ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 1, January 2022

Efficient Identification of DDoS Attacks on Cloud Computing Environment using Machine Learning Techniques

[1]Suzaifa, [2]Abdul Khader, and [3]Sareen Fathima

[1],[2],[3] Department of Computer Science and Engineering, Mangalore Institute of Technology, Badaga Mijar, Moodabidre, Mangalore, Karnataka, India

[1] *Corresponding author. Email: suzaifa.suzaifa@gmail.com,[2] khader_cse@bitmangalore.edu.in, [3] fathimasiddiq@ymail.com

Abstract*— The internet gives us beneficial data and information for personal, social, and economic growth. The security issue will be the main challenge on the internet. As the cloud technology is more popular, the first advantage of a cloud is flexibly scales to satisfy a variable demand, blew up instantly, blew down when demand moderates– in subsequent. So it requires expanse shielding from a DDoS attacks to challenge interlude effects of a DDoS Attacks. DDoS attacks are the malicious endeavor to interrupt normal traffic in a web property i.e., critical attacks that negotiates the accessibility of the network. On the cloud environment, aiming at efficiently analyzing and detecting the DDoS attacks, we propose four machine learning techniques i.e., the Naive Bayes, Support Vector Machine, K-nearest neighbor, and Random Forest.

Index Terms*— Efficient Identification, Cloud Computing Environment, Machine Learning Techniques, DDoS Attacks

I. INTRODUCTION

Cloud Computing Environment. Cloud computing is the platform which gives individuals to share data, provide facilities, and store data. It furnishes organizations with a springy structural design that delivers a productive outline for computing. By countless associations, the appropriation of the distributed computing condition, the tons of dangers, and difficulties have accompanied it [1]. Distributed computing servers have a value accessible on the web [2]. So, the problems like client protection, information spillage, and verification stay probably the greatest test to distributed computing conditions [10].

As the framework, and resource sharing of cloud computing architecture have introduced the viably countered challenges but, more concerns have been commendably disputed the framework [3]. The security of cloud computing includes several concerns mainly in access control information and network, cloud structure. it is not informal to impose entire security actions as there are diverse security anxieties of the diverse users [10].

DDoS Attack. Distributed Denial of Service attack, in short called DDoS attack [4]. Grounded on the application and network protocol, the DDoS attacks are classified. The ICMP, TCP, and UDP protocols do transport or network-level attacks. The HTTP DDoS bouts were executed taking place in low, and high amount situations each one of them partaking an outstanding influence on the victim. The victim concedes by slow, and requests which outcomes in an enervation of possessions or information in low rate situations while in high rate bout floods, a prey with the more amount of requests [10].

The DDoS bouts on cloud computing condition is primarily

an application layer that conveys requests ensuing the communication protocol. Which are then hard to distinguish the network layer because of their pattern suites a legitimate request. Thus, making the traditional defense systems improbable [10]. Request and session flooding bouts, asymmetric attacks, and slow response are the various categories of DDoS flooding attacks. All the flooding attacks generate traffic, which looks like an authentic user. which becomes harder for the objective to recognize amongst attack, and authentic traffic thus stalling the services of the authentic user.in a denial of service to a legal user in this kind of attack. To discover exposures, and accomplish flooding attacks by appeal methods, the DDoS bouts on the cloud are formed.

Machine Learning Techniques. Machine Learning techniques is the education of a processer algorithms that enhances spontaneously over event happened. Also, noticed as a subsection of artificial intelligence. Training data is machine learning procedures that make an arithmetic model built on selected information, in direction to create a prediction or resolution unaccompanied by overtly programmed. Machine learning systems encompass a broad range of solicitations, such as image processing and email

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 1, January 2022

straining, where it is laborious or impracticable to evolve standard algorithms to execute the required chores [6].

Here two forms of machine learning procedures. They have supervised learning and unsupervised learning systems. The Supervised learning technique is in which contains pre-existing labels and unsupervised learning technique is in which contains no pre-existing labels. Here we use supervised machine learning algorithms where we map an input to get output by analyzing the trained and tested data [10]. It is a tremendously accurate and reliable method. It also provides a strong tool to classify the data and assists to predict the results for unseen data [7].

