NNACS360 : A Neural Network Algorthm to perform Behavioral Analysis for Cloud Secure 360

Thiruchendhil Arasu, Senior Director, QwikCilver Solutions Pvt. Ltd. Platform Quality Engineering, Bangalore, India. tcarasu@gmail.com

Dr. E. George Dharma Prakash Raj, School of Computer Science and Engineering, Bharathidasan University, Trichy, India. georgeprakashraj@yahoo.com

Murali Krishna, Data Scientist, DELL – Bangalore, Inda. murali.aakas@gmail.com

Abstract. Now a days computing had transformed into a different model that provide services that can be delivered in mannered utilities such as water, power and network. In such well-defined model, users are allowed to access services that are based on their requirements regardless of where the services are hosted.

There are huge benefits that is offer to the organizations which prefer this cloud computing technology, but with their equal benefits comes the associated risk factors related to security. DDOS attack is one mere threat for cloud computing environments which bothers end user and lead them to huge financial loss.

This paper talks on the cloud security enhancement through deep learning Neural Network techniques. Some parameters that we have considered like inside and outside cloud attack with scaled user rating remains the same from our previous papers but the implementation through deep learning mechanism paves way for easy computing with higher accuracy in predicting the magnitude of DDOS attack. The non-linearity behavior prevailing with the data overcomes proposed Neural Network Algorithm.

Determining DDOS attack characteristics based on these parameters are complex, time consuming and costly with tradition multi-variate causal model. Therefore, in order to eliminate these disadvantages and complexities, the prediction of DDOS attack can be modeled using deep learning algorithm like neural network with automated feature selection method. Artificial Neural Network algorithm works like a biological neuron which receives input from different sources and handles nonlinear operations within data and predicts the final output with best fit and higher accuracy.

***Keywords. Cloud Computing; Data Privacy; Security; Virtualization; Deep Learning; Predictive Analytics*.**

I. INTRODUCTION

Cloud-based services are ideal for businesses with trending or fluctuating bandwidth demands. If your needs increase it's easy to scale up your cloud capabilities, drawing on the service's remote servers. Likewise, if you need to scale down again, the flexibility is baked into the service. This level of agility can give businesses using cloud computing a real advantage over competitors. With the introduction of numerous cloud based services and geographically dispersed cloud service providers, the risk associated with respect to attacks like DDOS had increased which incur huge financial loss to the entity which is adapted to this computing mode.

Cloud Secure 360 is equipped with an advanced machine learning capability that can identify threats based on the user behavior over internet usage and their pattern is classified based on the geography attached to that user. This service is expected to provide critical behavior based profile information to the ISP's, threat monitoring tools on the cloud and organizations that do business in the internet etc.

The implementation model for Cloud Secure 360 is based on subscription where subscribers will be proactively informed on a regular interval on the behavior based profile data accessed from a particular area within a city or the next lowest level of Dynamic Host Configuration Protocol (DHCP) server.

 DOI : 10.18231/2454-9150.2019.0320 © 2019, IJREAM All Rights Reserved.

Cloud Secure 360's success depends on the continuous collaboration with the ISP's on the internet user data in terms of the unique customer ID and the originating IP address to profile the behavior patterns at an area level and subsequently nailed down to the individuals who are trying to cause disruption to the Content Security Policy (CSP) operations in the form of Denial of Service (DoS) or Distributed Denial of Service (DDoS).

Key Question points in this paper are 1.How to predict magnitude of DDOS attack with higher accuracy using deep learning mechanism 2. How to incorporate the nonlinear pattern associated to the existing data with our prediction results. 3. Handling the huge volume of data inflow using deep learning algorithms

This paper is organized as follows. Section II gives a report on the various related work done similarly to this work. This is followed by the already proposed works in Section III followed by the Extended proposed work in Section IV. The experimentation and analysis is given in Section V and Section VI concludes the paper.

II. Related Work

There has been good amount of related work done in the area of behavior based user profiling for the cloud security. A rank clustering system, CloudRank [1], is proposed by Sakyajit et al that takes into account cloud user preference data to characterize cloud user behavior and also identify groups of users with similar behavior in an unsupervised manner. The user groups are determined based on fitting mixture models on the cloud user preference observations. A preference can be anything that a system designer would like to include to characterize high-level user requirements such as demands on performance, cost, security, availability, etc. CloudRank can be useful for: (i) cloud providers to target their service offerings according to the user groups through appropriate customization of services pertaining to the user groups typical requirements; (ii) recommendation systems or a marketplace to determine which offerings best suit certain user groups; and (iii) prediction of any new users behavior based on their preference information.

