Continuous Identity Verification in Cloud Storage Services using Behavioural Profiling

Burhan Al-Bayati[1, 2], Nathan Clarke[1, 3], Paul Dowland[3], Fudong Li[4]
[1]Centre for Security, Communications and Network Research, Plymouth University, Plymouth, UK
[2]Computer Science Department, Science College, Diyala University, Diyala, Iraq
[3]Security Research Institute, Edith Cowan University, Perth, Australia
[4]School of Computing, University of Portsmouth, Portsmouth, UK
burhan.al-bayati@plymouth.ac.uk
n.clarke@plymouth.ac.uk
p.haskelldowland@ecu.edu.au
fudong.li @ port.ac.uk

Abstract: Cloud storage services have become immensely popular because they enable users to remotely store their data over the Internet. However, this has led to a lack of physical control to protect their information with an increasing vulnerability to potential attacks. Well-known service providers including Dropbox and Apple iCloud have suffered from attacks, leading to sensitive customer information being exposed. A key issue is that the cloud services rely upon a simple authentication login and remain accessible to users afterward for significant periods of time. Thus, arguably more intelligent security measures are required to support the security of the system. Behavioural profiling is one technique that has been applied successfully with a variety of technologies for continuous user verification, telecommunication misuse and credit card fraud. However, the implementation of such a technique in cloud storage services has not been studied. This paper investigates the application in cloud storage services to detect misuse post initial login. A private dataset was collected from a cloud storage service (Dropbox) containing real user interactions of 30 participants over a six month period (totalling 91,371 log entries). A series of experiments have been implemented on the dataset using a supervised machine learning algorithms to examine the feasibility of classifying the normal and abnormal users' behaviour. On average, the best experimental result achieved an EER of as low as 5.8% with six users experiencing an EER equal to or less than 0.3%. The results are very encouraging and indicate the feasibility of detecting misuse in cloud computing services.

Keywords—Continuous identity verification; misuse; transparent; behavioural profiling; Dropbox; cloud computing services

1. Introduction

In the past few years, cloud computing has become a new paradigm for hosting and delivering services over the Internet. With the rapid developments in the information technology, the amount of digital information has dramatically increased. Cloud storage services have become particularly attractive for users (both individuals and enterprises) by offering data storage to meet differing levels of demand. Customers can upload, download, update, remove, and share data with directly accessing information through online web applications from anywhere at any time. The flexibility, accessibility, simplicity, efficiency, scalability and pay as you go that are offered by cloud providers has proven very successful (Forbes 2015). For example, Dropbox has more than 500 million users' accounts, 200,000 businesses and 1.2 billion files uploaded daily (Dropbox 2017).

There is no doubt that cloud storage services can provide a flexible and convenient way for customers to access their data. However, customers still have concerns about how to protect the data stored remotely in these services from unauthorized access, with reports suggesting it is the biggest barrier to the adoption of cloud computing services. For example, a survey carried out by Garter showed that more than 70% of CTOs believed that data security and privacy concerns were the main reason of hindering them of using the cloud service(Chou 2013). Also, studies conducted on users' data stored in the cloud, found that 88% of potential cloud customers were worried about their data security (e.g. who has access to their personal data) (Fujitsu 2010, Erin Griffith 2014).

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 access and misuse the service and user information. Many incidents have targeted popular cloud computing service providers, for example:

- Serious incidents on the Microsoft Azure cloud computing platform led to a massive collapse and outage of the service for 22 hours, with a loss of 45% of user data (Chen and Zhao 2012).
- Dropbox was hacked in July 2012; usernames and passwords of many users were stolen from third-party websites; hackers got access successfully to customers' accounts and misused their data (BBC News 2016).

- Many Apple iCloud accounts were compromised in 2014 as more than 20,000 passwords were stolen, resulting in user's personal photographs, specifically celebrities, being leaked online (Gupta 2015).
- Google's Gmail server faced attack in 2016; more than 272 million email addresses and passwords were stolen (Yadron 2016)

Cybercriminals can obtain access to sensitive information of cloud services even though security controls were in place and dedicated security teams were allocated. Therefore, additional security techniques are needed to protect cloud storage services from being compromised and misused. This paper proposes a novel continuous identity verification system that protects cloud storage service users' data by operating transparently to detect unauthorized access.

