Proceedings of the International Conference on Industrial Engineering and Operations Management, Dubai, UAE, March 10-12, 2020

Detecting Anomalies in Users – An UEBA Approach

Raguvir. S
Research Assistant
Amrita Vishwa Vidyapeetham,
Bangalore 560035, Karnataka, India

Prof. Shekar Babu, PhD
Amrita Vishwa Vidyapeetham,
Bangalore 560035,Karnataka, India

Abstract

Large organizations across the globe are using advanced security solutions to protect and watch the users information. Even with such advanced solutions these companies are not able to protect or attacks. In addition to attacks one of the key aspects is users behaviour and detecting anomalies when the users are utilizing the systems on the network as well the patterns in their behaviour. Lack of proper monitoring and controls implementation and data breaches are seen. The security professionals within the organizations as well as outside are grappling to solve these issues. One of the new approaches to information security is User Entity Behaviour Analytics (UEBA). One of the biggest challenges with incident response is the large amount of data that the system environment has generated and how to accommodate and analyse the data. Analytics within the area of information security is a new area. Analytics professionals are working on creating rules and correlation aspects, in addition to trends and behaviour patterns with respect to the users behaviour and their approach. One of the key focus areas of UEBA is on users actions and behaviours. Behaviours, users access as well as their usage anomalies are popular and interpretation of these anomalies or malicious activities is very critical. UEBA approach is a viable approach in the area of security to detect user behaviour anomalies using methods like statistical analysis and machine learning. The paper aims to show how analytics and specifically UEBA can help in users patterns and any anomalies within these patterns. With the focus on user behaviours and the analytics related to user behaviours the authors look at the insights, benefits and the utilization of resources in the area of security. The various parameters analysed for the users are user name, IP Address, time of usage, date of usage. The data was analysed over a period of 3 months. The researchers developed patterns using a visualization dashboard and used mining, script and processing of raw data before developing visual analytics. The various anomalies were highlighted from the different patterns.

Keywords

Security; User Entity Behavior Analytics (UEBA); Anomalies, Visual Analytics; Cybersecurity

1.Introduction

In enterprise security one of the key aspects is to detect the user accounts that are compromised and also the insiders within the company whose intents are malicious. The problem gets complicated when we have many scenarios and varied networking environments across the company. However, such issues could be tackled if the actions of the compromised user are quite different from the users daily duties. However, we could track the actions of the user's

actions over a certain period of time and accordingly develop a baseline profile and explore for any deviations from these baseline which can be flagged off as a potential anomaly.

In this paper, the researchers explore the various patterns of users, the users frequency, the various users locations and how with the use of User and Entity Behavior Analytics (UEBA) concepts the researchers are exploring patterns and activities. The authors apply basic statistical methods on large data set to identify if there are any anomalies in the behavior of users, by looking through their patterns.

The application of anomaly detection and the techniques and advanced approaches like machine learning to solve detection problems is not new. Research work on anomaly detection can be dated back to 1987 (Denning, 1987). There is also recent work found in the area of anomaly detection (Chandola *et al*, 2009), however there is little applications that can be found in industry. Currently in the industry there are interests regarding user's behavior and the analysis but with a lot of skepticism (Pinto, 2014), (Gates *et al.,* 2007), (Rieck *et al,* 2011) gives a gist of assumptions and the various details related to domain, data and operational aspects. (Sommer and Paxson, 2010) gives characteristics across security domain.

Hence, in UEBA it is quite important to develop relevant use-cases and a clear defined scope. Anomaly detection can only helps in identifying users events.

2.UEBA Definition

UEBA is a holistic approach to ensure an organization has a top-notch security while detecting the users that might cause a security breach in the system. UEBA can identify normal and abnormal behavior for both humans & computers to provide complete visibility. UEBA uses big data, data analytics, artificial intelligence, machine learning and much more to know whether there is any deviation from the usual patterns thereby arriving at the anomalies that could be a potential threat to the organization. ("What is User Entity Behaviour Analytics? A definition of UEBA, Benefits, How it works and more")

3.Benefits of UEBA

In UEBA, tracking of user entities are done. Currently, the preventive measures like Firewalls are not 100% fool proof. Hackers & attackers enter the system at any point, hence the need for the detection is very important for minimal damage.

