Misuse Detection in a Simulated IaaS Environment

Burhan Al-Bayati[1, 2], Nathan Clarke[1, 3], Paul Dowland[3] and Fudong Li[1, 4]

[1] Centre for Security, Communications and Network Research, Plymouth University
Plymouth, UK
[2] Computer Science Department, College Science, Diyala University
Diyala, Iraq
[3] Security Research Institute, Edith Cowan University
Perth, Western Australia
[4] School of Computing, University of Portsmouth
Portsmouth, UK
{burhan.al-bayati, n.clarke, p.dowland,
fudong.li}@plymouth.ac.uk

Abstract. Cloud computing is an emerging technology paradigm by offering elastic computing resources for individuals and organisations with low cost. However, security is still the most sensitive issue in cloud computing services as the service remains accessible to anyone after initial simple authentication login for significant periods. This has led to increase vulnerability to potential attacks and sensitive customer information being misused. To be able to detect this misuse, an additional intelligent security measures are arguably required. Tracking user's activity by building user behaviour profiles is one technique that has been successfully applied in a variety of applications such as telecommunication misuse and credit card fraud. This paper presents an investigation into applying behavioural profiling in a simulated IaaS-based infrastructure for the purposes of misuse detection by verifying the active user continuously and transparently. In order to examine the feasibility of this approach within cloud infrastructure services, a private dataset was collected containing real interactions of 60 users over a three-week period (totalling 1,048,195 log entries). A series of experiments were conducted using supervised machine learning algorithms to examine the ability of detecting abnormal usage. The best experimental result of 0.32% Equal Error Rate is encouraging and indicates the ability of identifying misuse within cloud computing services via the behavioural profiling technique.

Keywords: Continuous identity verification, misuse, behavioural profiling, IaaS, cloud computing services

1 INTRODUCTION

According to the Cisco Global Cloud Index, by 2019, more than 80% of all data centre traffic will be cloud traffic, and around 86% of all amount of processing will be achieved in cloud infrastructure services [1]. Moreover, cloud Infrastructure as a Service (IaaS) is a vital underlying infrastructure model that supports all other cloud services. However, IaaS was reported as the most vulnerable cloud model [2]. As a result, customers would have concerns about unauthorised access to their information that is remotely managed in these services.

Due to the online nature of those services, authentication provides the primary security control to prevent misuse by relying upon point-of-entry based passwords. By stealing customers' login credentials, hackers can gain illicit access and misuse the service and user information. Many incidents have targeted popular cloud computing service providers, for example:

- According to Cloud Security Alliance, a number of security incidents occurred to a British telecom provider (TalkTalk) in 2014 and 2015, resulted in disclosing four million of their customers' personal information [3].
- The Microsoft Azure cloud computing platform faced a serious security incidents in March 2009, led to a massive collapse and outage of the service for 22 hours, with a loss of 45% of user data [4].
- Dropbox was hacked in July 2012; usernames and passwords of many users were stolen from third-party websites; these stolen credentials helped hackers to get access successfully to customers' accounts and misused their data [5].
- Apple iCloud was compromised in 2014 as more than 20,000 passwords of its customer accounts were stolen, resulting in user's personal photographs, specifically celebrities, being leaked online [6].
- Google's Gmail server faced attack in 2016; more than 272 million email addresses and passwords were stolen [7].

It is clear from these incidents that users' sensitive information within cloud computing services can be abused by cybercriminals even though security controls were in place and dedicated security teams were allocated. Therefore, additional security techniques are required to protect cloud services from being compromised and misused. This paper proposes a novel continuous identity verification system as a solution to protect cloud infrastructure service by operating transparently to detect abnormal access. Behavioural profiling can provide a continuous and transparent assessing user's identity while they interact with cloud services. By creating user behaviour profiles, it can be identified people based upon the way in which they interact with cloud computing services. Therefore, the current user's activities (e.g. time of opening the service) are compared with an existing user's template - which is generated from historical usage—by implementing a machine learning algorithm such as Random forest. The comparison result will be determined if the current user is legitimate or not. Therefore, the misuse in this paper means, any abnormal usage that is diverted from a profile while users interact with their cloud services.