The remainder of the paper is structured as follows: The next section introduces the state of the art of using behavioural profiling to detect the anomaly usage within various technologies. Section 3 presents the experimental methodology. A series of comprehensive experimental studies to evaluate the applicability of using behavioural profiling with cloud storage service (Dropbox) 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

Author(s)	Activity	Client/Server	#Participants	Performance (%)	Method	Purpose
(Moreau et al 1997)	Telephony	Server	600	DR=90, FRR=10	Supervised Neural Networks	Fraud detection
(Burge and Shawe-Taylor 1997)	Telephony	Server	110	DR=75,FRR=40	Unsupervised Neural Network	Fraud detection
(Samfat and Molva 1997)	Mobility	Server	400	DR=82.5,FRR=40	Distance	Fraud detection
	Telephony			DR=80,FRR=30	Rule-base	
(Buschkes et al 1998)	Mobility	Server	None	DR=87.5	Bayes Decision Rule	IDS
(Sun et al 2004)	Mobility	Server	None	DR=87.5,FRR=15	High order Markov	IDS
(Sun et al 2006)	Mobility	Server	None	DR=89,FRR=13	High order Markov	IDS
(Hall et al 2005)	Mobility	Server	50	DR=50,FRR=50	Instance based learning	IDS
(Hilas and Sahalos 2005)	Telephony	Server	5000	DR=80	Statistical machine learning	Fraud detection
(Hilas and Sahalos 2007)	Telephony	Server	5000	DR=80	decision trees	Fraud detection
(Hilas et al 2014)	Telephony	Server	5000	DR=80	Genetic Programing method	Fraud detection
(Ogwueleka 2009)	Telephony	Server	180	FRR=3	self-Organizing Map and Probabilistic models	Fraud detection
(Qayyum et al 2010)	Telephony	Server	300	DR=70	Neural Network	Fraud detection
(Yazji et al 2011)	Mobility	Server	100	DR=81	cumulative probability and Marko properties of trajectories	IDS
(Yazji et al 2014)	Mobility	Server	178	DR=94	cumulative probability and Marko properties of trajectories	IDS
(Subudhi and Panigrahi 2015)	Telephony	Server	94	DR=97	SVM	Fraud detection
(Shi et al 2011)	Telephony, SMS,	Client	50	DR=95	Probability	Authentication

Author(s)	Activity	Client/Server	#Participants	Performance (%)	Method	Purpose
	Browsing, Mobility					
(Damopoulos et al 2012)	Telephony, SMS, Browsing	Client	35	DR=98.5,EER=1.6	Bayesian network , RBF, KNN, Random Forest	Authentication
(Li et al 2010)	Telephony, Device Usage, Bluetooth network scanning	Client	30	EER=13.5, 35.1,and 35.7	RBF network	Authentication
(Li et al 2011)	Application, Telephony, SMS	Client	76	EER=13.5, 2.2, 5.4	Neural network	Authentication
(Li et al 2014)	Application Usage	Client	76	EER=9.8	Rule base	Authentication
(Fridman et al 2017)	Text, App, Web and location	Client	200	EER=3	SVM	Authentication
(Aupy and Clarke 2005)	Way of using PC	Client	21	EER= 7	Neural Network (FF-MLP)	Authentication
(Yazji et al 2009)	File access activity and network event	Client	8	DR=90, FAR=14, FRR=11	K-Means Clustering	Authentication
(Salem and Stolfo 2011)	File access activity	Client	18	FAR=1.1	SVM	Insider detection
(Yang 2010)	Web Browsing	Server	100	DR=91	support-based, lift-based profiling	Identification
(Abramson and Aha 2013)	Web Browsing	Server	10	EER= 24	SVM	Authentication

*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, IDS: Intrusion Detection System

