IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-6, June 2018 E-ISSN: 2347-2693

Scaled User Rating Algorithm to Perform Behavioral Analysis for Cloud Secure 360

Thiruchendhil Arasu[1], E. George Dharma Prakash Raj[2], Murali Krishna[3]

AVA Digital Solutions–Bangalore, India.
School of Computer Science & Engineering, Bharathidasan University–Trichy, India.
Digital DELL – Bangalore, India.

Email : tcarasu@gmail.com, georgeprakashraj@yahoo.com, murali.aakas@gmail.com Tel.: +91-9686477617

Available online at: www.ijcseonline.org

Accepted: 06/Jun/2018, Published: 30/Jun/2018

Abstract. Cloud industry has reached a critical mass in the past few years, with many cloud service providers fielding competing services. Despite the competition, some of the security mechanisms offered by the services to be similar, indicating that the cloud industry has established several "best-practices," while other security mechanisms vary widely, indicating that there is also still room for innovation and experimentation. The cloud industry had grown in the fast few years, with so many providers focusing on cloud services. Besides huge competition the security mechanisms that is provided by all these vendors had shown many good practices but that paves a way for many new innovative experiments. This paper mainly focusses on improving the security mechanism against DDOS attacks. With the existing system we are not able to predict the magnitude of DDOS attack as the causes vary across different situation. So, resolving this security issue becomes much more complex in real time situation. One important reason for DDOS attacks can be because of fake users creating spoofed request. Apart from that there are also additional attacks which are made within cloud environment and outside cloud environment, so security mechanisms must be tightened. There is also some hidden pattern which prevails on user surfing through websites based on their frequency and content visited which is also required to establish furthermore security based on user behavior. The aim of the paper is to predict the magnitude of DDOS attack which is bonded with a two-fold solution 1. Capturing trust rating for a user visiting a website considering his frequency and website safety ranking based on a Scaled User Rating Algorithm. 2. Considering parameters that helps in figuring DDOS attack pattern based on both internal and external attacks within the cloud environment. The aims defined in this paper help us in figuring out a malicious behavior of user based on his surfing pattern and contents that he had referred, which in turn help us in expecting a possible DDOS attack. In addition to that we are also trying to find possible parameters that could be a reason for DDOS attack analyzing threats that had happened within and across cloud environments in the past.

Keywords. Cloud Computing; Data Privacy; Data Protection; Security; Virtualization; Monitoring; Deep Learning; Predictive Analytics; Scaled User Rating

I. Introduction

Cloud computing is a demanding technology in the present era. It is an environment which describes web world as a massive space that is pre-installed in computing which exists as a service that is readily available in a web environment to be shared. With the end users cloud computing is provided on a Pay-per-Use-On-Demand mode that helps them in conveniently accessing shared IT resources via the Internet.

As it gains more popularity with its flexible characteristics there is also a mere threat on the opposite edge which make user feel unsecured with these web environments. Some attacks like DDOS is still a biggest threat to the end user in completely trusting this flexible web-based computing environment.

Cloud Secure 360 is an engine that is designed to protect the cloud environment from these DDOS attacks. It acts as a warning or alarm system that can predict the DDOS attack pre hand. So that loss that is supposed to happen in future because of DDOS attack for organizations can be prevented. 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). It also tries capturing user behavior when visiting a website and frequency of their visits to malicious content is also captured in a continuous time interval. Cloud Secure 360 is enabled with advanced machine learning capability which can make inferences with the continuous flowing data. It can help us in bifurcating between good and risk users based on behavior in online streams. It also helps in predicting the possible risk factors within and across cloud environments considering wide range of cases that had occurred in the past in both inside and outside cloud environments.

Key Question points in this paper are:1. On how much or what % of risks can be mitigated through this approach. 2.Are there other solutions from Trusted Third Party (TTP) on the cloud that can be integrated to the Cloud Secure 360 to reduce the risk % and improve the overall customer experience and the brand image of the organization?3. What parameters really contribute to DDOS attack in a cloud environment? 4. Can risk factors be identified from both internal and external cloud environments? 5. How can website rating be scaled based on risk parameters?