The remainder of the paper is organised as follows: Section 2 introduces the state of the art in behavioural profiling. Section 3 presents the experimental methodology. A series of comprehensive experimental studies that evaluate the applicability of using behavioural profiling with cloud IaaS are presented in Section 4. Section 5 discusses the impact of the experimental results, and the conclusion and future directions of this work are presented in Section 6.

2 RELATED WORK

A variety of studies have investigated behavioural profiling from a number of security perspectives, including intrusion detection, fraud detection, and authentication across different technologies (e.g. mobile phone system, network, computer system, and web browsing). Table 1 provides an analysis of these studies.

Table 1. Related behavioural profiling studies

Studies	Activity	Platform	#Participants	Performance (%)	Method	Purpose
[8]	Mobility	S	50	DR=50,FRR=50	Instance based learning	IDS
[9]	Telephony	S	5000	DR=80	Genetic Programing method	FD
[10]	Telephony	S	180	FRR=3	self-Organizing Map, Probability	FD
[11]	Telephony	S	300	DR=70	Neural Network	FD
[12]	Mobility	S	100	DR=81	cumulative probability and Marko properties	IDS
[13]	Mobility	S	178	DR=94	cumulative probability and Marko properties	IDS
[14]	Telephony	S	94	DR=97	SVM	FD
[15]	Telephony, SMS, Browsing, Mobility	C	50	DR=95	Probability	Au
[16]	Telephony, SMS, Browsing	C	35	EER=1.6	Bayesian network, RBF, KNN, RF	Au
[17]	Telephony, Device Usage, Bluetooth network scanning	C	30	EER=13.5, 35.1, and 35.7	RBF network	Au
[18]	Application, Teleph-ony, SMS	C	76	EER=13.5, 2.2, and 5.4	Neural network	Au
[19]	Application Usage	C	76	EER=9.8	Rule base	Au
[20]	Text, App, Web and location	C	200	EER=3	SVM	Au
[21]	Way of using PC	C	10	EER= 7	Neural (FF-MLP)	Au
[22]	File access activity and network event	C	8	FAR=14, FRR=11	K-Means Clustering	Au
[23]	File access activity	C	18	FAR=1.1	SVM	Au
[24]	Web Browsing	S	100	DR=91	support-based, lift-based profiling	I
[25]	Web Browsing	S	10	EER= 24	SVM	Au

*S: Server, C: Client, DR: Detection Rate, FRR: False Reject Rate, FAR: False Accept Rate, EER: Equal Error Rate, SVM: Support Vector Machine, KNN: K-Nearest Neighbours, RBF: Radial Basis Function, RF: Random Forest, FD: Fraud Detection, IDS: Intrusion Detection System, Au: Authentication, I: Identification

Early research focused mainly on IDS and fraud detection via identifying the user behaviour activities during the interaction with telephony services, such as calling and mobility [8-14]. Several classification algorithms were successfully developed to handle various attributes of defined and undefined attacks. A number of these studies implemented a large dataset (as shown in Table 1) that achieved accuracy of a Detection Rate (DR) ranging of 50% to 80%. However, the performance was not clear because the studies only show the False Rejected Rate (FRR) without including the False Accept Rate (FAR). In comparison, more recent studies focused on transparent authentication through modelling application usage to alleviate device misuse [15-20]. Much more information can be gathered from user activities while interacting with these applications (e.g. phone calls, emails, websites visits, and calendar activities). These activities were exploited to build an accurate behavioural profile which can be investigated to increase the accuracy level of the security system for the device or applications. The best accuracy result achieved in these studies were by [16] through applying four classification methods on 35 participants with an overall Equal Error Rate (EER) of 1.6%. The Random Forest classifier achieved the highest accuracy algorithm in this work with a true positive rate of 99.8% and an accuracy rate of 98.9%.