Machine Learning techniques such as Naive Bayes, Support Vector Machine, K-nearest neighbor, and Random Forest are used for classification of the network they are supervised machine learning. Naïve Bayes algorithm is the modest probabilistic assortment deployed in smearing Bayes theorem. Support Vector Machine algorithm castoff classification and deterioration analysis. K-nearest neighbor algorithm is in which no parameter technique is castoff classification and regression. Random Forest is the collaborative learning system for the classification, deterioration, and additional chores.

II. RELATED WORK

Distributed Denial of Service Attack Detection using Machine Learning Systems in the Cloud Computing Environments. The underlying endowment protocols and technologies carry vulnerabilities and bugs which is exposed to doors for invasion by an invader. As Distributed Denial of Service are the utmost periodic that impose stern harm and disturb the cloud performance. Marwane Zekri. [1] developed C4.5 algorithms in 2017 which is adapted to perceive DDoS bouts, provides additional exact outcomes in assessment with the other machine learning systems.

Detection System of HTTP Distributed Denial of Service Attacks in a Cloud Environment Based on the Information Theoretic Entropy and Random Forest. Existing HTTP Distributed Denial of Service bout recognition structures is confronted by a large sum of web traffic created by the bouts, high false and low detection accurateness positive rates. Mohamed Idhammad. [5] in 2018 developed, to evaluate the entropy of a network heading structures of arriving web traffic by using a time-based sliding window system. When an estimated entropy outstrips its average the pre-processing and grouping tasks are triggered. Associated with sole classifiers that verified unswervingly on a dataset, the system achieved high-performance detection. The outcome of the proposed proposal achieves acceptable results with the running period of 18.5s and 99.54%, an FPR of 0.4% accuracy.

Analysis and Recognition of DDoS Attacks on Cloud Computing Environment using Machine Learning Methods. The Distributed Denial of Service bouts comes under the class of a dangerous bouts, that negotiates the accessibility of a network. These bouts have become an enlightened and endure to raise at quick leap so to identify and pledge the attacks has become a stimulating chores. Abdul Raoof Wani in 2019 [10] proposed various the machine learning system such as the Support Vector Machine, Naïve Bayes, and the Random Forest for the classification by generating a fresh dataset with Invasion Detection Method. The SVM algorithm shows the high performance of the dataset. The total accuracy of the Support Vector Machine is 99.7%, the Random Forest is 97.6%, and Naïve Bayes is 98.0%.

Analysis of Cybersecurity Threats in Cloud Applications using Deep Learning Methods. Cybersecurity coercions in the cloud environments is integrated enterprise applications since the fields of IoT and telecommunications. S. A. Sokolov, T. B [8] proposed in 2019 which is a method based on the Deep Neural system for an analysis of the cybersecurity extortions in a cloud application. The suggested system uses four neural classifiers for spam comments, unsolicited mail, images, and network traffic in which the attained outcomes are equivalent to immediate methods.

Network Intrusion Recognition using Supervised Machine Learning Method with Feature Selection. The intrusion discovery method that exists currently can identify the recognized attacks. Identifying zero-day bouts or new attacks remains a study theme due to more false-positive rate of current organizations. Kazi Abu Taher, Md. Mahbubur Rahman [9] in 2019 proposed while classifying network traffic, an Artificial Neural Network (ANN) based on machine learning through the wrapper feature selection outplay the support vector machine method. To estimate the performance, the NSL-KDD dataset castoff to a categorize web traffic to utilize ANN and support vector machine supervised learning algorithms. The perusal of an outcome illustrates that the prototypical erected using wrapper feature selection and artificial neural network outpaced all the other replicas in categorizing web traffic properly with a recognition rate of 94.02%.

A. Motivation

Cloud Computing services are frequently brought through HyperText Transfer Protocol. This enables admittance to reduces costs and services for the end-users and providers. Nevertheless, these growths a compulsion of the cloud services aspect to a HTTP

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 1, January 2022

DDoS bouts. The HTTP request approach is frequently castoff to generate several situations of a HTTP DDoS attacks such as Slow or Flooding attacks and Low attack and address web server vulnerabilities. Current HTTP DDoS recognition methods are disputed by a large amount of web traffic generated by the bouts, high false-positive rates, and low detection accuracy. Distinct machine learning procedures can be used to discover an attack but the Naive Bayes, support vector machine (SVM), K-nearest neighbor, Random Forest algorithms show an effectual application in the field of network security and performance.