4.Functions of UEBA

1. Detect insider threats: To detect the employees within the organization who could steal the data & information. Therefore, help in detecting data breaches
2. Detect compromised accounts
3. Detect breach of protected data
4. Detect Cyber attackers by discovering security abnormalities ("What is User Entity Behaviour Analytics? A definition of UEBA, Benefits, How it works and more")

5.Literature Review

Table 1: Literatures on UEBA

Sl.No	**Author, Year of publication, Title of article, Journal/report where published**	**Objectives/major issues discussed in article**	**Variables/ methodologies used**	**Findings of the study**
1	Maroun Touma, Elisa Bertino, Brian Rivera, Dinesh Verma and Seraphin Calo "2017,Framework for behavioral analytics in anomaly identification".	The authors in this paper, propose a framework on the role of policies and behavior security in a coalition setting with emphasis on anomaly detection and individual's deviation from group activities.	New constraints in the form of policies	They compared repetitive tasks that occur when two autonomous systems interact; we can detect anomalous behavior and prevent it from recurring.

2	Madhu Shashanka, Min-Yi Shen and Jisheng Wang, "2016, User and Entity Behavior Analytics for Enterprise Security, IEEE International Conference on Big Data"	The authors in this paper, provide an overview of an intelligence platform that address threats and incident investigation. The authors specifically, focus on the tracking of users behaviors and their location through IP addresses.	SVD based algorithm	The authors developed a solution and to detect anomalies which would help security analysts.
3	Baoming Tang, Qiaona/Joanna Hu, and Derek Lin, "2017, Reducing False Positives Of User-to-Entity First-Access Alerts for User Behavior Analytics, IEEE International Conference on Data Mining Workshops"	In this paper, the authors use factor techniques for data normalization to make prediction using enterprise logs. They developed methods to reduce false positives in users behavior.	Factorization machine based method	The authors show how their methods can reduce false positives.

4	Jamie Graves, 2017, How machine learning is catching up with the insider threat, Cyber Security: A Peer-Reviewed Journal	This paper aims to show how and why machine learning capabilities can help organizations to reduce these inadequacies, providing an essential extra element of protection	Machine learning	They concluded that human element is still required to analyze events, confirm that they are malicious and provide further forensic detail; machine learning has the potential to take us leaps and bounds ahead when it comes to detecting, analyzing and responding to the insider threat.
5	Danilo Ciscato, Mark Fabbi, Andrew Lerner, 2017, Magic Quadrant for Data Center Networking, Gartner Reprint	This article focuses on Network solutions for enterprise data centers are evolving to support better agility and cloud architectures. Enterprise I&O leaders should evaluate multiple vendors, and focus on comparing architectures, software features and infrastructure integration capabilities, not hardware specifications		
6	Gorka Sadowski, Avivah Litan, Toby Bussa, Tricia Phillips, "2018, Market Guide for User and Entity Behavior Analytics".	This article speaks about UEBA and how it was classified, their features, vendors etc.		

From the above Table-1 the authors have developed an approach to analyze and understand UEBA through the various literature papers and manuscripts the different researchers have studied. The above format and method developed by the authors above is across multiple dimensions and trying to map and relate to the existing methods across the industry. The authors analyzed in depth the parameters like issues, methodology and findings that the researchers have specifically in UEBA.

Since UEBA is a very recent and a nascent field there are very few research studies by researchers across the globe. Hence, the authors have very limited number of papers for the above analysis.

6. UEBA Implementation

6.1 How UEBA Works

The premise of UEBA is actually very simple. An employee's username and password can be easily stolen, yet it is substantially harder to emulate the individual's typical behavior once inside the network. For instance, let's say we steal someone's password and username. We would still not be able to act precisely like that person once in the system, unless given extensive research and preparation. Hence, when that person's username is logged in to the system and his behavior is different from the ordinary behavior that is when UEBA alerts start to sound.

Another relatable analogy would be if the credit card were stolen. A thief can pickpocket a wallet and go to a top of the line shop and start spending thousands of dollars using the credit card. If the spending pattern on that card is different from the thief's, the company's fraud detection department will often recognize the abnormal spending and block suspicious purchases, issuing an alert to the owner of the card or asking him to verify the authenticity of a transaction.

6.2 UEBA Architecture

One of the clear architecture is the one mentioned in the IBM Watson's Security (IBM QRadar Advisor with Watson: Revolutionizing the Way Security Analysts Work, 2017).