Further studies focused on the generation of user behaviour profiles from desktop computer usage to detect any illegal access to the device [21-23]. A number of features were extracted to build user behaviour profiles in the computer system, including applications being used, the time and interval of accessing files, and websites being visited. The accuracy of these studies was around 7% of EER. However, the number of participants was limited (ranging from 8 to 18 users) which does not reflect an accurate performance in practical sense. From the server side perspective, studies focused on building a user identifier by using their web surfing activities from numerous log files of websites [24,25]. A user behaviour profiling was created based on spending time on various topics of the website, site names, number of pages, starting time and duration time of sessions. Accurate user behaviour profiles have been built to detect illegitimate usage. The best performance achieved by [24] including 100 participants with an DR of 94%. However, the study did not involve all users in the practical experiment, only a few users who had at least 300 sessions in the dataset were selected to test the system. Therefore, it is difficult to be implemented this system for solving large scale problems.

As demonstrated by existing literature, the behavioural profiling technique has been applied successfully across different technologies including mobile phones, computers (client and server) to improve the system security level. However, to the best author's knowledge, no prior work that utilises the behavioural profiling technique has been studied regarding cloud infrastructure services.

3 EXPERIMENTAL METHODOGY

The main aim of this study is to focus upon understanding to what degree behaviour profiling can be successfully applied to verify the users via their usage within cloud infrastructure services – understanding whether it is the genuine user or not in order to provide a basis for a security system to respond. Therefore, a series of experiments were conducted on a real dataset to examine the impact of a number of factors on the performance of the machine learning algorithms. These include two further research questions:

- Does the volume of data and sampling selection for training and testing have impact on performance and classification algorithms?
- How much data and time are required to generate a user template within a specific criteria?

In order to achieve these experiments, users' interactions with the cloud infrastructure application are required to be collected. However, the collection for users' activity in cloud computing services proved to be problematic. To the author's best knowledge there are no public datasets that would be used for this study; and cloud providers would be unwilling to provide such access directly due to privacy and security concerns. Whilst it is possible to create IaaS based images and have a population of participants use these machines for a specified period, it was felt this might result in behavioural patterns that do not truly reflect user's normal activities. Consequently, a decision was made to capture users' interactions within their own personal computers simulating the environment of a cloud infrastructure service. Software was developed and installed on the participants' computers. This software works in the background of the computer operating system in non-intrusive manner. Activities of 60 participants (including PhD researchers and undergraduate students) were obtained resulting in a private dataset containing 1,048,195 users' interactions. The interactions comprise the following information: the start and the end time of applications being used (e.g. Excel, Word and MatLab) and web services (URLs) being visited. The data has been anonymised to protect the participants' privacy and ethical approval was sought and obtained from the authors' institution. Table 2 demonstrates a sample of user actions within the dataset.

Table 2. User activity with personal computer

Day	Hour	Minute	Second	App/URL	Event
2	9	10	8	Word	Focus
2	9	21	16	Word	Lost focus
2	9	21	20	Endnote	Focus
2	9	23	44	Endnote	Lost focus
2	10	15	30	Paint	Focus
2	10	45	23	Paint	Lost focus
2	10	45	30	v2wLG+lIc…	Focus
2	10	49	13	v2wLG+lIc…	Lost focus

The hour, minute with applications and URLs have been used were selected as the main features set in this study which could provide a good level of pattern recognition amount the users. In order to make those features acceptable by classification algorithms, the symbolic-valued attributes (e.g. name of applications and URLs) were enumerated into numerical attributes and into the range of 0-1 [26].

The dataset of each user was divided into two sets: the first set was utilised to create a user behaviour profile for training the classification algorithm whereas the second set of the user's data was utilised to assess the performance of the algorithm (i.e. test data). Also, at no point a sample is used for both training and testing. Due to the verification nature, the data is classified into 2-class problem (i.e. either belong to legitimate users or impostors). One user acts as the genuine user, whilst all the rest users are treated as impostors. This procedure is then applied sequentially to all the other users in order to ensure they have the opportunity to represent as an authorised user. Whilst the purpose of this classification is misuse detection, the approach is determining this through verifying the authenticity of the user. As such, the results will be presented in the form of FAR, FRR and EER – standard performance metrics that are widely used in biometrics. The EER is widely used as a key metric to evaluate the performance (when the FAR and FRR are equal).

