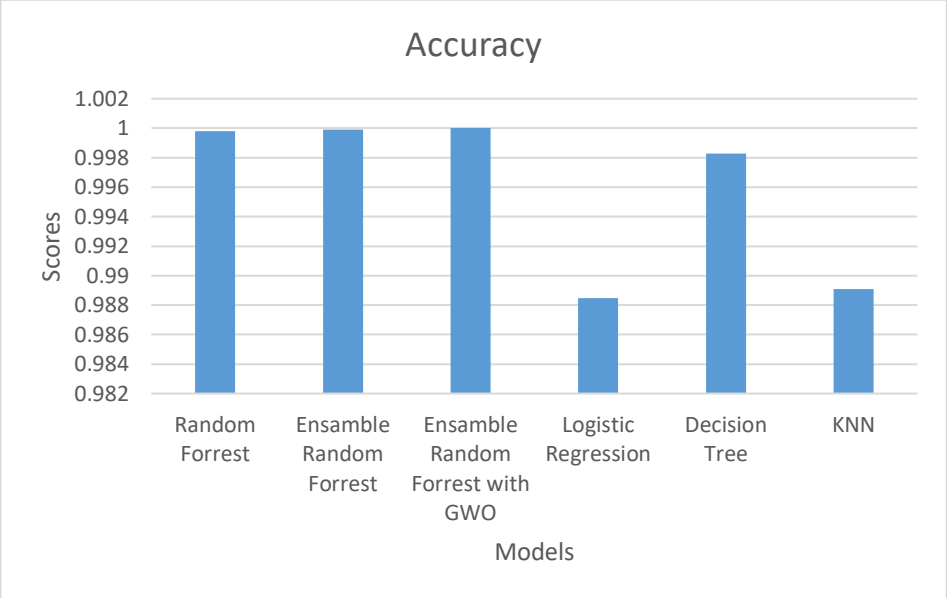

Figure 4.14: Comparison of proposed technique with state of art techniques in the literature

5. Conclusion

In this study, we have proposed a novel hybrid algorithm for DDoS attack detection. DDoS attacks continue to be a difficult obstacle in cloud computing environments, and in order to combat them, sophisticated simulations which are used in this research are required. The Random Forest Algorithm have been used to detect irregular traffic flow and Grey Wolf Optimization is utilized for optimization of performance. Based on the findings that were achieved, we have evaluated that the Random Forest classifier method is better to detect DDoS attacks since it produced results than those produced by other machine learning techniques. At last it is concluded that random forest algorithm is more capable and suitable for DDoS attack detection. It is predicting highest accuracy with 99.96% in comparison of other algorithms such as logistic regression, decision tree and KNN algorithm.

References

1. M. Chhabra and et al., "A novel solution to handle ddos attack in manet," 2013.
2. B. Gupta, P. Chaudhary, X. Chang, and N. Nedjah, "Smart defense against distributed denial of service attack in iot networks using supervised learning classifiers," Computers & Electrical Engineering, vol. 98, p. 107726, 2022.
3. T. Wedge, "The basics of digital forensics," Computers and Security, vol. 31, no. 6, p. 800, 2012.
4. J. Wang and et al., "Pcnncec: Efficient and privacy-preserving convolutional neural network inference based on cloud-edge-client collaboration," IEEE Transactions on Network Science and Engineering, 2022.
5. R. Li, J. H. Fan, and X. B. Wang, "Technique of constructing cloud computing platform based on ubuntu enterprise cloud," in Advanced Materials Research, vol. 482. Trans Tech Publ, 2012, pp. 713–716
6. A. Bonguet and M. Bellaiche, "A survey of denial-of-service and distributed denial of service attacks and defenses in cloud computing," Future Internet, vol. 9, no. 3, p. 43, 2017
7. G. Grispos, T. Storer, and W. B. Glisson, "Calm before the storm: The challenges of cloud computing in digital forensics," International Journal of Digital Crime and Forensics (IJDCF), vol. 4, no. 2, pp. 28–48, 2012.
8. S. Singh, Y.-S. Jeong, and J. H. Park, "A survey on cloud computing security: Issues, threats, and solutions,"