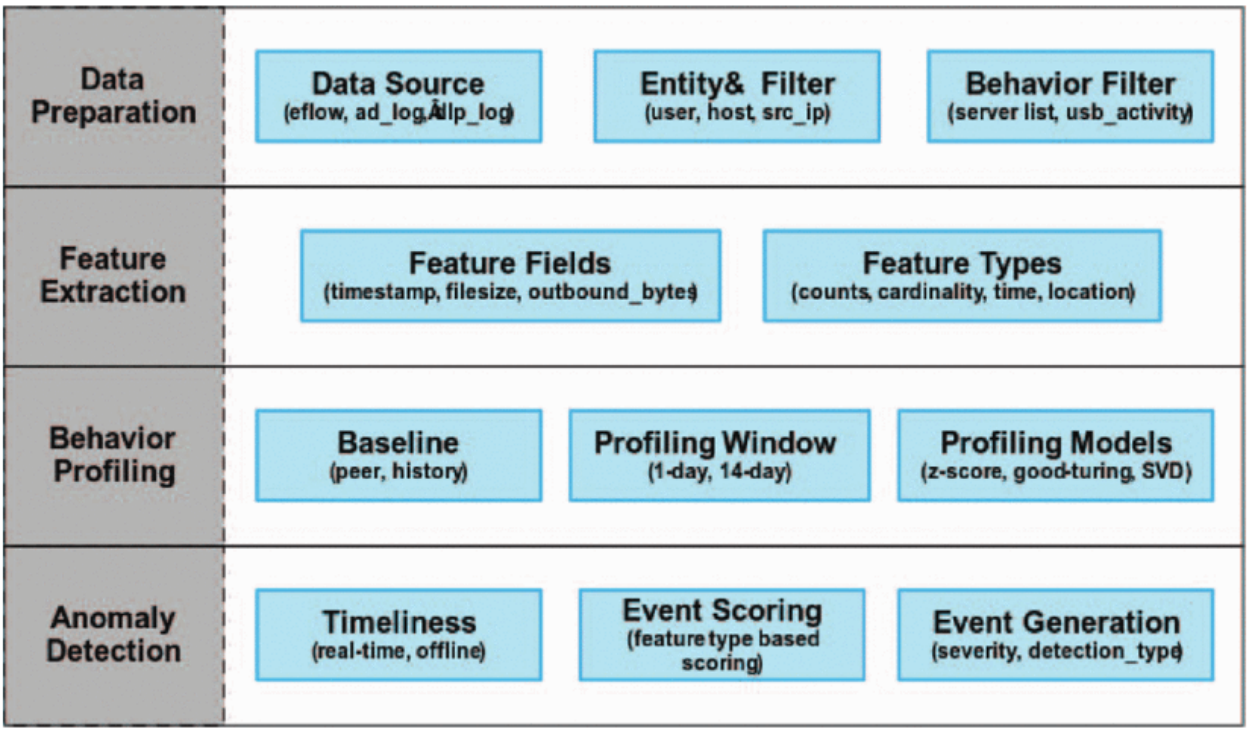

Figure 1. General Architecture of EUBA

6.3 Competitive landscape in the area of Information security

The global user and entity behavior analytics (UEBA) market is projected to grow from USD 131.7 million in 2016 to USD 908.3 million by 2021, at a CAGR of 47.1% between 2016 and 2021. The North American region is estimated to account for the largest share in the global entity behavior analytics market in 2016. The Asia-Pacific user and entity behavior analytics market is showing positive trends as several companies and industries are adopting user and entity behavior analytics solutions at various levels to strive in the market and to increase their productivity. The Asia-Pacific user and entity behavior analytics market is expected to witness exponential growth and is projected to grow at the highest CAGR during the forecast period. ("What is User Entity Behaviour Analytics? A definition of UEBA, Benefits, How it works and more")

This is due to increasing demand for user and entity behavior analytics solutions and services in this region. Moreover, rapid growth in the usage of web and mobile applications in the Asia-Pacific region and the need to protect these applications from vulnerabilities have resulted in increased demand for user and entity behaviour analytics solutions that identify security gaps in the network infrastructure and web and mobile applications, and help in reducing risks associated with them. ("What is User Entity Behaviour Analytics? A definition of UEBA, Benefits, How it works and more")