Two experiments were developed to investigate whether the collected data can be used to differentiate those users whom generated by. The first experiment implemented two classification algorithms on the given dataset: Random Forest (RF) and Classification And Regression Trees (CART). Decision tree algorithms are fast in identifying unknown instances in a large dataset and can easy deal with discrete attributes particularly handling outliers [27]. Therefore, the first classifier was selected due to its outstanding performance that was achieved in the previous studies (as illustrated in Table 1); while the second method was selected based on the study by [27] that conducted on different classification algorithms and the CART was a one of the best ten algorithms that achieved the highest performance. The configuration of the two selected classifiers remained as default except the number of trees of RF was modified into 30 as this number achieved a better accuracy among other configurations. Three splitting approaches for training and testing data were applied on these algorithms: 50/50, 66/34 and 80/20 in order to examine the impact of data volume on the overall performance. For each data volume setting, two sample selection methods were utilised: a random sample selection across the dataset and a time series sample selection (i.e. samples are selected sequentially as in a reality sense). The outcome of this experiment would explore how the performance of the system is affected by investigating the nature of different classification approaches. The optimal classifier can also be identified based on the findings of this experiment. Additionally, the comparison between the accuracy of the result of each data volume would give a better understanding of the nature of user behaviour profiles with the impact of the sample selection on the performance of the algorithms.

The second experiment focused upon exploring how much training data is required to generate a user template with an acceptable level of performance. For the purposes in

this study an EER of 10% was set. In practice, a user profile would need to be created based upon time-series rather than random sample selection. As such, time-series sample selection was applied to achieve the goal of this experiment; the first day's data was used for training and the data from remaining days was employed for testing, then the data from first and second days was used for training and the remaining data from the rest days was utilised for testing and so on.

4 Experimental Results

4.1 VARIOUS TRAIN/TEST SET RATIO WITH TWO SAMPLING METHODS

The results of this experiment (as illustrated in Table 3) are encouraging to support the idea of verifying the genuine user or identifying misuse of unauthorised access to cloud infrastructure services.

Table 3. Performance of classification algorithms

Classifier	EER (%) Volume of data with time series selection			EER (%) Volume of data with random selection		
	50/50	66/34	80/20	50/50	66/34	80/20
RF	15.35	13.18	12.40	4.07	3.57	3.09
CART	8.51	7.35	6.55	0.69	0.44	0.32

As seen in Table 3, the nature of the classifier had a significant impact on improving the system performance. The CART algorithm achieved a higher accuracy than the RF method regardless the amount of data being allocated to training and testing. This includes the time series and random sample selections with the highest accuracy of 0.32% EER.

From the sample selection perspective, Table 3 shows that the random sample selection achieved better performance than the time series selection within both classifiers and across all volumes of training and testing data split. This can be attributed to the high probability of selecting a variety of user activities across the entire usage range whilst employing the random sample selection. It is also worth highlighting that the change in performance with both types of sample selection (random and time series) gets better as the amount of training data increases; decreases of 2% and 0.37% in EERs can be observed for time series and random selection respectively. As the proportion being reduced is significant, this suggests that the nature of user behaviour across the three-week collection period is likely to be relatively changeable. Therefore, care must be taken to ensure appropriate template renewal procedures are developed to maintain performance levels.

The classifiers' overall average performance in terms of data volume (as illustrated in Table 3) also shows that the training phase with large sample volumes achieved better performance than those with smaller data volumes. Based upon the overall average

individual performance using the CART classifier with 80/20 of data splitting and random sample selection, the trend line regression approach, as illustrated in Fig.1, also supports the same idea. Users with high volumes of interaction achieved better performance than users with fewer interactions. This support the idea put forth by prior research that more volume would provide better accuracy [17, 24, 26]. Additionally, it is logical as the classifier can learn more about the pattern usage of a user by acquiring a large volume of data, leading to better performance.