Li-Jun-Jain et. al. in their paper [2] have proposed a dynamic trust evaluate method to deal with cloud user's behavior. using Entropy method to reflect the essential regular pattern of user's behavior evidence, making the evaluate way become a dynamic model, weaken the subjectivity of simply using and Analytic Hierarchy Process (AHP), moreover, still need AHP to make the result fit people's subjective experience. Hence, this paper discusses on the integrate algorithm that combine Entropy Method and AHP, in this way, the final evaluate value will keep the balance between objective and subjective and provide quantitative analysis foundation for security control. The analysis shows that the dynamic trust evaluate method can effectively distinguish user's abnormal behavior.

The paper from Xiaoming Ye et. al. proposes an anomalous behavior detection model [3] based on cloud computing. Virtual Machines (VMs) are one of the key components of cloud Infrastructure as a Service (IaaS). The security of such VMs is critical to IaaS security. This research into VM security issues, especially regarding VM network traffic anomalous behavior detection, remains inadequate. This paper proposes a model that uses Software-Defined Networks (SDN) to implement traffic redirection. The model can capture inter-VM traffic, detect known and unknown anomalous network behaviors, adopt hybrid techniques to analyze VM network behaviors, and control network systems. The experimental results indicate that the effectiveness of this approach is greater than 90%, and prove the feasibility of the model. This allows to ensure that the system is running in the expected environment, the monitoring probes have not been tampered with, and the integrity of measurement data provided is maintained. Overall this gives a basis for increased confidence in the security of running parts of the system in an external cloud-based environment.

The paper from Xin Lu et. al. addresses the issue of credibility authentication of user behaviors in the cloud computing environment. The Paper proposes a user behavior credibility authentication model [4] built on the characteristic-based random Petri network to assess the behavior contract credibility of the users accessing the cloud service resources. This model first makes dimensional normalization of the behavioral residue data of the users and uses the decision tree ID3 algorithm to characterize the behavioral residue data of the users to check such data against the behavior authentication sets, so as to determine the credibility of the compliance of the user behaviors with the contract. User behaviors are dynamic and random, so this paper proposes the status deduction function of the random Petri network to analyze the credibility of the compliance of user behaviors with the contract. Then the credible degree is calculated to make quantitative assessment of the user behaviors' credibility. The simulation experiments indicate that this model is able to reliably assess user behaviors' credibility in the cloud computing environment and is a certain improvement in terms of accuracy and efficiency of credibility authentication compared with the traditional models.

Insider attack is the most devastating threat due to the familiarity of the underlying system to the insiders. The proposed approach by Mahesh Babuet. al. mitigates this threat by a host based user profiling technique [5] where a key stroke dynamics is used for analyzing the user behavior and a retraining approach is also proposed as the imposter patterns are absent at the time of registration. Nowadays, securing the transmitted data is the most important challenging areas of development and research

 DOI : 10.18231/2454-9150.2019.0320 © 2019, IJREAM All Rights Reserved.

in modern communication. Users are able to communicate over an insecure channel using cryptography, so an attacker cannot decrypt and understand the original message. Public key cryptography requires large computational power, huge time consumption and complexity. An Artificial Neural Network (ANN) is used in order to overcome these problems.Insider threats still remain as one of the major concerns. Lucky Nkosiet. al discusses an approach that can help in identifying insiders behaving in a malicious way, which may lead to an attack. A rule learning algorithm [6] was used in learning the behavior pattern of users, in order to build user profiles. A Matching algorithm was then used to match the historical behavior of the user with the current behavior, in order to identify users that masquerade in the system as normal users. The obtained results show that it was possible to identify insiders that masquerade in the system by observing their behavior patterns.

Ngugi, Benjamin and Beverly K. Kahn[7] proposed that the Behavioral biometrics, like biometric typing patterns, have the potential to make another level of security to cloud but this research identified some deficiencies in performance quality. Two research streams for improvements have emerged. The first approach attempts to improve performance by building better classifiers, while the second attempts to attain the same goal by using richer identifying inputs. Both streams assume that the typing biometric patterns are stable over time. This study investigates the validity of this assumption by analyzing how students' typing patterns behave is considered as one particular parameter for enhancing security layer.