III. METHODOLOGY

A. Naive Bayes

Naive Bayes mechanism on the ideology of provisional probability is given by the Bayes theorem P(R|S) = P (S|R) P (R) P (S).

The Bayes theorem gives a provisional probability of an event R given additional event S has ensued. Bayes theorem computes the provisional probability of the incidence of an event created on earlier information of circumstances that might be associated with the incident. Naive Bayes classifier is very modest and easy to contrivance and needs fewer training data.

As it's speedy, it can be castoff in a present prediction. Which is extremely accessible with several analysts and a data point. Naive Bayes is not sensitive to inapplicable features. Handles together continuous and discrete information.

Bayes theorem conditions a subsequent bond, given class variable b and reliant on feature vector b_1 through a_p:

$$P(b|a_1\ldots\ldots,a_r) = \frac{p(b)p(a1,\ldots\ldots,ar|b)}{p(a1,\ldots\ldots,ar)}$$

By means of the naive provisional independence statement that

$$P(a_1|b, a_1\ldots\ldots,a_{q-1}, a_{q+1},\ldots,a_r) = p(a_q|b),$$

for all q, this relationship is shortened to

$$P(b|a_1,\ldots\ldots,a_r) = \frac{P(b)\prod_{q=1}^{r} P(aq|b)}{P(a1,\ldots\ldots,ar)}$$

Since P (a_1…….a_r) is perpetual given the input, we can use the subsequent classification decree:

$$P(b|a_1,\ldots\ldots,a_r)\alpha\, P(b)\prod_{q=1}^{r} P(aq|b)$$

$$\Downarrow$$

$$b=argmax_b P(b)\prod_{q=1}^{r} P(aq|b)$$

and we can use Maximum A Posteriori (MAP) approximation to appraisal P(b) and P(a_i|b); the previous is the comparative frequency of class b in a training set. Diverse naive Bayes classifiers vary mostly by the conventions they create a concerning distribution of P(a_i|b). Instead of their seemingly over-simplified conventions, naive Bayes classifiers are operated pretty well in numerous real-world circumstances, well document classification and spam filtering. They need a lesser number of training data to appraisal the required constraints.

Naive Bayes trainees and classifiers can be tremendously wild related to extra erudite approaches. The decoupling of the class provisional feature distributions means that each dispersal can be individualistically projected as a one-dimensional distribution. This is shot aids to improve the problems stopping from the curse of dimensionality.

On another side, while naive Bayes is recognized as an attired classifier, it is recognized to be an immoral estimator.

B. Support Vector Machine

Support Vector Machine is the lone of a grade machine learning systems used in the anomaly network intrusion recognition, spam straining, and pattern recognition.

All other cases can be removed without altering it. Classes are not linearly separated. Boundary depends on very few points. The execution of SVM universally substitutes all lost standards and converts normal attributes into a binary one.

It also standardizes all the attributes by default. The factors in the production are not as native data. This is important for an interpreting a classifier, which is based on the normalized data.

SVM can discover the comprehensive optimal result by accomplishment of linear separation verdict an optimal hyper-plane that parts into 2 classes. A neighbouring data to the hyper-plane are the support vectors. Receiving the features, the predicted class is acknowledged.

To distinct the 2 classes of data themes, several likely hyperplanes might be selected. An impartial, to discover the plane which is the maximum border, i.e., a maximum distance amongst data points of the both classes. Exploiting the boundary distance affords some strength so that upcoming data points can be classified with additional sureness.

Figure 1 represents the explanation of SVM

$x=b+\sum\alpha_i y_i(i).a$ where $\sum\alpha_i y_i a(i).a$ is a sum over support vector $a(i)$.

C. The drawbacks of support vector machines comprise:

The amount of features is abundant bigger than the number of samples, evade over-fitting by choosing Kernel functions and regularization period is critical.