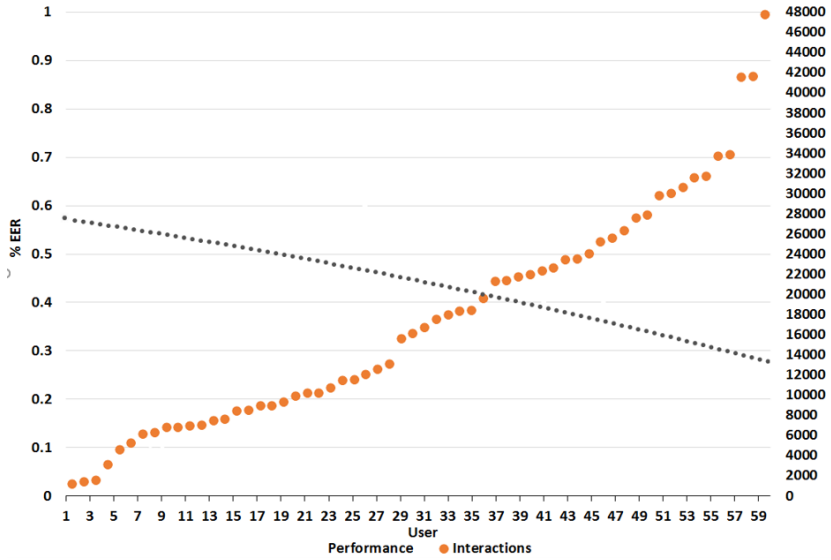

Fig. 1. Average performance based on volume of data

4.2 TIME AND VOLUME OF DATA REQUIRED FOR GENERATING USERS TEMPLATES

This experiment focused upon the amount of data and time are required for each user to generate a user template based on predefined criteria (10% of EER). The CART classifier was chosen for this experiment due to its outstanding performance in the first experiment. In a practical sense, the data split between training the classifier and testing the performance was selected based on using the daily basis as a time window, as mentioned previously. Fig. 2 demonstrates the statistical distribution (min, median, and max) of the performance of all users across 20 days.

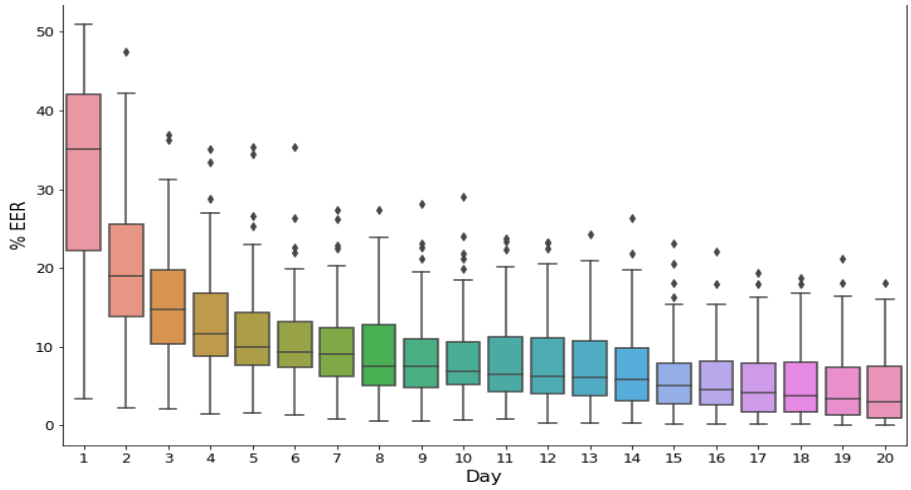

Fig. 2. Distribution of users' performance across 20 days

Fig. 2 shows that the classifier achieved a significantly higher performance for larger volume of training data than the low volume of samples for the training phase, specifically within the first five days. Therefore, based on the overall distribution of users' result accuracy, it suggests that at least five days of user data are needed as an overall average time to profile individuals within the given criteria. However, it can be seen on the chart that there is also a variation among actual users; some users would need less than five days and others would need more to generate the template. Therefore, further investigation is required to determine the actual time and interactions required for each user. Fig. 3 demonstrates the minimum days and interactions needed for each user to build suitable user behaviour profiles.