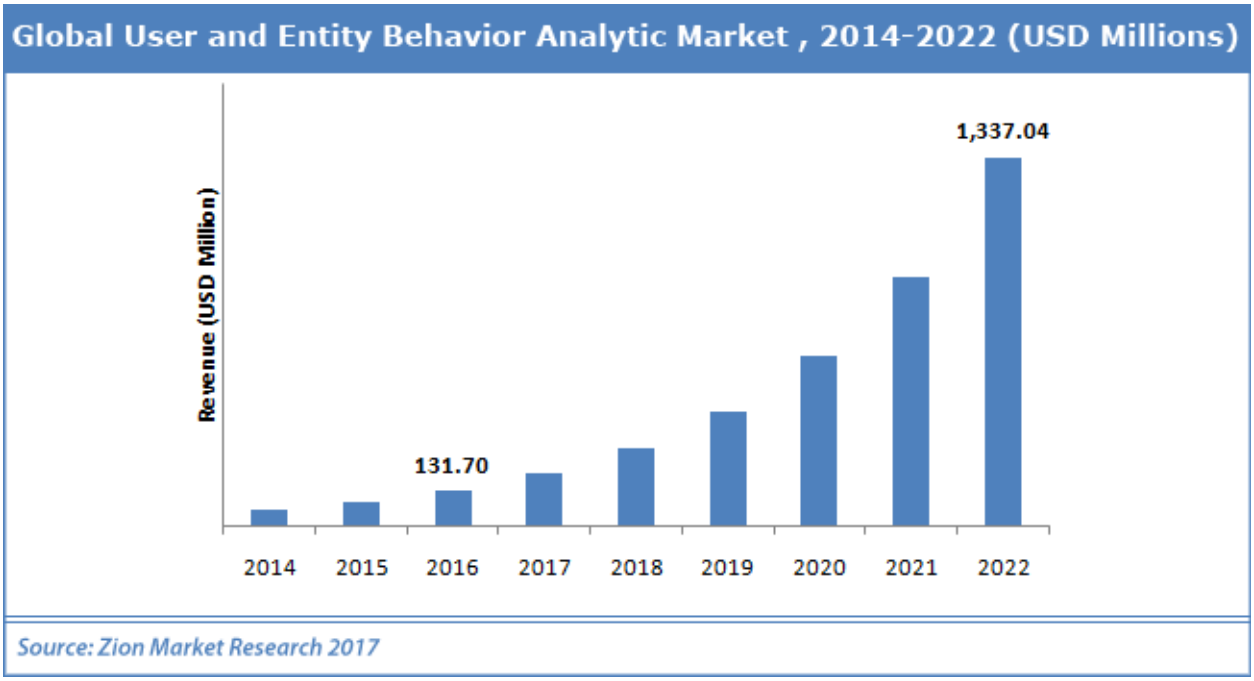

Figure 2. Growth Trend of UEBA

6.4 UEBA Solutions working

UEBA solutions identify patterns in typical user behavior and then pinpoint anomalous activities that do not match those patterns and could correspond with security incidents. UEBA solutions explore patterns in behaviors by applying statistical methods.

Figure - 3 below shows the Difference between UEBA from cyber security.