Early research focused mainly on IDS and fraud detection based on identifying the user behaviour activities during the interaction with mobile's services, such as calling and mobility (Moreau et al 1997, Burge and Shawe-Taylor 1997, Samfat and Molva 1997, Buschkes et al 1998 , Sun et al 2004, Sun et al 2006, Hall et al 2005, Hilas and Sahalos 2005, Hilas and Sahalos 2007, Hilas et al 2014, Ogwueleka 2009, Qayyum et al 2010, Yazji et al 2011, Yazji et al 2014, Subudhi and Panigrahi 2015). In comparison, more recent studies have focused on transparent authentication through modelling application usage to alleviate device misuse (Shi et al 2011, Damopoulos et al 2012, Li et al 2010, Li et al 2011, Li et al 2014, Fridman et al 2017). Much more information can be gathered from user activities while interacting with these applications (e.g. phone calls, GPS locations, SMSs, emails, websites visits, and calendar activities). These activities have been exploited to build an accurate behavioural profile which can be investigated to increase the accuracy level of the security system for the device or application itself.

There is a further studies focused on the generation of user behaviour profiles from desktop computer usage to detect any illegal access to the device (Yazji et al 2009, Salem and Stolfo 2011). 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, websites being visited. While the server side perspective, studies focused on building a user identifier by using their web surfing activities from numerous log files of websites (Yang 2010, Abramson and Aha 2013). 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. An accurate user behaviour profiles have been built to detect illegitimate usage.

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 utilizes the behavioural profiling has been studied regarding cloud storage services.

3. Experimental methodology

The aim of this study is to focus upon understanding to what degree behaviour profiling can be used to verify individuals within cloud storage services – understanding whether it is the legitimate user or not provides a basis for the system to respond. Therefore, a series of experiments were conducted on users of cloud storage services to examine different factors that can affect the performance of the classification algorithms. These include:

- Investigate the nature of different classification approaches to explore how the performance is affected.
- Explore the impact of the volume of data for training and testing on the performance of the system.
- Understand the effect of time series rather than random sample selection on the accuracy of a decision.

In order to conduct these experiments, users' activities within the cloud storage service are needed, in terms of both quality and quantity. Due to the privacy and security concerns, it is challenging to get a workable dataset from cloud storage providers. Also, to the best of the author's knowledge, no public dataset on user's cloud activities is available. Dropbox, Google Drive, One Drive and Box are all examples of widely popular cloud storage services. Dropbox was chosen for this research as it is one of the most popular cloud storage services (Erin Griffith 2014, CloudRAIL 2017) and importantly it provides simply access to users' interactions records. This made is possible to access and capture logs of Dropbox activity over prolonged periods of time. By downloading user's historical activities, a private dataset that contains 91,371 unique interactions from 30 users over a six month period was obtained. For each user interaction, the following information is available: timestamp of the action (day, hour, and minute), the file type (e.g. pdf, jpg, and docx), user's action (i.e. add, edit, delete, move, and rename).

In order to make those features acceptable by classification algorithms, the symbolic-valued attributes (e.g. file type and user action) were enumerated into numerical attributes and into the range of 0-1 (Sola and Sevilla 1997).

The records of each user were divided into two sets: the first set was used to generate a profile for training; while the second set was used to evaluate the classifiers' performance (referred to as the test set). Classification is based upon a 2-class problem; legitimate or impostor, where one user acts as the legitimate user, with the remaining users acting as impostors. This is then repeated to ensure all users have the opportunity to act as an authorised user. The corresponding False Acceptance Rate (FAR), False Reject Rate (FRR) are computed and are used to determine the Error Equal Rate (EER). The Equal Error Rate (the point at which the FAR and FRR are equal) is used as the key performance metric.

The first experiment explored how the performance of the system is affected by investigating the nature of different classification approaches. The findings of this experiment would also help to identify the optimal classifier. Four supervised machine learning algorithms were selected: Feed-Forward Multi-Layered Perceptron (FF MLP), Random Forest (RF), Support Vector Machine (SVM) and Classification And Regression Trees (CART). The first three approaches were selected based on the highest performance that achieved with the previous studies (as illustrated in Table 1) whereas the fourth method was based on the study by (Wu et al 2008) that conducted on different classification algorithms. Also default configurations are used for all selected classifiers. The experiment utilised a 66/34 splitting for the training and testing data with random selection across the dataset. At no point is a sample used for both training and testing.