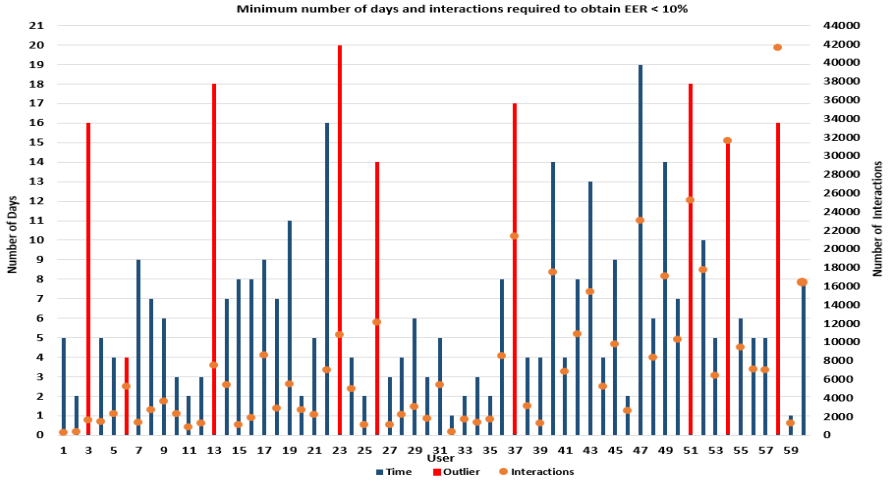

Fig. 3. Time and volume of data required for generating users' template

Fig. 3 shows the time and volume of interactions required to generate a user's template are different among users. For example, some users achieved the selected criteria in equal or less than two days training with lower interactions compared with other users (i.e. less than 1,500 interactions), such as users 2, 11, 32, 33, 35, and 59. In comparison, some users needed more than the half of the given period of training with high interactions (more than 17,000 interactions) to achieve the goal, such as users 40, 47, 49 and 52. Moreover, some users (outliers) did not achieve the criteria even though they had a long period of training with the highest interactions among most other users (more than 21,000 interactions) such as users 37, 51, 54, and 58. This case is a common issue for behavioural based biometrics as users' behaviour tend to change over time and under different external circumstances, which can impact negatively on the sample collection and classification performance. As less control over users and the environment exists, more care is needed when considering their implementation in a verification system. Therefore, it demonstrates that the time and volume factors worked for 80% of the user population. However, these factors may not be always necessary for determining appropriate discriminatory information for users that can help to generate a suitable temple to support the classifier for achieving the correct decision.

5 Discussion

The experimental results reveal that cloud infrastructure service users based on a simulated environment can be identified via their activities with a high degree of accuracy. In addition, although the prior work within similar environment (computer desktop) in [22, 23, 24] implemented a small dataset with limited number of participants (8-18), the overall outcome was approximately 7% of EER. While this study applied a larger dataset containing more than million samples with 60 participants—in comparison to the prior art—and the performance was better with the best EER of 0.32%, suggesting the usefulness of the proposed technique.

From an individual classifier performance perspective, the experiment showed that the CART algorithm achieved 0.32% EER and outperforms the RF with random sampling and 80/20 training and testing data splitting. This would allow other factors such as time taken to compute, computational overhead, and memory requirements to be considered as part of the selection. Also, the overall result accuracy of the large volume of data had a positive impact. Users' performance improved with more frequent activities/interactions across both classification algorithms. Moreover, the performance results with random sample selection also achieved better accuracy than the time series selection. This indicates user behaviour is changeable over time and therefore care must be taken to ensure appropriate template renewal procedures are developed regularly to maintain levels of performance.

For the time and data volume required to generate a user template, the experiment revealed that five days can be considered as the average time for generating useable user behaviour profiles, as shown in Fig. 2. However, the five days as a static thresh-

old is not a definitive criterion for creating a user template. A number of users needed less than five days while others needed more as illustrated in Fig. 3. Also the large volume of data for training is not always guaranteed to perform with better accuracy than the low volume of activity for all users. Therefore, further statistical analysis was applied by selecting users for representing the best and worst cases. Based on Fig. 4, User 32 was selected as the best case because the user achieved the criteria in the shortest time (one day) and lowest interactions (383 interactions). User 58 was selected as the worst case as they did not achieve the criteria even though the user had the longest time (20 days) with the highest interactions (42,624 interactions). The standard deviation was calculated for the main features (time, applications and URLs) as illustrated in Fig. 4. The figure shows that User 32 had a tidy pattern of usage, which could make the classifier more able to identify the user while User 58 did not seem to have consistent usage. These changes in user behaviour can have a negative effect on the performance of classifiers because their activities are so diverse.