SVMs unreliably offer probability appraisals, these are intended by means of utilizing fivefold cross-validation.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 1, January 2022

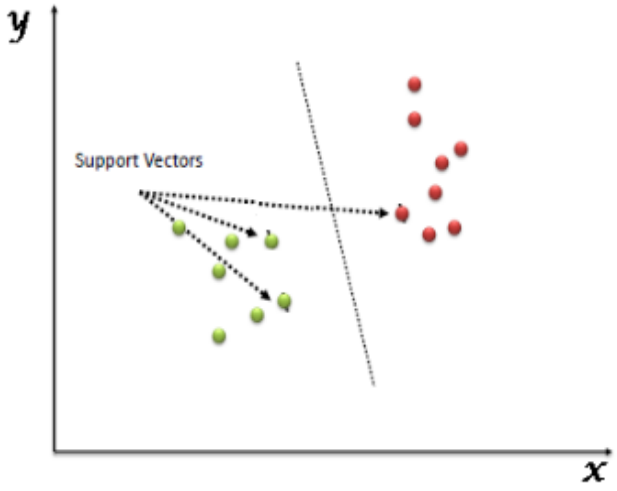

Fig. 1 Example of SVM

D. K-nearest neighbor

Knn is the data-centric approach, that reserves all the accessible cases and classifies the new data or case based on a harmony measure. It is used for classification because that is the primary area.

The algorithm is multipurpose. We select a value of a k i.e.; sum of a nearest neighbors, where we select the least distance. The least distance measure by Euclidean distance i.e., square root of a sum of the difference amongst the new point x and existent point y.

KNN algorithm accepts the alike things happen in near immediacy. In further disputes, identical things are closed to each other.

Selecting a right value for K

To choose the K particular for a data, run a KNN procedure numerous periods with dissimilar values of a K and select a K that decreases the number of faults that encounter while upholding an algorithm's capability to precisely make predictions when it is assumed data it has not seen earlier.

A Drawback of KNN is

The algorithm gets knowingly slower as several instances and/or predictors/sovereign variables rise.

E. Random Forest

The random forest algorithm is made out of decision trees. They are informal to practice, informal to build, and informal to understand.

Trees have one feature that stops them from being the ideal tools for predictive learning. Random Forests combine the simplicity of decision trees with flexibility which results in a vast improvement in accuracy. Prototypes are intended to give data about a relative amongst variables and classification.

Random Forest algorithm calculates immediacies amongst twosomes of cases which can be used in a grouping, finding outliers or provides stimulating outlooks of data. The proficiencies of overhead can be prolonged to unlabelled data, outlier detection, foremost to data views, and unsupervised clustering. It agreements a new technique of discovering variable connections. It turns competently on huge databases.

Random Forest algorithm can handle thousands of input variable deprived of a variable removal. It gives appraisals of what variables are significant in a classification. Also produces a core dispassionate appraisal of a generalization fault as a forest building progresses.

Random Forest algorithm has an active technique for appraising lost data and upholds accuracy when a huge proportion of data is lost. Random Forest has approached, complementary error in a class population unstable datasets. The created forest can be protected for the upcoming use of the other data.

F. Remarks of Random Forest

Random forests underfit. we can track numerous trees as we need. The random forest is reckless. Running on the data set with 500,000 cases and 1000 variables, formed 1000 trees in 11 minutes on an 800Mhz engine. For the huge data sets, the main memory necessity of storing the data itself, and 3 integer arrays with alike dimensions as data. If proximities are considered, stowage requirements produce as the number of cases times the sum of a tree.

G. Each tree is grown as follows:

If a sum of cases in a training set is R, sample R cases at a random, but with a spare, from a new data. This model will be training set for increasing a tree.

If near are S input variables, the number s<<S is specified such that, at the individual node s variables are selected at random out of the S and the best divided on this s is used to split the node. The value of an s is detained constant throughout a forest growing.

For each tree is grown to the biggest extent likely. There will be no clipping.

H. Performing DDoS Attack using Torshammer

The DDoS attack is powerful and operated this command in Kali Linux Operating System. At first, we check for the website for port number and domain name or IP address.