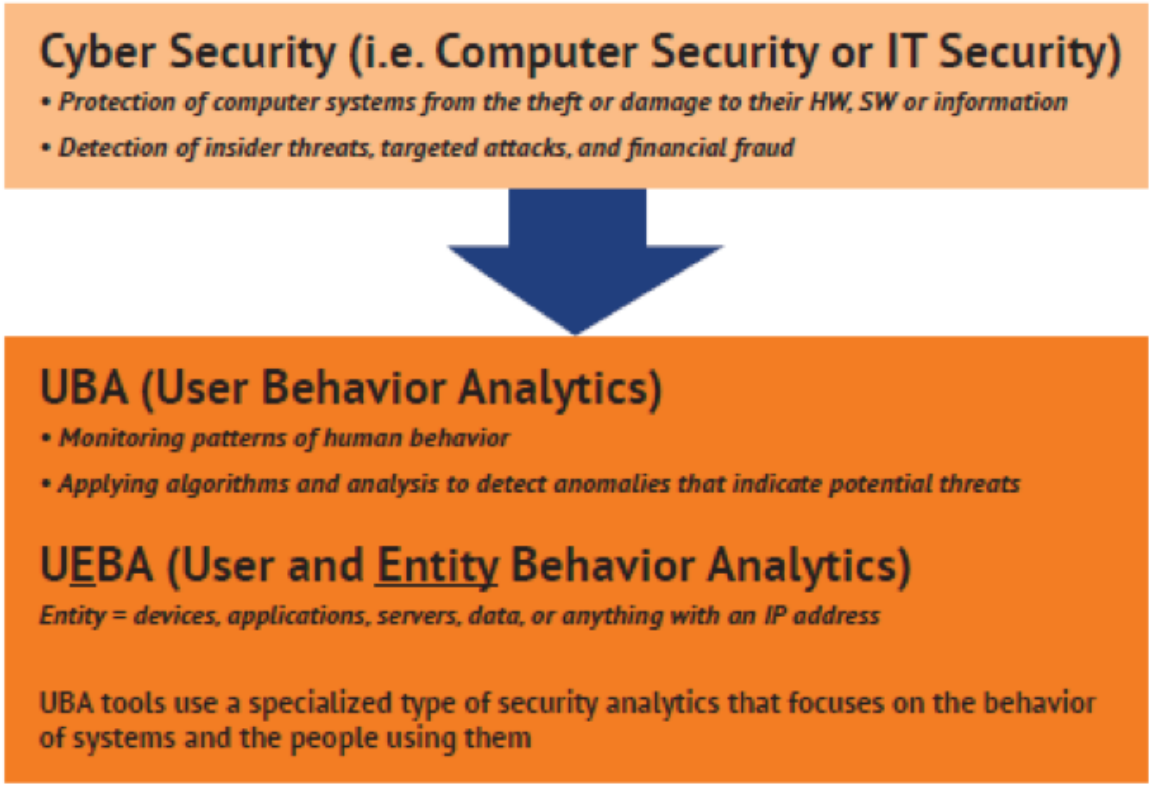

Figure 3. Difference of UEBA from Cyber Security

7.Findings

The data provided contained the list of users and the IP addresses that they accessed. Location details can be obtained from the IP address by using IP2location on R. This gives more details about the IPs, some of the details include city, state, country and the coordinates of the location. Based on the obtained data as shown in Table 2, the following graphs were made.

Table 2. Data after obtaining locations

IPs	vpnuser	abrv	country	region	city	zip	lat	long	gmt
1.136.104.147	User 24801	AU	Australia	Victoria	Melbourne	3000	-37.814	144.96332	+10:00
1.136.106.192	User 24801	AU	Australia	Victoria	Melbourne	3000	-37.814	144.96332	+10:00
1.136.107.93	User 24801	AU	Australia	Victoria	Melbourne	3000	-37.814	144.96332	+10:00
1.136.108.232	User 24801	AU	Australia	Victoria	Melbourne	3000	-37.814	144.96332	+10:00
1.136.109.182	User 24801	AU	Australia	Victoria	Melbourne	3000	-37.814	144.96332	+10:00
1.136.109.97	User 24801	AU	Australia	Victoria	Melbourne	3000	-37.814	144.96332	+10:00
1.136.111.154	User 24801	AU	Australia	Victoria	Melbourne	3000	-37.814	144.96332	+10:00
1.144.109.65	User 3662	AU	Australia	Queensland	Gold Coast	4217	-28.00029	153.43088	+10:00
1.152.104.117	User 24801	AU	Australia	New South Wale	Blakebrook	2480	-28.76667	153.23333	+10:00
1.152.104.228	User 24801	AU	Australia	New South Wale	Blakebrook	2480	-28.76667	153.23333	+10:00
1.152.104.81	User 24801	AU	Australia	New South Wale	Blakebrook	2480	-28.76667	153.23333	+10:00
1.152.105.67	User 24801	AU	Australia	Tasmania	Gawler	7315	-41.18388	146.15913	+10:00
1.152.110.250	User 24801	AU	Australia	Western Australi	Perth	6003	-31.95224	115.8614	+08:00
1.160.147.46	User 29555	TW	Taiwan, Province	T'ai-wan	Hsinchu	300	24.80361	120.96861	+08:00
1.171.93.93	User 29555	TW	Taiwan, Province	T'ai-wan	Hsinchu	300	24.80361	120.96861	+08:00
1.186.150.134	User 8009	IN	India	Maharashtra	Lohogaon	412207	18.6	73.91667	+05:30
1.186.150.135	User 8009	IN	India	Maharashtra	Lohogaon	412207	18.6	73.91667	+05:30
1.186.150.144	User 8009	IN	India	Maharashtra	Lohogaon	412207	18.6	73.91667	+05:30
1.186.150.146	User 8009	IN	India	Maharashtra	Lohogaon	412207	18.6	73.91667	+05:30
1.186.150.148	User 8009	IN	India	Maharashtra	Lohogaon	412207	18.6	73.91667	+05:30

Figure 4 shows how activity of users varies from different countries. It can be observed that India covers a major area of the chart at 69%. This means that a large portion of the recorded data is from India. Next highest usage comes from The United States.