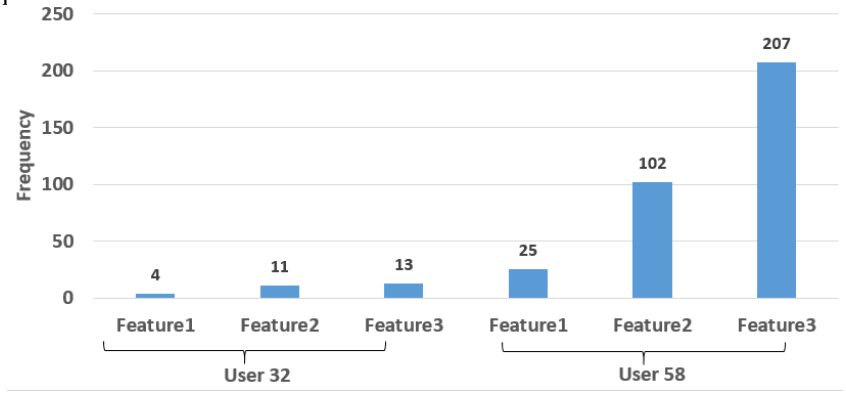

Fig. 4. Standard deviation of features' User 32 and 58

One reason for the worst case could be that part of the dataset was collected from early PhD research students who normally they use a variety of applications and websites during their initial research period. These variations and changes in user behaviour can affect negatively building an accurate picture of user usage pattern. Therefore, an additional mechanism is required to analyse the data deeply rather than relying on the time and volume of data alone to provide sufficient discriminatory information for creating these templates.

6 Conclusions

The results successfully demonstrate the ability to correctly distinguish users based on their interactions derived from a simulated cloud infrastructure service environment. Accurate user-behaviour profiles can be built to help distinguish between the normal and abnormal usage with high accuracy. Subsequently, the approach proved a highly

promising solution for applying user-behavioural profiling as a supporting technique to validate users after the initial point-of-entry authentication. This can contribute and guide the system to identify a misuse of cloud services in a continuously and friendly manner.

Future work will focus on developing mechanisms for understanding the nature of the user activities traffic more deeply in order to make sure appropriate user-behaviour profiles can be generated and when and how template renewal should be undertaken.

REFERENCES

1. Cisco. Cisco Global Cloud Index: Forecast and Methodology, 2016–2021 White Paper. (2016).
2. Abdallah, E. G., Zulkernine, M., Gu, Y. X. & Liem, C. TRUST-CAP: A Trust Model for Cloud-Based Applications. in *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)* **2,** 584–589 (IEEE, 2017).
3. Cloud Security Alliance. The Treacherous 12 Cloud Computing Top Threats in 2016. *Security* 1–34 (2016).
4. Chen, D. & Zhao, H. Data Security and Privacy Protection Issues in Cloud Computing. *2012 Int. Conf. Comput. Sci. Electron. Eng.* **1,** 647–651 (2012).
5. Walters, R. *Cyber Attacks on U.S. Companies in 2016*. *The Heritage Foundation: Issue Brief* **No. 4636,** (2016).
6. Cameron, D.: Apple knew of iCloud security hole 6 months before Celebgate. *The Daily Dot* (2014). Available at: https://www.dailydot.com/debug/apple-icloud-brute-force-attack-march/. (Accessed: 27th February 2018)
7. Danny Yadron.: Hacker collects 272m email addresses and passwords, some from Gmail | Technology | The Guardian. *Theguardian* (2016). Available at: https://www.theguardian.com/technology/2016/may/04/gmail-yahoo-email-password-hack-hold-security. (Accessed: 10th March 2018)
8. Hall, J., Barbeau, M. & Kranakis, E.: Anomaly-based intrusion detection using mobility profiles of public transportation users. *WiMob'2005), IEEE Int. Conf. Wirel. Mob. Comput. Netw. Commun. 2005.* **2,** 17–24 (2005).
9. Hilas, C., Kazarlis, S., Rekanos, I. & Mastorocostas, P.: A genetic programming approach to telecommunications fraud detection and classification. *Proc. 2014 Int. Conf. Circuits, Syst. Signal Process. Commun. Comput.* 77–83 (2014).
10. Ogwueleka, F.: Fraud Detection In Mobile Communications Networks Using User Profiling And Classification Techniques. *J. Sci. Technol.* **29,** 31–42 (2009).
11. Qayyum, S., Mansoor, S., Khalid, A., Halim, Z. & Baig, A. R.: Fraudulent call detection for mobile networks. *2010 Int. Conf. Inf. Emerg. Technol.* 1–5 (2010). doi:10.1109/ICIET.2010.5625718
12. Yazji, S., Dick, R. P., Scheuermann, P. & Trajcevski, G.: Protecting Private