We use tool Torshammer which is inscribed in python which can give harm to insecure web servers in an occurrence that's why it is a sluggish post tool. Here we use cloud computing services. The foremost feature of this tool is, it works in the application layer.

Sample of DDoS attack is shown below in figure 2.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 1, January 2022

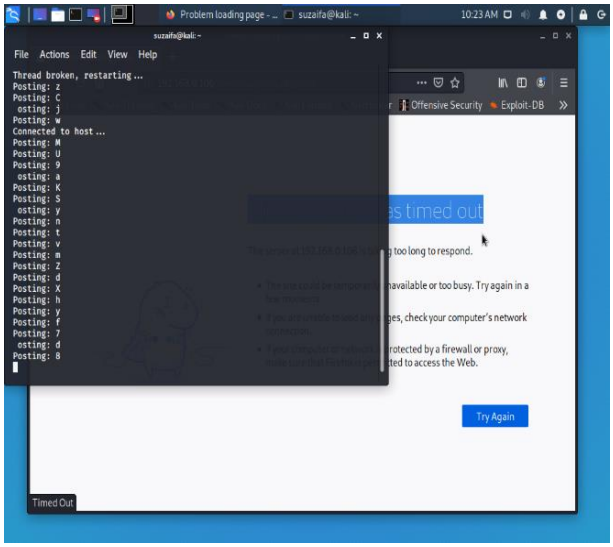

Fig 2. Sample of DDoS Attack

IV. CALCULATION ON PERFORMANCE PARAMETER

To Categorize the traffic network, whether it is suspicious, normal, or unknown we do this approach of checking parameter performance.

	Predicted Class POSITIVE (Suspicious)	Predicted Class NEGATIVE (Unknown)	Predicted Class NEGATIVE (Normal)
Actual Class POSITIVE (Suspicious)	TP	FN	
Actual Class NEGATIVE (Unknown)	FP	TN	
Actual Class NEGATIVE (Normal)			

Fig 3. Confusion matrix

Figure 3. above represents the confusion matrix. In which it contains three predicted classes and three actual classes. The predicted classes contain positive suspicious, negative unknown, negative normal and the actual classes contain positive suspicious, negative unknown, negative normal.

The **False Negative(FN)** is combined with predicted negative unknown, predicted negative normal and actual positive suspicious.

The **True Positive(TP)** is combined with predicted positive suspicious and actual positive suspicious.

The **True Negative(TN)** is combined with predicted negative unknown, predicted negative normal, actual negative unknown and actual negative normal.

The **False Positive(FP)** is combined with predicted positive suspicious, actual negative unknown and actual negative normal [11].

Confusion Matrix

Recall = (TP)/(TP+FN)

Precision = (TP)/(TP+FP)

Accuracy = (TP+TN)/(TP+TN) + (FP+FN)

Specificity = (TN)/(TN+FP)

F measure = (2TP)/(2TP+FP) = (FN)

System Design

A. Attack Generation

The DDoS bouts were accomplished in a cloud environment. The DDoS bout was engendered in a safe environment using attacking or the Torshammer tool. The figure 4 below shows the DDoS Attack on Cloud [12].

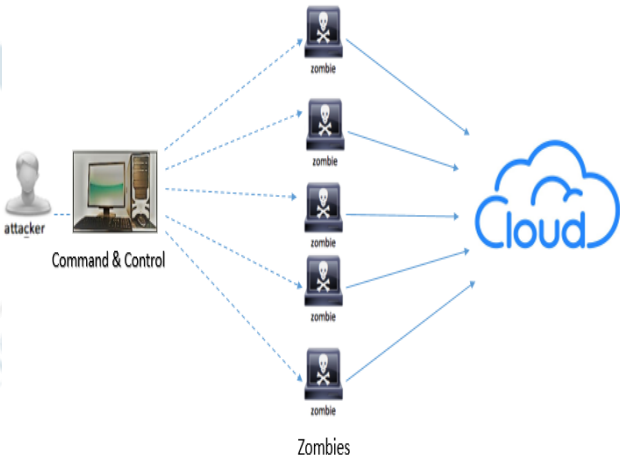

Fig 4. Depicts the DDoS Attack on Cloud

B. Sequence diagram

A sequence diagram demonstrates various members and associations linking them by the positioning of messages. Grouping outline exhibits a combination of an outline with an on-screen nature similar to a utilization case yet focusses additional on the exchange of messages amongst the associates. The objects and Classes are tangled in the sequence diagram. The sequence diagram is in which tangled in a picture of a sequence of messages switched amongst the entities needed [13].