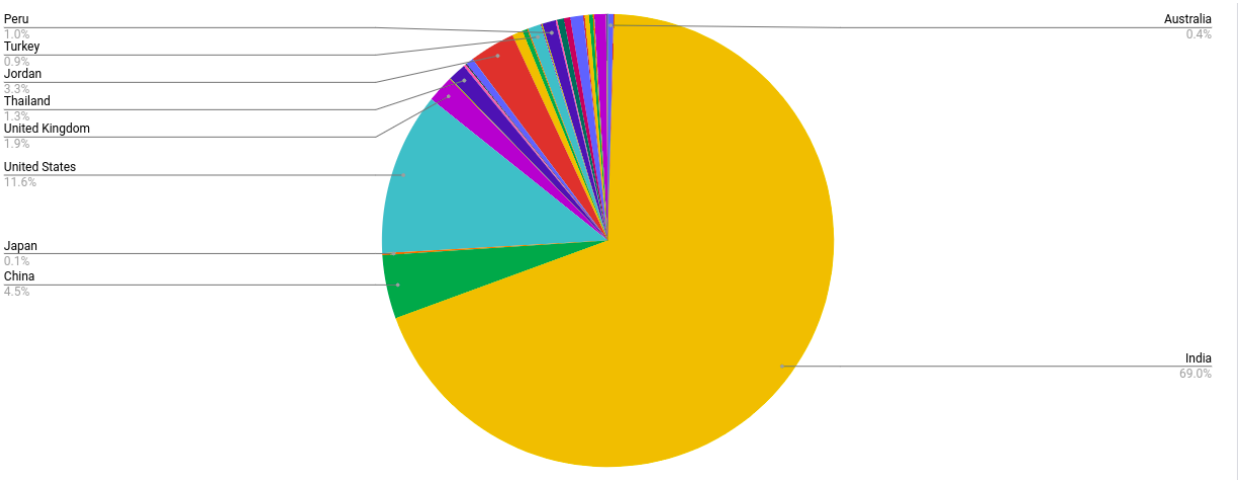

Figure 4. User activity according to country

Usage rate of different users can be understood from Figure 5. User 8728 has the highest usage, while most other users have not accessed more than 50 times and a few above 100.

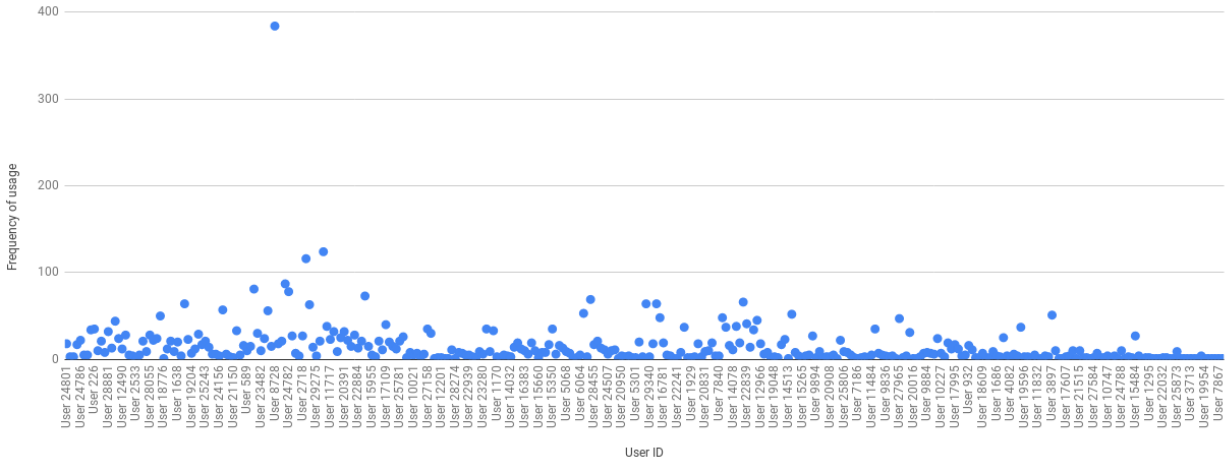

Figure 5. Frequency of access by users

Figure 6 gives an understanding of how user activities are in different cities across the world. According to the legend, areas marked with a shade of red imply high access while those with a green shade indicate low access from the areas. The map shows two areas with larger markers in red in the south asian region. These would be Mumbai and Bangalore. These two regions can be associated to the high access rate from India (from figure 1). There are also other markers in the south asian region. In Asia, other regions that indicate a lot of activity are China (Beijing, Shanqui, Shanghai and Shunyi) and Thailand (Bangkok). The australian region shows activity around the coast including Melbourne, Perth, Gold Coast and Gawler. Activities seem to reduce around middle east and Europe. Cities in Jordan that have been recorded were Amman, Al Mafraq. In Europe, almost all the observed activity comes from the UK (London,Snodland and Blakebrook). There is a significant increase in activity in North America, The US being the primary user. In the US, the number of cities along the east and west coast that were observed were higher, while the mid- region has a lower number.