The second experiment focuses on investigating the impact of training and testing data split upon the performance. In addition to the 66/34 split, 50/50 and 80/20 splitting approaches were investigated for training/testing with random sample selection. Regarding the classifier, the classification algorithm that achieved the best performance from the first experiment was selected. The comparison between the accuracy of the result given differing levels of training data would provide a better understanding of the nature of user behaviour profiles and the volume of data necessary to achieve an appropriate level of performance. The first two experiments sought to random sample data across the dataset, in order to understand the general feasibility of the approach. However, in practice a profile would need to be created based upon time-series (i.e. those it collected first rather than samples it has yet to capture). As such, the third experiment sought to evaluate the impact of applying time series on the performance. In order to understand the effect of the two metrics on the performance, the similar volume of data for the training and testing sets of the previous experiment is applied. The accuracy of each volume is compared with the accuracy of the volume of the previous experiment.

4. The Experimental Results

4.1 Classification Algorithms

The overall results of this experiment are presented in Table 2. Generally speaking, the results are very encouraging to support the idea of verifying the legitimate user or unauthorised access to data stored in cloud storage services, with EERs that are aligned to similar results in other applications from the prior work.

Table 2: Performance of classification algorithms

Classifier	EER (%)
SVM	20.27
RF-25 trees	9.93
FF MLP Neural Network-65	6.98
CART	6.02

The nature of classifier utilised does have an impact; however, with the exception of SVM, the variation in performance is not overly significant – suggesting the classifier itself is not overly key. As seen in Table 2, the CART algorithm achieved the highest accuracy amongst the others with an EER of 6.02%. This would allow other factors such as time taken to compute, computational overhead, memory requirements to be considered as part of the selection.

A more detailed analysis of the classifiers was undertaken, to determine what impact optimisation would have. The results from the FF MLP and RF methods are demonstrated in Figure 1 and Table 3 respectively. For the FF MLP classifier, the best result of EER 6.98% was achieved by using 65 neurons; while with RF approach, the best performance of EER 9.93% was obtained when 25 trees was used. Neither SVM or CART had any parameters to optimise.

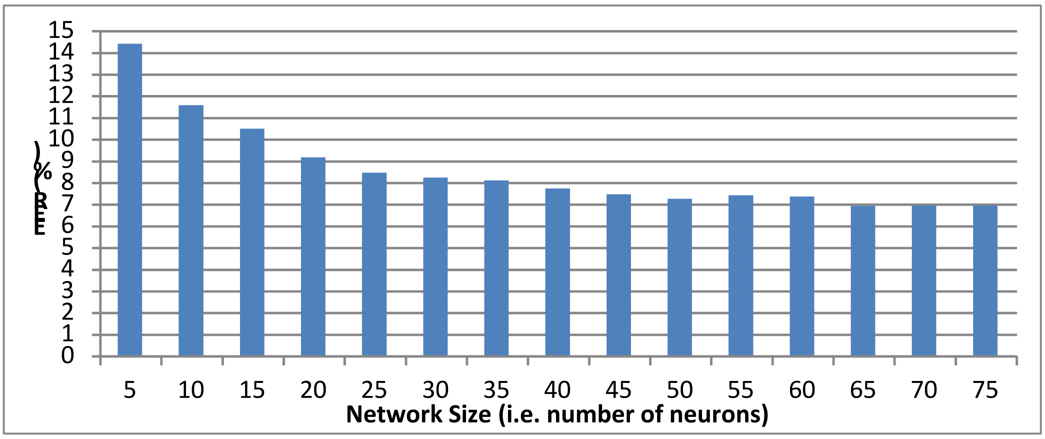

Figure 1: Performance of FF MLP with different network configurations

Table 3: Performance of RF with trees

Number of trees	EER (%)
5	10.39
10	10.41
15	10.21
20	10.18
25	9.93
30	10.26
35	10.43

Prior research has shown the volume of data per user has a significant impact upon performance. As such, an analysis was performed. Users were divided into two groups based on their interactions, as illustrated in Table 4. The first 15 users belong to the users' group who have less interactions (i.e. equal to or less than 2000 interactions) whereas the remaining 15 users belong to the users' group who have more interactions (i.e. more than 2000 interactions). The selection of 2000 was felt sufficient to separate the groups yet ensure a suitable number of participants were left in each group. Based upon the overall average performance from these two groups, it shows that users who have more interactions achieved a better performance than the users with less interactions by using the RF, FF MLP and CART classifiers.