The figure 5. is shown below for sequence diagram.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 1, January 2022

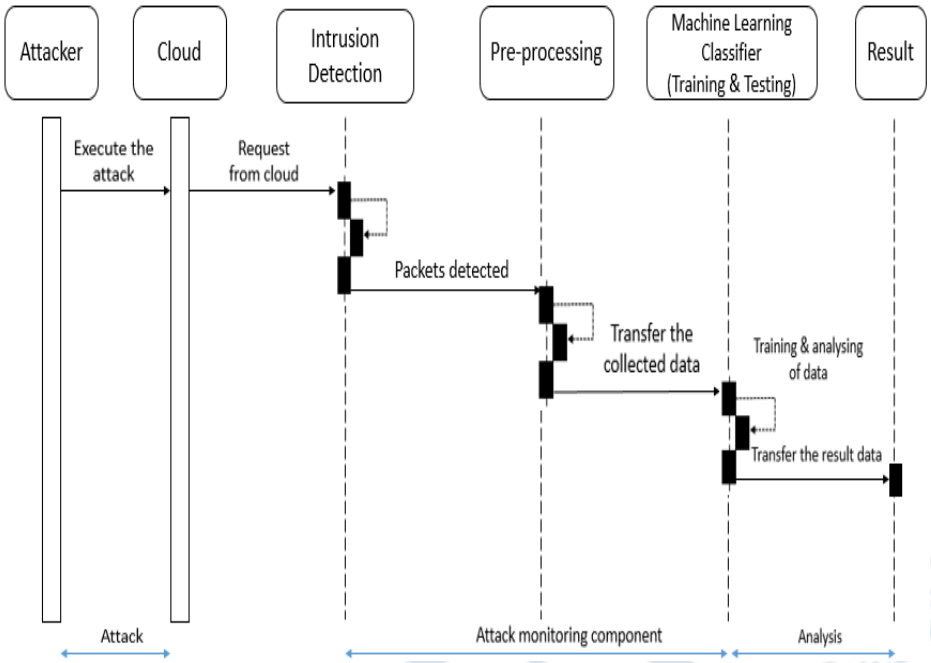

Fig 5. Sequence Diagram

C. Flow of the System

The information rivulet framework is evident between the utmost crucial presenting equipment. It is employed to illustrate the outline sections. The framework segments include the outline procedure, component that confederates with the outline, the details utilized, and the data streams in the outline.

The flow diagram specifies how the data explores the outline and how it is accustomed by an evolution. It is the graphical strategy that exemplifies a data stream and the changes are associated as data transfers from a benefaction to produce.

In the project flow of the project starts from the attack i.e., the attack is generated by the attacker to the cloud.

Interrupt by the network traffic when the HTTP request is done. so there will be an HTTP attack

Attack is detected by the network analyzer Wireshark.

Dataset is created using HTTP packet filter

Dataset is trained in the Weka tool.

After training, the testing is done using cross-validation 12-folds.

Then the result is classified for better performance.

Figure 6. of a flow diagram shown below

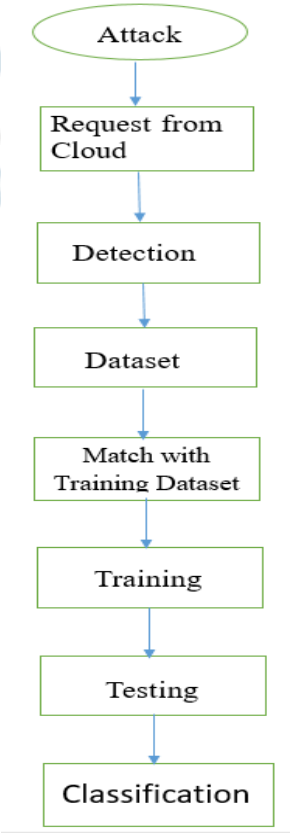

Fig 6. Flow diagram

D. Experimental Setup

The DDoS attack is achieved using the torshammer tool in the Kali Linux 2020.1 on OwnCloud in Ubuntu 19.4. Screenshot of OwnCloud page is shown below figure 6. The attack is recorded by the network analyser, Wireshark in the attacking system. The database