Most possible factors that can contribute for a DDOS attack is captured and which can reveal us hidden pattern that exists in these attacks. With the pattern and attributes that is figured it is easy to predict the DDOS attack in a specific cloud environment. Predicting these attacks pre-hand can help the organization to plan their security measures at appropriate time which help them saving huge financial loss. This paper proposes a Scaled User Rating Algorithm to perform Behavioral Analysis for Cloud Secure 360

This paper is organized as follows. Section I contains the introduction of Cloud Computing. Section II contains a report on the various related work done similarly to this work, Section III contains the Proposed work, Section IV contains the Extended proposed work, Section V contains the experimentation and analysis and Section VI contains the Conclusion.

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 usergroups; 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 can 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 Babu et. 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 approaches also proposed as the imposter patterns are absent at the time of registration. One observation made in cloud is that most of theadministration work involves command line interface rather than graphical user interface. Since the command line interface requires a lot of key strokes, the proposed approach is well suitable for this environment. If the abnormality in the user behavior is detected, the system is locked so that a malicious masquerader cannot do any modification in the name of others.

Insider threats remain as one of the major concerns. Threats from malicious insiders are often listed as dangerous threats by many researchers. However, this threat has not received the attention it deserves because many organizations turn out to be extra careful about external threats than insider threats. Lucky Nosier et. 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, to build user profiles. A Matching algorithm was then used to match the historical behavior of the user with the current behavior, 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 parameter for enhancing security layer.

Pin Shen Teh, 1 Andrew Beng Jin 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 unique to each individual and holds huge potential as personal identifier.

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 honeypot 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.

R.Piplode, P. Sharma and U.K. Singh. [10] proposed the study of threats, risks and challenges involved in cloud computing. Detailed description of security issues involved in cloud computing with root cause analysis of web application, accessibility and authentication of those respondent device vulnerabilities are discussed.

P. Rutravigneshwaran. [11] proposed that machine learning methods can help in learning the malicious pattern involved in intrusion detection system. Anomaly detection system creates a database of normal behavior and any deviations from the normal behavior occurred are stored in Misuse Detection system database, an alert is triggered regarding the occurrence of intrusions and if there is a change with normal behavior. These patterns are identified and trained to machine learning model which helps in classifying the threats that is likely to happen.

III. Proposed Work

The aim of 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 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 everyone 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 consider the websites he visits that have rating >3 but only those that have rating <3.

A sample data set for website ratings is considered and analyzed. A Sample dataset for website rating is given in Table 1.

Table 1 Sample data set 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.com	9
www.encyclopedia.com	8.3
www.youtube.com	7

Table 2 Individual 1'sbehavioral 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 3's behavioural 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	
Websites	**Website Scores**	

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 analyzed and 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: Scaled User Rating Algorithm

The methods discussed above in this paper and in the papers referenced, provide solutions/methods to protect the cloud environment from attacks both inside and outside the cloud. The methods above are extremely merit-worthy on their own, but one drawback that all suffer from is that these methods solve only one problem at a time instead of having a check or mechanism that acts like a warning system for all the types of attacks on the cloud – both from inside and from outside. This is the gap that this paper also aims to fill. The Pseudo Code for the Behavioral Analysis and Scaled User Rating is given below.

Pseudo Code for Behavioral Analysis and Scaled User Rating:

X = input values of user behavior data and scaled user rating
Y = magnitude of DDOS attack
Step1: Collect parameters that contribute to inside and outside attack based on user behavior
Step2: Collect the individual user DHCP score and frequency of their visit to a website
Step3: Scale user rating algorithm
For (i in 1: length(users)) {

1. Fetch internal and external website visits with DHCP score and frequency
2. If (DHCP_score< 2 & frequency > 1){

user_DHCP_rating<- $Pp+10(1-P)\ (1-e^{-f/Q})$)
#p = website rating
#f = Frequency of user visiting website
#Q = importance/weight you attach to the notion "Frequency"

}
user_DHCP_rating<- DHCP_score(normal average)
}
Step4: Build multi variate causal model and predict magnitude of DDOS attack based on user behavior scaled rating with internal and external cloud attack parameters

All cloud service providers are focused on a common goal – increase the coverage of their security and

reduce the risk for their consumers against attacks on the cloud environment. CSPs will be benefitted by a mechanism that alerts them of the possibility of multiple forms of attacks rather than one. A comparison of the different forms of internal cloud attacks is performed using a multiple variate causal model.

The different forms of attacks that we are focusing on are using – user contract behavior, user's storage preference data, user behavior patterns inside the cloud and anomalous behavior patterns inside the cloud. The parameters used are keystroke speed, demands on performance, cost, security, availability, compliance contract score/credibility score and packets transferred. Using these parameters, the multiple regression predicts the Risk Factor for different combinations of the 7 parameters. The Risk Factor will tell the CSPs that their environment is, for example, 50% at risk based on the certain combination of the parameters.