Journal of Network and Computer Applications, vol. 75, pp. 200–222, 2016
9. P. Nelson, "Cybercriminals moving into cloud big time, report says," Network world, 2015.
10. Z. Hui, "A design of distributed collaborative intrusion detection model," in 2011 6th International Conference on Computer Science & Education (ICCSE). IEEE, 2011, pp. 99–101.
11. S. Dong, K. Abbas, and R. Jain, "A survey on distributed denial of service (ddos) attacks in sdn and cloud computing environments," IEEE Access, vol. 7, pp. 80 813–80 828, 2019.
12. G. Somani, M. S. Gaur, D. Sanghi, M. Conti, M. Rajarajan, and R. Buyya, "Combating ddos attacks in the cloud: requirements, trends, and future directions," IEEE Cloud Computing, vol. 4, no. 1, pp. 22–32, 2017.
13. N. Z. Bawany, J. A. Shamsi, and K. Salah, "Ddos attack detection and mitigation using sdn: methods, practices, and solutions," Arabian Journal for Science and Engineering, vol. 42, no. 2, pp. 425–441, 2017.
14. A. K. Soliman, C. Salama, and H. K. Mohamed, "Detecting dns reflection amplification ddos attack originating from the cloud," in 2018 13th International Conference on Computer Engineering and Systems (ICCES). IEEE, 2018, pp. 145–150. VOLUME 4, 2022 5 M.Rahaman et al./ Cyber Security Insights Magazine, Vol 04, 2022
15. A. Amjad, T. Alyas, U. Farooq, and M. A. Tariq, "Detection and mitigation of ddos attack in cloud computing using machine learning algorithm," EAI Endorsed Transactions on Scalable Information Systems, vol. 6, no. 23, pp. e7–e7, 2019.
16. S. Dong, K. Abbas, and R. Jain, "A survey on distributed denial of service (ddos) attacks in sdn and cloud computing environments," IEEE Access, vol. 7, pp. 80 813–80 828, 2019.
17. Balyan, A. K., Ahuja, S., Lilhore, U. K., Sharma, S. K., Manoharan, P., Algarni, A. D., ... & Raahemifar, K. (2022). A hybrid intrusion detection model using ega-pso and improved random forest method. *Sensors*, *22*(16), 5986.
18. Fan, G. F., Zhang, L. Z., Yu, M., Hong, W. C., & Dong, S. Q. (2022). Applications of random forest in multivariable response surface for short-term load forecasting. *International Journal of Electrical Power & Energy Systems*, *139*, 108073.
19. Makhadmeh, S. N., Alomari, O. A., Mirjalili, S., Al-Betar, M. A., & Elnagar, A. (2022). Recent advances in multi-objective grey wolf optimizer, its versions and applications. *Neural Computing and Applications*, *34*(22), 19723-19749.
20. Gupta, I., Gupta, R., Singh, A. K., & Buyya, R. (2020). MLPAM: A machine learning and probabilistic analysis based model for preserving security and privacy in cloud environment. *IEEE Systems Journal*, *15*(3), 4248-4259. Link: - https://www.kaggle.com/datasets/cicdataset/cicids2017.
21. Kumar, V. and Kumar, R., 2015. An adaptive approach for detection of blackhole attack in mobile ad hoc network. Procedia Computer Science, 48, pp.472-479.
22. Kumar, V. and Kumar, R., 2015. An optimal authentication protocol using certificateless ID-based signature in MANET. In Security in Computing and Communications: Third International Symposium, SSCC 2015, Kochi, India, August 10-13, 2015. Proceedings 3 (pp. 110-121). Springer International Publishing.
23. Kumar, V. and Kumar, R., 2015, April. Detection of phishing attack using visual cryptography in ad hoc network. In 2015 International Conference on Communications and Signal Processing (ICCSP) (pp. 1021-1025). IEEE.
24. Kumar, V., Shankar, M., Tripathi, A.M., Yadav, V., Rai, A.K., Khan, U. and Rahul, M., 2022. Prevention of Blackhole Attack in MANET using Certificateless Signature Scheme. Journal of Scientific & Industrial Research, 81(10), pp.1061-1072.
25. Kumar, Vimal, and Rakesh Kumar. "A cooperative black hole node detection and mitigation approach for MANETs." In Innovative Security Solutions for Information Technology and Communications: 8th International Conference, SECITC 2015, Bucharest, Romania, June 11-12, 2015. Revised Selected Papers 8, pp. 171-183. Springer International Publishing, 2015.