Pin Shen Teh,1 Andrew BengJin Teoh,2,3 and Shigang Yue[8] proposed that keystroke dynamics refers to the process of measuring and assessing human's typing rhythm on digital devices. Such device, to name a few, usually refers to a computer keyboard, mobile phone, or touch screen panel. A form of digital footprint is created upon human interaction with these devices. These signatures are believed to be rich in cognitive qualities, which is fairly unique to each individual and holds huge potential as personal identifier. Existing network security prediction methods for the cloud environment are limited in terms of both accuracy and real-time performance. Many papers have been addressing these issues with a proposal for a method based on grey neural network to predict network security situations in cloud environments. First, we explore security factors for network security situation awareness based on classification and fusion techniques in order to generate awareness indexes. Through this, we establish a hierarchical index system for network security situation. Then, a method is elaborated that combines grey theory and neural networks to predict network security situations by analyzing the features of grey and neural networks that combine high accuracy and real-time performance.

Lee, K., Caverlee, J., And Webb, S. [9] proposed that spammers, content polluters, and malware disseminators can be easily identified using a honeypot-based approach in online social systems. The core objective here is to fix a social honeypots for harvesting deceptive spam profiles from social networking communities; and applying Statistical analysis of the properties of these spam profiles for creating spam classifiers to actively filter out existing and new spammers. The harvested spam data that is considered for analysis contains signals that are strongly correlated with observable profile features like content, friend information, posting patterns, hours visited and frequency staying in websites.

Miller, E. [10] proposed that Average rating works fine if the data always have a ton of ratings, but if a term has 2 positives and 0 negatives and suppose the other term has 100 positive ratings and 1 negative rating. This algorithm puts item two (tons of positive ratings) below item one (very few positive ratings) which is statistically not 100 percent but we need to balance the proportion of positive ratings with the uncertainty of a small number of observations so that the rating system can be improvised with scaled weights.

Khoi KhacNguyen ,Dinh Thai Hoang , DusitNiyato , Ping Wang , Diep Nguyen , and ErykDutkiewicz [11] had proposed a preventive approach to detect and isolate cyber threats before they can cause serious impacts to the mobile cloud computing system. In this paper, they had proposed a novel framework that leverages a deep learning approach to detect cyberattacks in mobile cloud environment. The use of ANN can identify attackswhen rules are not known. Patterns are recognized and recent actions happened with the usual behaviour are compared by a neural network approach,also, NN is adapted to certain constraints,in order to resolve many issues even without human intervention. Misuse are consistently detected by neural networks,also, the recognition of malicious events are improved. This makes the system enhancing flexibility against intrusions in order to be able to protect their entire organization

JürgenSchmidhuber, The Swiss AI Lab IDSIA, IstitutoDalleMolle di Studisull'IntelligenzaArtificiale, University of Lugano & SUPSI, Galleria 2, 6928 Manno-Lugano, Switzerland [12] had proposed that neuron is a non-linear transformation of a linear sum of inputs: $y = f(wT x + b)$ .An array of neurons taking the same input x form a new layer on top of the input in a neural network: $y = f(WT x + b)$ which contributes in treating non linearity within data adding dynamic weights to layers and node predicting with higher accuracy.

III. PROPOSED WORK

The aim of the this paper is to predict the magnitude of the DDoS attack that can occur in a DHCP server when the DHCP is servicing customers with different patterns of

 DOI : 10.18231/2454-9150.2019.0320 © 2019, IJREAM All Rights Reserved.

behaviors. Behaviors are defined for individuals based on their online activity depending on what kind of websites they visit. Scores are generated for customers based on the websites they visit and in turn, the scores are generated for DHCP servers based on the scores of customers that they service. The following Assumptions are considered in this research model.

Assumption 1: The customer ID of each individual will be attached with their IP address. The IP address keeps changing but the customer ID remains the same. This assumption is critical in our model because we want to be able to trace the hacker based on a single unique factor. This can be achieved by collaborating with the ISPs and getting the Customer IDs linked to the users.