Some more parameters with respect to the users are also collected when visiting a website such as Age, Employment history, Region and the time zone that they work. Because inappropriate access timing of website can also help us in deriving a malicious website visit because most of the spoof hackers prefer inappropriate work timings which can also help us in predicting security issues with higher accuracy.
Some parameters with respect to user behavior are :

- Keystroke speed
- perf_dem
- cost_dem,
- security_dem
- availability_dem
- Credibility_score
- packets_transferred

In addition to the existing approach a dynamic way of assigning weights to user net score value is implemented. For instance, every website has different rating mechanism and becomes complex in scaling up those values to a normal range. In such case take the website rating (weight as weighted mean), divide those values by half of it [this is to make the rating scale between common range. For example, if the scale should be defined between [0, 5]. Add the value which is made into half with $5(1-e^{-f})$. So, the formula becomes DHCP score = $5p/10+5(1-e^{-f/Q})$
Where
p = website rating
f = Frequency of user visiting website
Q = appropriate number that shows what importance/weight you attach to the notion "Frequency"

For example : A user visiting a website has 3 times a revision score of 6 and 2 times a revision score of 7. Then p= (3.6+2.7)/5=6.4 if we take Q=10 then $5(1-e^{-5/10}\approx 3.88)$ so the total score is 3.2+3.9=7.1 approximately 7
On the other hand, if somebody has 20 scorings of 6 then p=6 and $5(1-e^{-20/10})$ ≈4.58 so the final score is 3+4.6 rounded giving 8.
The choice of Q depends on what you call "few", "moderate", "many". As a rule of thumb consider a value M that you consider "moderate" and take Q=−M/ln (1/2) ≈1.44M. So, if you think 100 is a moderate value the take Q=144. Finally, you can also replace the equal weight on quantity and quality by a skewed one so that the final formula becomes:
DHCP score =Pp+10(1−P) $(1-e^{-f/Q})$)
Where P∈ [0, 1] (in the original formula we had P=0.5).

The equation of the causal model that was run is
Risk factor = 0.105 + (4.3367)*keystroke speed + (-0.0311)*perfdem + (-3.8008)*costdem + (-0.300974)*securitydem + (0.450001)*availabilitydem + (5.789111)*credibilityscore + (2.45678001)*packetstransfered + 0.452(Scaled_User_Rating)

V. Experimentation and Analysis

The experimentation is done in R programming language using causal machine learning model for DDOS attack and risk factor calculation. Fig 1 shows the Plot if the data points.

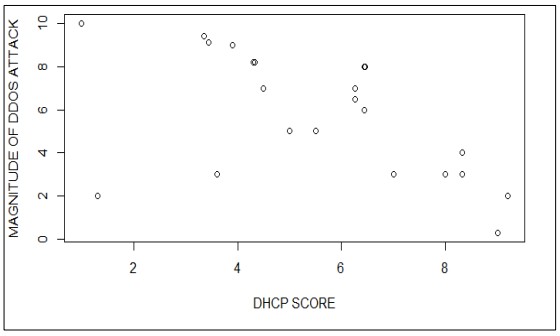

Fig. 1 Plot of the Data Points

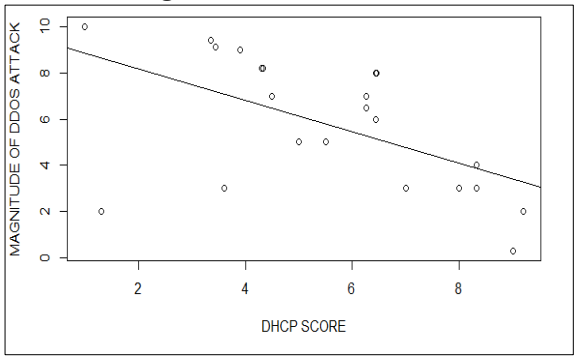

Fig. 2 Fit of Prediction Line

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 DHCP rating calculated with our previous papers does not consider the scaling impact. In order to establish a scaled score between the DHCP rating a normalized rating system which incorporate user frequency of visiting a website is implemented using an scaling equation

$$Scaled\_User\_Rating = Pp+10(1-P)\ (1-e^{-f/Q}))$$

With tool (p * average) + 10 * (1 - p) * (1 - ((Math::E)**(-number_scores / q))) When I set q = 5