However, this was not always true on a per user basis. For example, although Users 17, 18 and 21 have more interactions than many other users, they achieved a low performance than many users with low interactions such as Users 2, 4, 9, and 13. Further investigation suggests that those three users used Dropbox as a backup solution by uploading photos within a period of time which can be carried out automatically by a computer rather than the users themselves, creating difficulty for the classifiers to differentiate between user actions and computer generated activities. Similarly, some low usage users got a better accuracy than many more active users. For instance, User 2 achieved less than 2% of EER across the most approaches and User 13 got an EER close to zero. When looking to the usage of these users, it is found that they worked constantly on specific file types that majority of the rest users do not use. This unique pattern of usage made the classifier more able to discriminate them from others. Based upon the result, it suggests that users who have more interactions achieve better performance in general. However, the uniqueness of interactions can be a key factor to build discriminative patterns for users, which can make classifiers more accurate to distinguish between them.

Table 4: Users' performance with different classifiers

User	No. of Interactions	EER % based on classifier algorithms			
		SVM	RF	FF MLP	CART
1	549	19.08	8.88	10.32	5.13
2	585	2.11	3.47	1.95	1.96
3	652	6.23	9.68	4.65	3.35
4	677	2.52	1.56	1.43	2.02
5	726	26.80	6.55	7.21	3.63
6	764	23.79	8.17	13.09	5.71
7	797	3.02	28.29	4.78	14.89
8	1146	23.76	36.09	15.75	23.73
9	1370	10.67	2.36	4.70	1.22
10	1413	31.49	6.17	3.86	3.86
11	1462	13.47	10.08	4.89	5.36
12	1656	25.14	33.00	12.49	16.64
13	1714	11.25	0.11	0.22	0.02
14	1765	32.30	18.75	16.84	10.95
15	1988	39.35	19.86	15.02	15.56
Av*	**1,150,933**	**18.06**	**12.87**	**7.81**	**7.6**
16	2250	30.22	4.69	7.81	4.08
17	2373	43.34	20.13	13.08	12.63
18	2487	28.53	19.87	9.13	12.27
19	2799	10.22	2.71	4.84	1.26
20	2879	17.54	0.56	2.40	0.54
21	2960	20.02	25.05	8.34	14.38
22	3226	28.17	1.33	3.57	0.99
23	3464	3.08	1.26	1.54	0.53
24	3568	31.55	3.95	13.88	1.91
25	4858	25.28	3.94	4.30	3.34
26	5780	7.13	0.15	0.19	0.15
27	6440	13.21	1.65	1.85	1.37
28	7263	29.06	11.20	7.78	5.89
29	14985	29.99	7.69	10.79	6.64
30	15013	19.81	0.68	2.79	0.75
Av**	**5,356,333**	**22.48**	**7.00**	**6.15**	**4.45**

Av*: Average from the first group, Av**: Average from the second group

4.2 Volume of Data for Training and Testing

This experiment studied the impact of the volume of data for training upon the performance. The CRT classifier was chosen for this experiment due to its (best) performance from the first experiment; also the data splitting between training the classifier and testing the performance was set to 50/50, 66/34 and 80/20. Table 5 illustrates the performance of all users across the selected volumes of data.

Table 5: Performance based on volume of data with random selection

User	Volume of data		
	50/50	66/34	80/20
1	7.28	5.13	5.09
2	2.59	1.96	1.65
3	7.30	3.35	4.26
4	1.84	2.02	1.14