Assumption 2: There are individual scores/ratings for the websites that are present on the World Wide Web. The websites are classified/ranked based on how safe or friendly they are from an information security standpoint. For example, a website like www.khanacademy.com will have a rating of 9 on a scale of 1-10 and a google search result for "How to hack a network" will have a rating of 2. Hackers will not be as explicit/specific in their online activity as mentioned but this approach gives a solid starting point to mitigate the risk that hackers pose based on the hacker's past behavior and online activity.

An individual upon visiting any website is transferred a rating of the visited website's rating. In the above example, the individual has, let's say, visited a series of websites such as www.khanacademy.com, www.amazon.com and a google search of "learn statistics online" having ratings of 9, 7 and 8.5 respectively, then the individual's rating will be an average of the ratings of the 3 websites, 9+7+8.5 = 24.5 = 24.5/3 = 8.167.

Ratings for individuals will act as a profile for the person based on his/her behavior and will help in classifying the different individuals into safe/neutral/risk categories. If the individual rating is let's say, above 5, then the person can be viewed as a non-risk inducing person and <5 will be a risk-inducing person. As and when a website is visited by a person, the rating of the website gets attached to the rating of the person and a new average for the person is generated.

Note : If the person visits a website that has a rating of less than 3, which means that the person has visited a website that has potentially harmful information, then the average for the rating is calculated only with the ratings of that person that are less than 3 and the previous (if any) ratings > 3 are ignored.

For example – a person has visited khan academy, amazon, a google search result of "learn statistics" and a google search result of "how to initiate a DOS attack" then the assumption is that a single 'bad' score/website is enough to classify his/her intentions as risk-inducing. For this example, his rating instead of being, 9 + 7 + 8.5 + 2.5 = 27/4 = 6.75; will be just 2.5 which is the rating of the "unsafe" website which is the google search result of "how to initiate a DDOS attack". From this point onwards, the rating of this person will not take into account the websites he visits that have rating >3 but only those that have rating <3.

Table 1 Sample dataset for Website rating

Websites	Safety Rating
www.facebook.com	7
www.twitter.com	8
www.instagram.com	9.2
www.amazon.in	6.7
Google search result for "places in Bangalore to eat"	7.4
Google search result for "How to initiate a DOS attack"	0.5
Google search result for "Learn statistics online"	9.3
Google search result for "Hacking tutorials"	0.9
www.khanacademy.cm	9

Table 2 Individual 1's Behavioral score

Websites	Website Score	Individual's score
www.facebook.com	7	0.5+0.9 = 1.4/2 = **0.7** Since the individual has visited unsafe websites (rating<3), we take into account only unsafe websites from this point henceforth in categorizing their behavior.
www.twitter.com	8	
www.instagram.com	9.2	
www.amazon.in	6.7	
Google search result for "places in Bangalore to eat"	7.4	
Google search result for "How to initiate a DOS attack"	0.5	
Google search result for "Learn statistics online"	9.3	
Google search result for "Hacking tutorials"	0.9	

Table 3 Individual 2's Behavioral score

Websites	Website Scores	Individual's Score
www.facebook.com	7	7 + 8 + 9.2 + 6.7 + 7.4 + 9.3 = 47.6/6 = 7.93
www.twitter.com	8	
www.instagram.com	9.2	
www.amazon.in	6.7	
Google search result for "places in Bangalore to eat"	7.4	
Google search result for "Learn statistics online"	9.3	

 DOI : 10.18231/2454-9150.2019.0320 © 2019, IJREAM All Rights Reserved.

Since a way to rate individuals is devised, rating of DHCP servers that are servicing these individuals is considered next. The way this will be done is by averaging the ratings of individuals that the DHCP server is currently servicing and assigning this average rating to the DHCP server thereby giving a rating for each DHCP Server. A probable DHCP rating would look like what is given in Table 4.

Table 4 DHCP Server score

Individual Ratings through DHCP server	DHCP Rating
6	32.55/8 = 4.07
8	
0.2	
0.5	
1.3	
9.3	
6.3	
0.95	

Since the assumptions are known, the modelling exercise can be done. Consider Table 5 to be the data for the magnitude of DDOS attacks and the DHCP scores of the DHCP servers at the time when the DDOS attacks took place.