Data on Mobile Systems based on Spatio-temporal Analysis. in (2011).
13. Yazji, S., Scheuermann, P., Dick, R. P., Trajcevski, G. & Jin, R.: Efficient location aware intrusion detection to protect mobile devices. *Pers. Ubiquitous Comput.* **18,** 143–162 (2014).
14. Subudhi, S. & Panigrahi, S. Quarter-Sphere Support Vector Machine for Fraud Detection in Mobile Telecommunication Networks. *Procedia Comput. Sci.* **48,** 353–359 (2015).
15. Shi, E., Niu, Y., Jakobsson, M. & Chow, R.: Implicit authentication through learning user behavior. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **6531 LNCS,** 99–113 (2011).
16. Dimitrios Damopoulos, Sofia A. Menesidou, Georgios Kambourakis, M. P. & Gritzalis, N. C. and S.: Evaluation of anomaly-based IDS for mobile devices using machine learning classifiers. *Secur. Commun. Networks* **5,** 3–14 (2012).
17. Li, F., Clarke, N., Papadaki, M. & Dowland, P.: Behaviour profiling on mobile devices. *Proc. - EST 2010 - 2010 Int. Conf. Emerg. Secur. Technol. ROBOSEC 2010 - Robot. Secur. LAB-RS 2010 - Learn. Adapt. Behav. Robot. Syst.* 77–82 (2010).
18. Li, F., Clarke, N., Papadaki, M. & Dowland, P.: Misuse Detection for Mobile Devices Using Behaviour Profiling. *Int. J. Cyber Warf. Terror.* **1,** 41–53 (2011).
19. Li, F., Clarke, N., Papadaki, M. & Dowland, P.: Active authentication for mobile devices utilising behaviour profiling. *Int. J. Inf. Secur.* **13,** 229–244 (2014).
20. Fridman, L., Weber, S., Greenstadt, R. & Kam, M.: Active Authentication on Mobile Devices via Stylometry, Application Usage, Web Browsing, and GPS Location. *IEEE Syst. J.* **11,** 513–521 (2017).
21. Aupy, A. & Clarke, N.: User Authentication by Service Utilisation Profiling. in *Advances in Network and Communications Engineering 2* 18 (School of Computing, Communications & Electronics, University of Plymouth, 2005).
22. Yazji, S., Chen, X., Dick, R. P. & Scheuermann, P.: Implicit user re-authentication for mobile devices. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **5585 LNCS,** 325–339 (2009).
23. Salem, M. Ben & Stolfo, S. J.: Modeling user search behavior for masquerade detection. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **6961 LNCS,** 181–200 (2011).
24. Yang, Y.: Web user behavioral profiling for user identification. *Decis. Support Syst.* **49,** 261–271 (2010).
25. Abramson, M. & Aha, D.: User Authentication from Web Browsing Behavior. *Twenty-Sixth Int. FLAIRS Conf.* 268–273 (2013).
26. Sola, J. & Sevilla, J.: Importance of input data normalization for the

application of neural networks to complex industrial problems. *IEEE Trans. Nucl. Sci.* **44,** 1464–1468 (1997).

27. Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q, Motoda H, McLachlan GJ, Ng A, Liu B, Philip SY, Zhou ZH.: *Top 10 algorithms in data mining*. *Knowledge and Information Systems* **14.** (2008).