User	Volume of data		
	50/50	66/34	80/20
5	3.66	3.63	2.05
6	6.57	5.71	4.96
7	15.20	14.89	14.14
8	25.25	23.73	24.45
9	2.14	1.22	0.84
10	7.06	3.86	5.39
11	6.47	5.36	5.32
12	20.04	16.64	16.04
13	0.09	0.02	0.02
14	9.92	10.95	9.77
15	13.81	15.56	14.50
16	3.88	4.08	2.19
17	13.58	12.63	11.93
18	13.93	12.27	12.98
19	1.83	1.26	1.38
20	0.59	0.54	0.70
21	16.63	14.38	14.36
22	1.22	0.99	1.35
23	0.90	0.53	0.89
24	2.33	1.91	1.73
25	3.35	3.34	3.51
26	0.21	0.15	0.17
27	1.56	1.37	1.80
28	7.15	5.89	6.04
29	6.63	6.64	6.39
30	0.85	0.75	0.77
Average	**6.79**	**6.02**	**5.86**

As shown in Table 5, the training phase with a larger volume of samples achieves better performance than those with a smaller volume of data on average; the best result performance was 5.86% of EER achieved by using 80/20 splitting for training and testing respectively. This agrees with the prior research and suggests that larger volume of samples for training the classifier can have a positive impact on the overall performance. This is logical as the classifier can be trained more about user behaviour pattern by using a larger volume of data, leading to a better performance. However, it is also worth highlighting that the change in performance from 6.79% to a best case of 5.86% is not significant. This suggests that the nature of user behaviour across the 6 month collection period is likely to be relatively stable.

From an individual user's perspective, the increasing volume of data for the training stage has different impacts upon the performance. When increasing the training data volume to 66/34 and 80/20 splitting, a number of users performance improved and some stayed relatively stable - suggesting more data made little difference. In a practical sense, being able to understand which users have more stable or active profiles would be very useful in interpreting the classification decisions and in template retraining.

4.3 Time Series Sample Selection

In addition to the random sample that is used for the previous experiment (a standard methodological approach in feasibility studies), the impact of the time and natural changes in user behaviour over time is important to evaluate. The CRT classifier was with the data split for training and testing is the same manner as the pervious experiment (i.e. 50/50, 64/34 and 80/20). The results of the experiment are presented in Table 6.

Table 6: Performance of the different volume of data with time series selection

User	Volume of data		
	50/50	66/34	80/20
1	18.38	15.86	15.15
2	2.61	3.84	4.28
3	6.49	6.22	6.41
4	2.92	1.18	0.04
5	18.28	25.05	21.06

User	Volume of data		
	50/50	66/34	80/20
6	2.43	9.85	3.54
7	19.90	17.24	19.45
8	41.39	40.57	43.59
9	2.36	3.15	2.80
10	27.60	38.91	23.38
11	10.07	8.89	4.71
12	40.31	38.13	36.66
13	15.69	19.24	14.40
14	17.32	17.05	19.60
15	27.37	21.14	23.83
16	2.22	1.37	1.13
17	34.52	32.90	35.51
18	25.87	19.43	15.45
19	2.10	1.61	1.15
20	8.54	4.51	5.32
21	25.10	22.64	19.03
22	1.74	2.30	1.89
23	1.90	0.96	0.82
24	6.27	5.34	4.57
25	4.92	5.09	6.32
26	0.11	0.19	0.31
27	3.28	1.91	2.29
28	11.28	11.05	9.74
29	8.77	8.35	10.22
30	0.81	0.83	0.85
Average	**13.02**	**12.83**	**11.78**

As demonstrated by Table 6, the best performance is EER 11.78% and it is achieved by using the 80/20 data splitting for training and testing. Similarly to experiment 2, the nature of the data split has not had a significant impact upon performance; however, the results themselves have doubled. This suggests that over-time user behaviour does change and therefore care must be taken on ensuring appropriate template renewal procedures are developed to maintain levels of performance.

5. Discussion

The experimental results reveal that cloud storage service users can be discriminated via their usage with a reasonable performance being achieved. Also, the outcome of this research is in line with the highest results that are achieved in the related works such as (Shi et al 2011, Aupy and Clarke 2005, Yazji et al 2014, Subudhi and Panigrahi 2015). In terms of the performance of each individual classifier, the CRT algorithm achieves 6.02% EER and outperforms the other three chosen classifiers (i.e. SVM, RF and FF MLP). From an individual users' perspective, on average users who have more frequent activities/interactions acquired better results than those who have the less interactions across most classification. However, users with fewer interactions did also achieve a good level of performance. For example, when examining the interactions of those users (e.g. Users 4, 9, and 13), they have a unique way of using Dropbox (particularly unique file types). Therefore, a good pattern (uniqueness) from the users' interactions can also affect the performance of the classifiers even though the number of users' activities is low.