Table 5. DHCP score vs DDOS

DHCP score	Magnitude of DDOS attacks
6.257	7
6.45	8
6.257	6.5
8.33	3
8.33	4
6.44	6
6.44	8
3.35	9.4
3.45	9.1
4.3	8.2
1.3	2
3.6	3
5.5	5
4.5	7
3.9	9
1	10
5	5
7	3
9	0.3
9.2	2
8	3
4.34	8.2

The idea is to fit a Cloud Security 360 model to this data and come to a prediction/formula that fits this data the best and can be used as a predictor for the future work based on the historical data.

The model that has been chosen is the linear regression model. This is due to two reasons. Firstly, a supervised learning methodology is used here since the existing/historical labelled data is already available. The data available is to be analyzedand produce an inferred function which can be used for mapping future values. Secondly,our aim is to get a mathematical relation between the DHCP scores and the magnitude of DDOS attacks on a continuous scale instead of a categorical scale.

IV. Extended Work : Neural Network Algorthm

The above sections give us a view of what can be achieved in the field of cloud security with the help of machine learning. With the advancements in the field of ML and with multiple learning algorithms in the picture, it is essential that we understand the pros and cons of each algorithm and choose the algorithm that best fits our needs.

Table 6 gives the pros and cons of the different classes of algorithms. The classes areRegression, Classification and Clustering. From each class there are examples taken and they are analyzed.

Table 6 Pros and Cons

S. No	Model	Pros	Cons
1	Linear Regression	Easy interpretability, minimal tuning, fast processing	Data set often has non-linear relationships. LR cannot capture complex relationships within the data
2	Naïve Bayes	Similarity classification	Does not work on small data set. Does not work with interdependent columns
3	K Means	Similarity classification	Needs to tweaked by a human consistently
4	Support Vector Machines	Both classification and regression	
5	Deep Learning – Neural Networks	Classifies text/audio/image very easily. Captures complex relationships within the data	Require large amounts of data. Require large computation capacity.

Our initial approach involved a multi variate causal model with factors that contributes to both inside and outside attacks of cloud. With an additional parameter of user visiting a website and the rating is captured using a scaled algorithm. But the complexity that exists in building a prediction model with multi variate regression is that the

 DOI : 10.18231/2454-9150.2019.0320 © 2019, IJREAM All Rights Reserved.

nonlinear relationships are not captured with high importance and computing. In addition to that nonlinear data applied with regression model requires an infinite number of possible functions and is more difficult to setup.

In future our data set is expected to expand exponentially both in length and breadth. To service the growing number of users who spend time online, more DHCP servers need to be installed. It could also be the case that our predictor variable – 'the online behavior of a user' does not completely predict the magnitude of DDOS attacks. In case of extending parameters or to establish more valid relationship with greater accuracy deep learning algorithms should be engagedby which we get a better understanding of each parameters and their impact on DDOS attack. These deep learning algorithms help in managing huge volumes of data very easily and interconnection patterns and updating weights of these interconnections happens much faster with generating appropriate output activation function

Based on the given constraints and needs, the Deep Learning Neural Network algorithms give us the best shot at achieving our solution. Deep Learning Neural Network algorithms are very useful for training and predicting on data sets that are very large and contain intricate relationships between the attributes. Deep Learning Neural Network is also very useful for working with data that is in the form of videos, images and text. The Pseudo Code for the Neural Network Algorithm is given below :

Pseudo Code for Neural Network Algorithm :

X = input values of both inside and outside cloud attacks with scaled user rating
Y = output values of DDOS attack scores
repeat{
for (i in 1:length(layer)){
for(j in 1:i){
1. weighted sum calculated for all inputs to node
2. add threshold and activation function is calculated
}
}
for(k in 1:length(output_layer)){
1. Calculate Error
}
for(l in 1:length(hidden_layers)){
for(j in 1:i){
1. node error rate calculated
2. adjust node weights
}
}
Error calculated
}

while(max(iteration)<user_defined) AND (Error function >user_defined)

The main reason for replacing a traditional model with Deep Learning Neural Network algorithm is because the only way to optimize a linearly separable model is to add third degree polynomials to their coefficients but in that way we indicate some assumptions about the data by defining objective function's structure. But in case of Neural network we create input layer that creates the linear separators for the data and hidden layer ANDs the regions that bounds some classes and last layer ORs all these regions. So in that case data should adopt a methodology that can incorporate a nonlinear way and should learn weights internally with defined functions. One greater advantage of Neural Network is that adding multiple features as plug in can be handled easily with greater accuracy whereas in case of regression it is opposed to "Curse of dimensionality". So going with Neural Network is the best feasible option for targeting higher accuracy.