Some parameters are collected which contributes to both inside and outside attack in cloud. Some variables has direct impact on the level of DDOS attack. In order to predict the risk factor involved in DDOS attack a predictive model is engaged. Here in our approach we build a multi variate regression model including all possible parameters we had collected.

```
> regression <- lm(ddos_mag ~ . , data = train_data)
> summary(regression)

Call:
lm(formula = ddos_mag ~ ., data = train_data)

Residuals:
     Min       1Q   Median       3Q      Max
-18863.7   -481.6     97.5    690.3  15781.1

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         2.844e+03  7.638e+02   3.724 0.000225 ***
keystroke_speed     9.287e-02  2.092e-03  44.382  < 2e-16 ***
perf_dem           -4.218e+01  5.849e+01  -0.721 0.471267
cost_dem           -6.812e+01  5.712e+01  -1.193 0.233709
security_dem       -1.945e-01  1.738e-02 -11.191  < 2e-16 ***
availability_dem   -8.436e+01  5.738e+01  -1.470 0.142324
Credibility_score   3.741e+01  5.872e+01   0.637 0.524416
scaled_user_rating  9.088e-01  2.951e-02  30.797  < 2e-16 ***
packets_transferred -4.313e+00  5.363e+00  -0.804 0.421803
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2917 on 391 degrees of freedom
Multiple R-squared:  0.8883,    Adjusted R-squared:  0.886
F-statistic: 388.6 on 8 and 391 DF,  p-value: < 2.2e-16
```

Fig. 3 Regression Result – R Tool

The results Figure. 3 thrown by causal model shows that keystroke speed, security_dem and Scaled_User_Rating shows high significance on magnitude of ddos attack. Their p value shows that these variables are highly significant which contributes to adjusted R squared relationship of 88 percent.

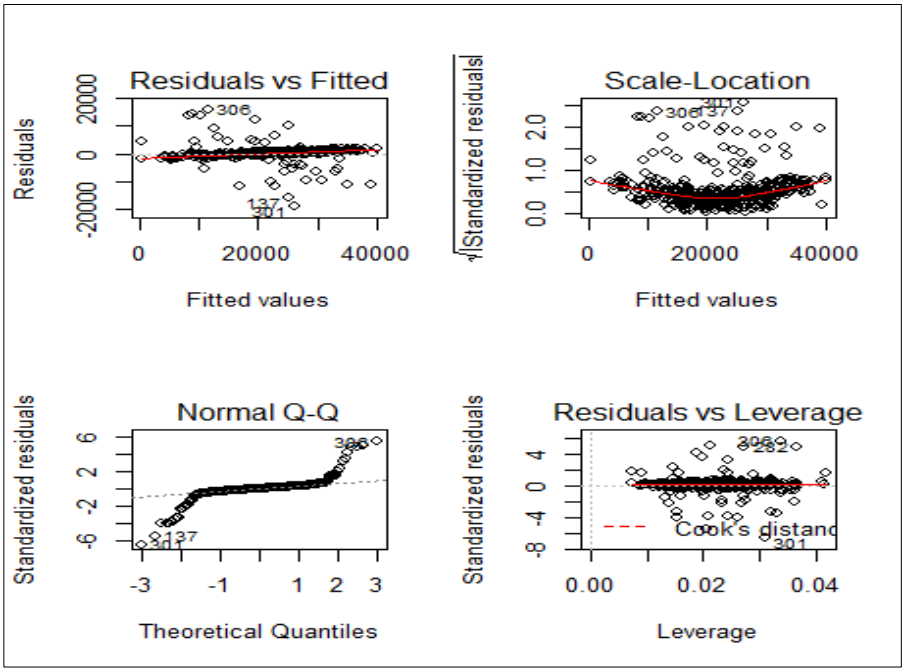

Fig. 4 Diagnostics Plot – R Tool

The residual plot shows that there is some nonlinear relationship between predictor variable and outcome variable because for a linear relationship between data there should be a pattern with equally spread residuals around a horizontal line without distinctive patters but in here all residual points are correlated with horizontal line. The scale location graph provides an equal residual spread across the range of predictors which satisfies the assumption of equal variance (homoscedasticity). There are some points evident from cook's distance that can impact the prediction results but as every point on our case is important those points cannot be removed as outlier. So, prediction must be taken care including those points. The RMSE value calculated comes around 23.89 which tells the difference between the actual test data and predicted test set.