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 1, January 2022

generated by Wireshark using packet filter i.e., HTTP packet is transferred to the server. Weka data mining utility tool which is an open-source used for testing and training of the two dataset i.e., large and small dataset. Large dataset is the csv file contains more number of instances and small dataset is the csv file contains less number of cases. Both the dataset is used for the classification of the machine learning algorithms such as the Naive Bayes, Support Vector Machine, K-nearest neighbor, and Random Forest machine learning algorithm. The testing is obtained using cross-validation 12 folds. Sample of dataset is shown in below figure 7.

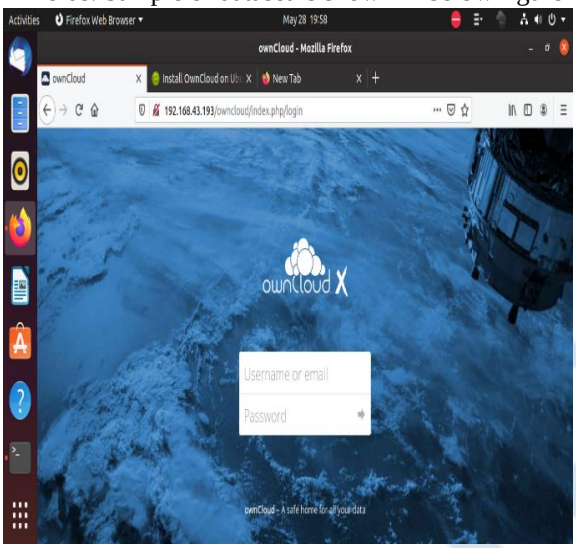

Fig 6. Screenshot of OwnCloud page

No.	Time	Source	Destinatic	Protocol	Length	Info	Class
75	32.12846	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
91	32.76132	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
98	33.13591	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
112	33.46241	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
120	33.84825	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
143	34.28921	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
153	35.17181	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
156	35.20332	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
178	35.65257	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
198	36.52611	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
209	37.20032	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
234	37.55348	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
236	37.61575	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
261	38.50928	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
282	38.59112	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
296	39.1393	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
313	39.63881	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
315	39.6389	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
328	39.76099	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious
353	43.10095	192.168.0.	192.168.0.	HTTP	690	HTTP/1.1	Suspicious

Fig 7. Sample of dataset

V. RESULTS

The confusion matrix is used to assess the classifiers and the outcomes is presented in below two tables for large dataset and small dataset

.

Table 1. Large dataset

	Naive Bayes	Support Vector Machine	K-nearest neighbor	Random Forest
Recall	0.999	0.999	1	0.999
Precision	0.998	0.999	1	0.999
Accuracy	0.9987	0.9993	0.9996	0.9993
Specificity	0.857	0.857	0.714	0.714
F measure	0.999	0.999	1	0.999

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 1, January 2022

Table 2. SMALL DATASET

	Naive Bayes	Support Vector Machine	K-nearest neighbor	Random Forest
Recall	0.971	0.985	0.971	0.971
Precision	0.968	0.971	0.957	0.957
Accuracy	0.9705	0.9852	0.9705	0.9705
Specificity	1	0.75	0.5	0.5
F measure	0.967	0.978	0.962	0.962

The accuracy of Naive Bayes is 99.87%, Support Vector Machine is 99.93%, K-nearest neighbor is 99.96% and Random Forest is 0.99.93%, when we do the pre-processing of large set of data. Hence for large dataset, the K-nearest neighbor algorithm have high accuracy.

The accuracy of Naive Bayes is 97.05%, Support Vector Machine is 98.52%, K-nearest neighbor is 97.05% and Random Forest is 95.58%, when we do the pre-processing of small set of data. Hence for small dataset, the Support Vector Machine have high accuracy. The recall, precision, specificity, and F-measure are correspondingly significant since the imbalanced data and should be put into deliberation. The DDoS attack is known from the timestamp and information given by Wireshark during packet filtering.