V. Experimentation and Analysis

The experimentation is done in R programming language using linear regression.Figure 1 shows the Plot of the data points for the Linear Regression model.

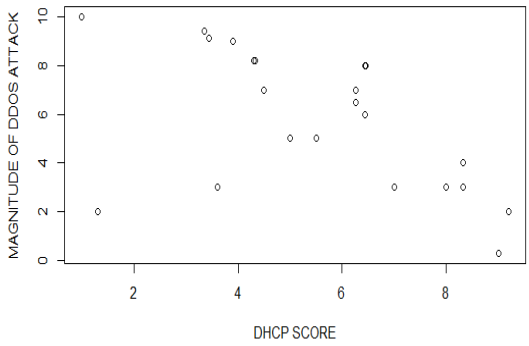

Fig. 1 Plot of the data points

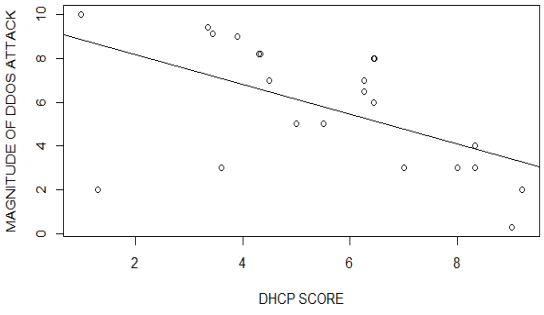

Fig. 2 Fit of prediction line

 DOI : 10.18231/2454-9150.2019.0320 © 2019, IJREAM All Rights Reserved.

Figure 2 shows the Fit of prediction line through the data points where 'x' represents DHCP scores and 'y' represents Magnitude of DDOS attacks.

From Fig 3 below, the summary of the linear regression function is found as

$$Y = -0.6785x + 9.5198$$

With R-squared as 0.2987.

The DHCP scores will be changing in real-time depending on the customers/users they would be servicing. With this linear regression equation the DHCP scores can be fed in real-time and check the predicted magnitude of DDOS attacks that can occur.

The previous paper focus on same set of parameters for predicting the magnitude of DDOS attack but with a multi variate causal model.The equation of the causal model that was run is Riskfactor = 0.105 + (4.3367)*keystrokespeed + (-0.0311)*perfdem + (-3.8008)*costdem + (-0.300974)*securitydem + (0.450001)*availabilitydem + (5.789111)*credibilityscore + (2.45678001)*packetstransfered + 0.452(scaled_user_rating)

With 88 percent accuracy and the Root Mean Square value comes around 23.89 which shows that there is huge variance existing between the actual test data and predicted test data. This huge variance exists because of nonlinear relationship that exists between data. Adding non parametric functions as weights to reduce the variance level is complex using a multi variate regression model. So a Deep Learning Neural Network mechanism handling high volumes of data and adjusting to non-linearity functions should be implemented.

The existing paper uses same parameters like inside and outside attack that happens in cloud environment and user behavior scaled rating which is also considered in our previous paper as impacting factors on DDOS attack but with an implementation change in deep learning mechanism, the followung is obtained.

Feature Selection:

The parameters which we had considered for our analysis capture hidden threats to a cloud environment but with the vast number of features and volume that we get it is hard to build model with higher accuracy. Some factors may be an added noise to a model and some features may be really important for our analysis. So it is required to undergo a feature selection process which can help us prioritizing factors that speaks more about our target variable rather considering all parameters for building model.

In the existing paper we had trained our data with random forest variable selection method. This black box supervised learning method with no assumptions that target variable has a linear relationship with predictor variables try creating a partial dependence plot. For every tree that is created in this random forest process, the prediction accuracy on the out-of-bag portion of all features is recorded. Then the same is done after permuting each impacting factors. The difference between the two accuracies are then averaged across all trees, and normalized by the standard error term. The Mean Squared Error is computed on the out-of-bag data for each tree, and then the same computed after permuting a feature. The differences are averaged and normalized by the standard error.