VI. Conclusion

This paper titled " Scaled User Rating Algorithm to perform Behavioral Analysis for Cloud Secure 360" focuses on the Behavioral Analysis based on historical online activity that can provide the user with a method to mitigate risk when it comes to DDOS attacks and prevent unnecessary damage to CSPs across the globe. It has also been seen that cloud threats can occur through different medium, so in this paper, multiple parameters like keystroke speed, security level, credibility score and packet transferred ratio through network are included to access the risk associated with different internal forms of attacks that could happen in the cloud. With the proposed algorithm the variance captured is less and in

turn contributes to 88 percent accuracy after including the Scaled User Rating algorithm with the user behavioral attributes. For future enhancements, Deep learning methods can be implemented to discovers intricate structure in large data sets. By using the back-propagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. In addition to its existing feature automated feature selection components can be attached to select parameters that contribute finding high significance on calculating the magnitude of DDOS attack.

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.

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

[3] Xiaoming Ye, Xingshu Chen, Haizhou Wang, XuemeiZeng, 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 KeyStroke Dynamicsto Protect Cloud from Malicious Insiders" IEEE International Conference on Cloud Computing in Emerging Markets, 2014

[6] Lucky Nkosi, Paul TarwireyiMathew O Adigun, "Detecting a Malicious Insider in the CloudEnvironment 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] R.Piplode, P. Sharma and U.K. Singh, "Study of Threats, Risk and Challenges in Cloud Computing", International Journal of Scientific Research in Computer Science and Engineering, Vol.1, Issue.1, pp.26-30, 2013

[11] P. Rutravigneshwaran, "A Study of Intrusion Detection System using Efficient Data Mining Techniques", International Journal of Scientific Research in Network Security and Communication, Vol.5, Issue.6, pp.5-8, 2017

Authors Profile

Thiruchendhil Arasu is a self-motivated, people focused, innovative, Strategic & Transformational complementing Business through the right technology solutions that impacts the Quality, Productivity & Cost. 24 Years' of IT experience with strong experience in Collaborative Innovation, DevOps, Predictive Analytics/Machine Learning, Quality Engineering, Performance Engineering, E-commerce [Dell.com] Performance testing in Production, Test Center of Excellence (Test CoE), Strategic & Operations Delivery, Infrastructure Management, Enterprise Supply Chain, Technology Pre-Sales, Application Development & Support, and Database Architecture & Administration. Led full stack E2E Performance engineering for the Dell Supply Chain starting from Dell Commerce Platform [Dell.com] to Finance [Procure to Pay] which included post sales support application portfolio [CRM] Integrated Dell.com, legacy EMC and legacy Dell into a single Dell IT Performance engineering organization. Program

management for INNOVATION across PAN INDIA through Cultural Shift, Capability and Collaborative transformation resulting in increased Patents.

Dr .E. George Dharma Prakash Raj completed his Masters Degree in Computer Science and Masters of Philosophy in Computer Science in the years 1990 and 1998. He has also completed his Doctorate in Computer Science in the year 2008. He has around twenty eight years of Academic experience and twenty years of Research experience in the field of Computer Science. Currently he is working as an Asst. Professor in the School of Computer Science, Engineering and Applications at Bharathidasan University, Tiruchirappalli, India. He is an Editorial Board Member, Reviewer and International Programme Committee Member in many International Journals and Conferences. He has published several papers in International Journals and Conferences related to Computer Science He has convened many National and International Conferences related to Computer Science. His Areas of Interest are Computer Networks, Big Data, Cloud Computing and Internet of Things.

A. Murali Krishna completed his Bachelors Degree in Information Technology in the year 2014 from Anna University and completed intense training on "ADVANCED ANALYTICS" from RVS ANALYTICS. He has around 4 years experience in Data Science practice across Retail, FMCG, Insurance and Health Care domains. Currently he is working as a Software Engineer in DELL International Services Private Limited,Bangalore, India - handling Dynamic Learning Driven Predictive Analytics for Performance Engineering. He had also worked as trainer for data scientist course of R and SAS studio for Simplilearn having taught inferential statistics and predictive modeling techniques for more than 700 working professionals with nationalities across the globe. His Areas of Interest are Machine Learning, Artificial Intelligence and Reinforcement learning.