VI. CONCLUSION

After The work is conceded out on a OwnCloud environment using Torshammer, DDoS attacking tool. Two datasets were used i.e., a large and small dataset, which is obtained from the network analyser, which is pre-processed in the Weka tool for the classification of four algorithms i.e., Naive Bayes, Support Vector Machine, K-nearest neighbor, and Random Forest. For large dataset, the K-nearest neighbor algorithm gives high performance with an accuracy of 99.96% and for small dataset the Support Vector Machine gives high performance with an accuracy of 98.52%.

The result shows that, all the algorithms used for a large dataset is more accurate compare to a small dataset. Hence when a large number of a dataset is incorporated, the performance is high. As OwnCloud uses Content Delivery Network, which protects against the DDoS attack, hence there will be no effect on OwnCloud due to attack. Further work can use different types of attacks, other tools for attack, and different feature selection techniques.

REFERENCES

[1]. Marwane Zekri, Said El Kafhali, Noureddine Aboutabit, and Youssef Saadi. "DDoS Attack Detection using Machine Learning Techniques in Cloud Computing Environments." IEEE, 2017.

[2]. Xiaoyong Yuan, Chuanhuang Li, Xiaolin Li. "DeepDefense: Identifying DDoS Attack via Deep Learning." IEEE, 2017.

[3]. Chenxu Wang, Tony T. N. Miu, Xiapu Lu, and Jinhe Wang. "SkyShield: A Sketch-based Defense System Against Application Layer DDoS Attacks." IEEE, 2017.

[4]. Bahman Rashidi, Carol Fung, and Elisa Bertino. "A Collaborative DDoS Defence Framework Using Network Function Virtualization." IEEE,2017.

[5]. Mohamed Idhammad, Karim Afdel, and Mustapha Belouch. "Detection System of HTTP DDoS Attacks in a Cloud Environment Based on Information Theoretic Entropy and Random Forest." Security and Communication Network. Hindawi. 2018.

[6]. Anteneh Girma, Mosses Garuba, and Rajini Goe. "Advanced Machine Language Approach to Detect DDoS Attack Using DBSCAN Clustering Technology with Entropy." Information Technology – New Generations, Advances in Intelligent Systems and Computing 558. Springer. 2018.

[7]. Jieren Cheng, Mengyang Li, Xiangyan Tang, Victor S. Sheng, Yifu Liu, and Wei Guo. "Flow Correlation Degree Optimization Driven Random Forest for Detecting DDoS Attacks in Cloud Computing." Security and Communication Networks. Hindawi. 2018.

[8]. S. A. Sokolov, T. B. Iliev and I. S. Stoyanov. "Analysis of Cybersecurity Threats in Cloud Applications Using Deep Learning Techniques." University of Telecom-munications and Posts, Department of Information technology. MIPRO. 2019

[9]. S. A. Sokolov, T. B. Iliev and I. S. Stoyanov. "Analysis of Cybersecurity Threats in Cloud Applications Using Deep Learning Techniques." University of Telecom-munications and Posts, Department of Information technology. MIPRO. 2019.

[10]. Kazi Abu Taher, Billal Mohammed Yasin Jisan, Md. Mahbubur Rahman. "Net- work Intrusion Detection using Supervised Machine Learning Technique with Feature Selection." International Conference on

Robotics, Electrical and Signal Processing Techniques. 2019.

[11]. Abdul Raoof Wani, Q.P. Rana, U. Saxena, Nitin Pandey, "Analysis and Detection of DDoS Attacks on Cloud Computing Environment using Machine Learning Techniques" IEEE, 2019.

[12]. Hongbin Luo, Member, Zhe Chen, Jiawei Li, Athanasios V. Vasilakos, Senior Member, "Preventing Distributed Denial-of-Service Flooding Attacks with Dynamic Path Identifiers" IEEE, 2017.

[13]. Rahul Chourasiya, Vaibhav Patel , Anurag Shrivastava, "Classification of Cyber Attack using Machine Learning Technique at Microsoft Azure Cloud", IRJES, 2018.