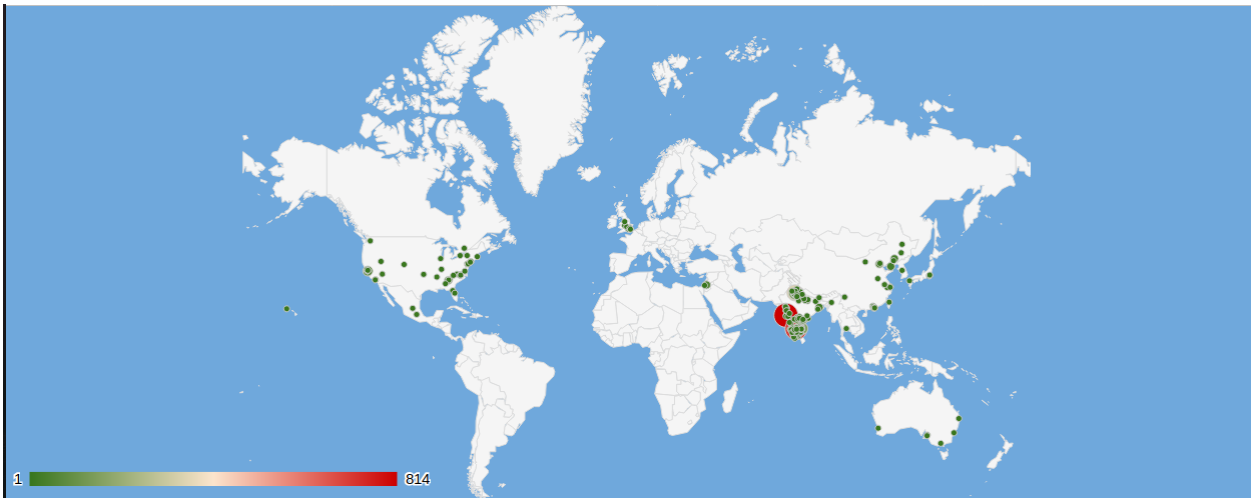

Figure 6. Frequency of access in cities

The higher number of access in India and the US can be accounted to the number of work spaces and clients in the countries.

Data related to IPs was consolidated and visualized into a graph in figure 7. All the IPs were grouped into different ranges and the count of IPs in the ranges were recorded in the graph. It is observed from the graph that the most used IP range is 100.0.0.0 and 149.255.255.255, reaching over 1500. The lowest range is 200.0.0.0 and 255.255.255.255.

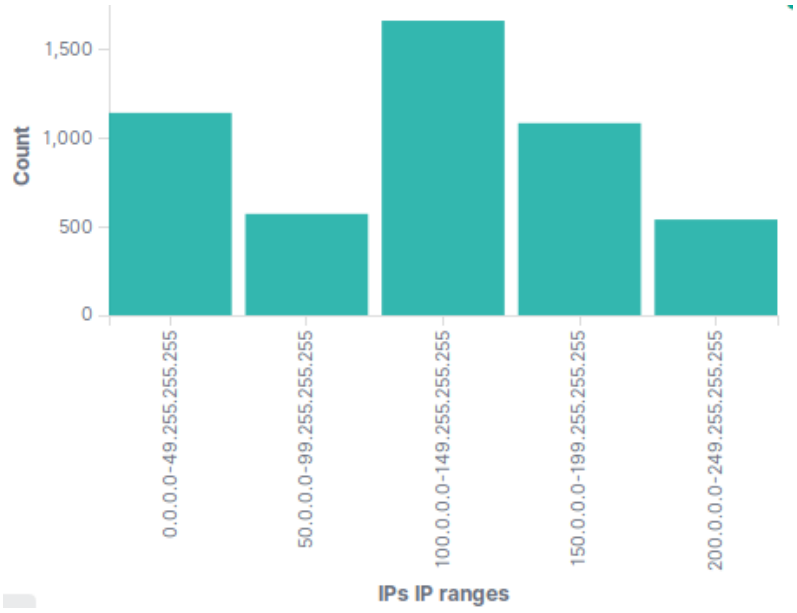

Figure 7. Distribution of IP ranges

8.Conclusions

From the analysis and the data we can observe that the authors focused on the behavior of the users usage of the corporate network from various locations. In addition, the authors also explored to find out the different times these users were utilizing these locations. The authors also wanted to explore how frequently the users are utilizing these corporate services. They also looked at from which locations were these services used and how frequently. The authors also looked at the density of users from various locations, using IP ranges and IP addresses.