The results of the second and third experiments show that the data split for training and testing the classifier and the timestamp factors have an impact upon the overall performance. As shown in these two experiments, a larger volume training data (i.e. 80/20 splitting) with random sample selection achieves better performance with 5.8% of EER on an average. However, regarding individual users, the performance for a number of users with more training data (i.e. 66/34 and 80/20) are not as good as the results being achieved by using less training data (i.e. 50/50 split for training and testing). One of the reasons could be that part of the dataset was collected from early stage PhD research students, and normally they conduct various activities and use different file types during the initial research period. Therefore, they might deal with specific files types and actions within the first period of their research, then other file types and actions with the next period. These changes in user's behaviour can affect the performance of classifiers because their activities are so diverse.

When applying the behaviour profiling technique in practice, the time series sample selection showed a significant difference over the random sampling. Therefore, user's templates needs to be updated regularly to ensure its quality for achieving a high level of system performance. However, the renewal of users' templates dynamically is not an easy task because it might need to avoid including impostor's behaviour with the legitimate behaviour. For example, an impostor might be accepted by the system over time as the genuine user as more and more impostor samples are included within the template renewal process. This problem needs to be managed carefully and correctly to avoid capturing of illegitimate usage, at the same time ensuring a user convenient level exists in the system for the legitimate user comparison for sample selection techniques

6. Conclusion & Future Work

The results have successfully demonstrated the ability to correctly discriminate between users based on their interactions derived from the cloud storage (Dropbox). Accurate user behaviour profiles can be built to help in distinguishing between the normal and abnormal usage. Classification algorithms experiments achieved a high accuracy with only the SVM not performing particularly well. Further experiments have shown that time-series versus random sampling of data for training does have a significant impact upon performance; however, the volume of training data less so. From individual's performance, many of participants achieved a very high performance where the system was capable to identify their interactions fully correct without any error. Subsequently, the approach proved a highly promising solution to applying user behavioural profiling as a supporting technique to validate the users after initial point-of-entry authentication. This can contribute and guide the system to identify a misuse of cloud services in continuously and friendly manner. However, there were a number of users who performed particularly poorly and in line with most behavioural-based applications, would not be suited to such a technique.

Future work will focus upon developing mechanisms for understanding where and when such an approach can be utilised (i.e. enabling the approach for users with a sufficiently stable profile) and when and how template renewal should be undertaken.

References

Abramson, M. and Aha, D.W., 2013, May. User Authentication from Web Browsing Behavior. In FLAIRS conference (pp. 268-273).

BBC News, 2016. Dropbox hack 'affected 68 million users. BBC News Web page. Available at: http://www.bbc.co.uk/news/technology-37232635 (accessed 17/10/17).

Burge, P. and Shawe-Taylor, J., 1997. Detecting cellular fraud using adaptive prototypes. Proc. AI Approaches to Fraud Detection and Risk Management, pp.9-13.

Buschkes, R., Kesdogan, D. and Reichl, P., 1998, December. How to increase security in mobile networks by anomaly detection. In Computer Security Applications Conference,1998. Proceedings.14th Annual (pp. 3-12). IEEE.

Chen, D. and Zhao, H., 2012, March. Data security and privacy protection issues in cloud computing. In Computer Science and Electronics Engineering (ICCSEE), 2012 International Conference on (Vol. 1, pp. 647-651). IEEE.

Chou, T.S., 2013. Security threats on cloud computing vulnerabilities. International Journal of Computer Science & Information Technology, 5(3), p.79.

CloudRAIL, 2017. Cloud Storage Report 2017 - Dropbox Loses Market Share But is Still the Biggest Provider on Mobile - CloudRail. . Available at: https://blog.cloudrail.com/cloud-storage-report-2017/ (accessed 19/11/17).