Based on this feature selection method the variable importance are captured and only features that contribute to maximum variance on the target variable is chosen as final neural network inputs. Figure 3 represents Variance Importance Plot and Figure 4 represents the variance importance table of all the features. Based on the maximum variance contributed the features are selected as input layers for neural network.Figure 5 shows the Neural Network Plot.

Features	Variance importance
packets_transferred	4.7255087
keystroke_speed	2.5771542
perf_dem	2.489199
Credibility_score	1.2124263
security_dem	0.6105834
cost_dem	0.3849314
availability_dem	-1.7620965
scaled_user_rating	-2.6760372

Fig. 3 Variance Importance Plot

 DOI : 10.18231/2454-9150.2019.0320 © 2019, IJREAM All Rights Reserved.

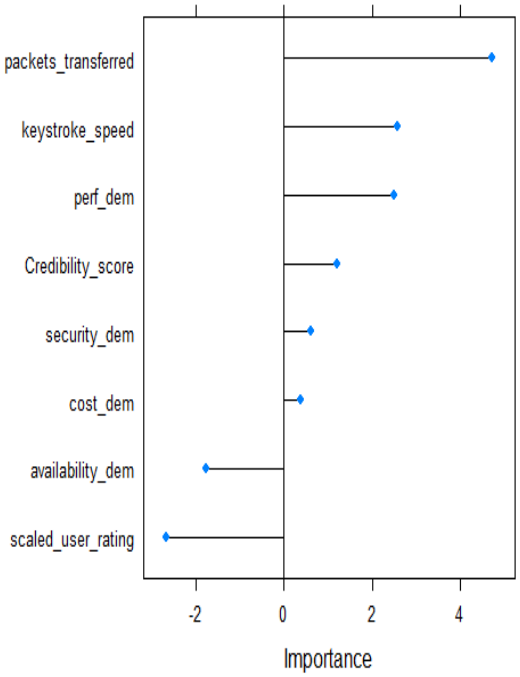

Fig. 4 Variance Importance Table

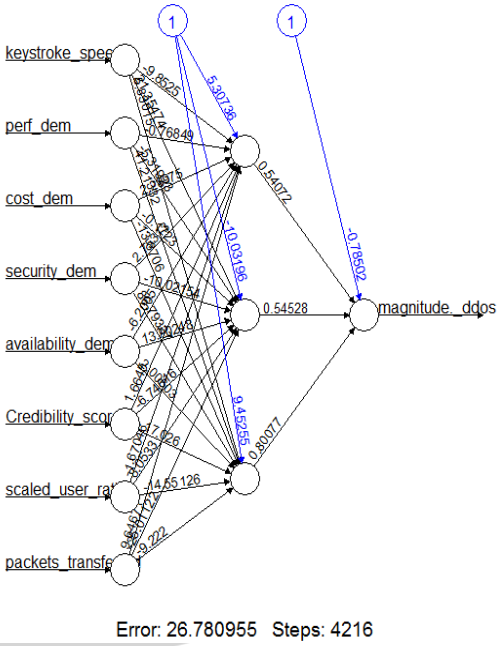

Fig 5 Neural Network Plot

The sample code for finding the Root Mean Square is given below :

```
# Random sampling
samplesize = 0.60 * nrow(data)
set.seed(80)
index = sample( seq_len ( nrow ( data ) ), size =
samplesize )

# Create training and test set
datatrain = data[ index, ]
datatest = data[ -index, ]
max = apply(data , 2 , max)
min = apply(data, 2 , min)
scaled = as.data.frame(scale(data, center = min, scale =
max - min))
trainNN = scaled[index , ]
testNN = scaled[-index , ]
set.seed(2)
NN = neuralnet(magnitude._ddos ~ keystroke_speed +
perf_dem + cost_dem + security_dem + availability_dem
+ Credibility_score
        + scaled_user_rating + packets_transferred,
trainNN, hidden = 3 ,linear.output = T )

plot(NN)
## Prediction using neural network
predict_testNN = compute(NN, testNN[,c(1:8)])
predict_testNN = (predict_testNN$net.result *
(max(data$magnitude._ddos) -
min(data$magnitude._ddos))) +
min(data$magnitude._ddos)
# Calculate Root Mean Square Error (RMSE)
RMSE.NN = (sum((datatest$magnitude._ddos -
predict_testNN)^2) / nrow(datatest)) ^ 0.5
```

The RMSE value calculated on the test data is 17.456 which contributes the greatest accuracy and based on the back ward propagation neural network the weights of all nodes at different layers are calculated and cost function happens in iterative way to reduce the error term between the actual and fitted data points thus enhancing high accuracy to a model.