9.References

Maroun Touma, Elisa Bertino, Brian Rivera, Dinesh Verma, Seraphin Calo, 2017, *Framework for behavioral analytics in anomaly identification*, Event: SPIE Defense + Security.

Madhu Shashanka, Min-Yi Shen, Jisheng Wang, 2016, *User and Entity Behavior Analytics for Enterprise Security, IEEE International Conference on Big Data.*

Baoming Tang, Qiaona/Joanna Hu, Derek Lin, 2017, *Reducing False Positives Of User-to-Entity First-Access Alerts for User Behavior Analytics, IEEE International Conference on Data Mining Workshops.*

Jamie Graves, 2017, *How machine learning is catching up with the insider threat, Cyber Security: A Peer-Reviewed Journal.*

Danilo Ciscato, Mark Fabbi, Andrew Lerner, 2017, *Magic Quadrant for Data Center Networking, Gartner Reprint.*

Gorka Sadowski, Avivah Litan, Toby Bussa, Tricia Phillips, 2018, *Market Guide for User and Entity Behavior Analytics.*

V. Chandola, A. Banerjee, and V. Kumar. *Anomaly detection:A survey. ACM Computing Surveys, Sep 2009.*

D. Denning. *An intrusion detection model. IEEE Trans. On Software Engg., 13(2),* 1987.

C. Gates and C. Taylor. *Challenging the anomaly detection paradigm: A provocative discussion. In Proceedings of the 2006 Workshop on New Security Paradigms,* NSPW '06, pages 21–29, New York, NY, USA, 2007. ACM.

A. Pinto. *Secure because math: A deep-dive on machine learning based monitoring. In Black Hat Briefings* USA, 2014.

K. Rieck. *Computer security and machine learning: Worst enemies or best friends? In SysSec Workshop (SysSec)*, 2011 First, *pages 107–110. IEEE,* 2011.

R. Sommer and V. Paxson. *Outside the closed world: On using machine learning for network intrusion detection. In Security and Privacy (SP), 2010 IEEE Symposium on, pages 305–316. IEEE,* 2010.

C. Brook "What is User Entity Behaviour Analytics? A definition of UEBA, Benefits, How it works and more" Available at: https://digitalguardian.com/blog/what-user-and-entity-behavior-analytics-definition-ueba-benefits-how-it-works-and-more

Global User and Entity Behavior Analytics Market (UEBA) Market Worth USD 908.3 Million by 2021 - Rapid Growth in Mobile Apps in APAC - Research and Markets Available at : https://www.businesswire.com/news/home/20160823006004/en/Global-User-Entity-Behavior-Analytics-Market-UEBA (2016)

IBM QRadar Advisor with Watson: Revolutionizing the Way Security Analysts Work. Available at: https://securityintelligence.com/ibm-qradar-advisor-with-watson-revolutionizing-the- way-security-analysts-work/ (2017)

Cargills Bank Ltd., https://www.ibm.com/case-studies/cargills-bank-ltd

Juniper Advanced Threat Prevention Products https://www.juniper.net/uk/en/products-services/security/sky-advanced-threat-prevention/

What Are The Biggest Challenges Facing The Cybersecurity Industry?
https://www.forbes.com/sites/quora/2017/09/15/what-are-the-biggest-challenges-facing-the-cybersecurity-industry/#7b5f84ab2d62

Anomaly Detection / Outlier Detection in Security Applications https://www-users.cs.umn.edu/~lazar027/anomaly_detection.htm

What Are the Most Common Cyber Attacks?
https://www.cisco.com/c/en/us/products/security/common-cyberattacks.html

Cyber Security Is A Business Risk, Not Just An IT Problem
https://www.forbes.com/sites/edelmantechnology/2017/10/11/cyber-security-is-a-business-risk-not-just-an-it-problem/#5c679bf77832 (2017)

User and entity behavior analytics: building an effective solution
https://www.slideshare.net/YolantaBeres/user-and-entity-behavior-analytics-building-an-effective-solution (2017)

Cisco Identity Services Engine Technology Partners https://www.cisco.com/c/en/us/products/security/identity-services-engine/technology-partners.html

Cisco to buy cyber security company Duo for $2.35 billion https://www.reuters.com/article/us-duosecurity-m-a-cisco/cisco-to-buy-cyber-security-company-duo-for-2-35-billion-idUSKBN1KN1LA (2018)

https://investor.cisco.com/investor-relations/news-and-events/news/news-details/2018/Cisco-Announces-Intent-to-Acquire-Duo-Security/default.aspx