V. CONCLUSION

NNACS algorithm is proposed in this chapter to enhance security using Neural Network Model to perform user behavior analysis. NNACS is proposed by considering all possible security threats – user behavior, inside and outside cloud attacks parameters in same model. Here, automated feature selection and variable weight assigning with neural nets to give high accuracy. Additionally, handling huge volumes of data and interpretation with specific to respective applications in cloud environment. NNACS gives us a quick understanding of how predicting magnitude of DDOS attack and attributes contributing to risk can help us tackle this DDOS attack problem. The accuracy of this current model had reduced the variance level to greater extent by handling the non-linearity characteristics that exist in data. The RMSE error had reduced from 23.89 to 17.456 by further increasing the accuracy with 6.2 percentage more contributing to an overall accuracy of 94.2 percentage.

REFERENCES

[1] Sakyajit Bhattacharya, Tridib Mukherjee, and KoustuvDasgupta, "CloudRank: A Statistical Modelling Framework for characterizing user behaviour towards targeted Cloud Management" IEEE Network Operations and Management Symposium, 2014

 DOI : 10.18231/2454-9150.2019.0320 © 2019, IJREAM All Rights Reserved.

[2] LI Jun-Jian, Li-Qin, "User's Behavior Trust Evaluate Algorithm Based On Cloud Model" IEEE Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control, 2015

[3] Xiaoming Ye, Xingshu Chen, Haizhou Wang, Xuemei Zeng, Guolin Shao, Xueyuan Yin, and Chun Xu, "An Anomalous Behavior Detection Model in Cloud Computing"Special Issue On Information Security ,Volume 21, Number 3, June 2016

[4] Xin Lu, Cheng du, China; Yue Xu, Cheng du, China "An User Behavior Credibility Authentication Modelin Cloud Computing Environment". IEEE International Conference on Information Technology and Electronic Commerce, 2014.

[5] Mahesh Babu, Mary SairaBhanu, "Analyzing User Behavior Using Key Stroke Dynamics to Protect Cloud from Malicious Insiders" IEEE International Conference on Cloud Computing in Emerging Markets, 2014

[6] Lucky Nkosi, Paul Tarwireyi Mathew O Adigun, "Detecting a Malicious Insider in the Cloud Environment Using Sequential Rule Mining", IEEE International Conference on Adaptive Science and Technology, 2014

[7] Ngugi, Benjamin, Beverly K. Kahn, and Marilyn Tremaine. "Typing biometrics: impact of human learning on performance quality. " Journal of Data and Information Quality (JDIQ) 2. 2 (2011): 11

[8] Teh, Pin Shen, Andrew BengJin Teoh, and Shigang Yue. "A survey of keystroke dynamics biometrics. " The Scientific World Journal 2013 (2013).

[9] LEE, K., CAVERLEE, J., AND WEBB, S. Uncovering social spammers: social honeypots + machine learning. In ACM SIGIR: Proceeding of the international conference on Research and development in Information Retrieval (2010)

[10] Miller, E. 2006. How not to sort by average rating. how-not-to-sort-by-average-rating.html.

[11] Khoi Khac Nguyen , Dinh Thai Hoang , DusitNiyato , Ping Wang , Diep Nguyen , and ErykDutkiewicz "Cyberattack Detection in Mobile Cloud Computing: A Deep Learning Approach" , arXiv:1712.05914v1 [cs.CR] 16 Dec 2017

[12] JürgenSchmidhuber, The Swiss AI Lab IDSIA, IstitutoDalleMolle di Studisull'IntelligenzaArtificiale, University of Lugano & SUPSI, Galleria 2, 6928 Manno-Lugano, Switzerland" Deep learning in neural networks: An overview"

 DOI : 10.18231/2454-9150.2019.0320 © 2019, IJREAM All Rights Reserved.