Danny Yadron, 2016. Hacker collects 272m email addresses and passwords, some from Gmail | Technology | The Guardian. Theguardian. Available at: https://www.theguardian.com/technology/2016/may/04/gmail-yahoo-email-password-hack-hold-security (accessed 25/11/17).

Damopoulos, D., Menesidou, S.A., Kambourakis, G., Papadaki, M., Clarke, N. and Gritzalis, S., 2012. Evaluation of anomaly-based IDS for mobile devices using machine learning classifiers. Security and Communication Networks, 5(1), pp.3-14.

Dropbox, 2017. About - Dropbox. Available at: https://www.dropbox.com/about (accessed 07/11/17).

Erin Griffith, 2014. Who's winning the consumer cloud storage wars?. Available at: http://fortune.com/2014/11/06/dropbox-google-drive-microsoft-onedrive/ (accessed 20/11/17).

Forbes, 2015. Roundup Of Cloud Computing Forecasts And Market Estimates, 2015. . Available at:

http://www.forbes.com/sites/louiscolumbus/2015/01/24/roundup-of-cloud-computing-forecasts-and-market-estimates-2015/ (accessed 27/04/15).

Fridman, Lex, Steven Weber, Rachel Greenstadt, and Moshe Kam. "Active authentication on mobile devices via stylometry,application usage, web browsing,and GPS location." IEEE Systems Journal 11, no. 2 (2017): 513-521.

Fujitsu, 2010. Personal data in the cloud: A global survey of consumer attitudes. Available at: http://www.fujitsu.com/downloads/SOL/fai/reports/fujitsu_personal-data-in-the-cloud.pdf.

Gupta U, 2015. Survey on security issues in file management in cloud computing environment. 2(11): 5. Available at: http://arxiv.org/abs/1505.00729.

Hall, J., Barbeau, M. and Kranakis, E., 2005, August. Anomaly-based intrusion detection using mobility profiles of public transportation users. In Wireless And Mobile Computing, Networking And Communications, 2005.(WiMob'2005), IEEE International Conference on (Vol. 2, pp. 17-24). IEEE.

Hilas, C.S., Kazarlis, S.A., Rekanos, I.T. and Mastorocostas, P.A., 2014. A genetic programming approach to telecommunications fraud detection and classification. In Proc. 2014 Int. Conf. Circuits, Syst. Signal Process. Commun. Comput (pp. 77-83).

Hilas, C.S. and Sahalos, J.N., 2007, September. An application of decision trees for rule extraction towards telecommunications fraud detection. In International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (pp. 1112-1121). Springer, Berlin, Heidelberg.

Hilas, C.S. and Sahalos, J.N., 2005, October. User profiling for fraud detection in telecommunication networks. In 5th International conference on technology and automation (pp. 382-387).

Li, F., Clarke, N., Papadaki, M. and Dowland, P., 2010, September. Behaviour profiling on mobile devices. In Emerging Security Technologies (EST), 2010 International Conference on (pp. 77-82). IEEE.

Li, F., Clarke, N., Papadaki, M. and Dowland, P., 2011. Misuse detection for mobile devices using behaviour profiling. International Journal of Cyber Warfare and Terrorism (IJCWT), 1(1), pp.41-53.

Li Li, F., Clarke, N., Papadaki, M. and Dowland, P., 2014. Active authentication for mobile devices utilising behaviour profiling. International journal of information security, 13(3), pp.229-244.

Moreau, Y., Verrelst, H. and Vandewalle, J., 1997. Detection of mobile phone fraud using supervised neural networks: A first prototype. Artificial Neural Networks—ICANN'97, pp.1065-1070.

Ogwueleka, F.N., 2009. Fraud detection in mobile communications networks using user profiling and classification techniques. Journal of Science and Technology (Ghana), 29(3).

Qayyum, S., Mansoor, S., Khalid, A., Halim, Z. and Baig, A.R., 2010, June. Fraudulent call detection for mobile networks. In Information and Emerging Technologies (ICIET), 2010 International Conference on (pp. 1-5). IEEE.

Salem, M. and Stolfo, S., 2011. Modeling user search behavior for masquerade detection. In Recent Advances in Intrusion Detection (pp. 181-200). Springer Berlin/Heidelberg.

Samfat, D. and Molva, R., 1997. IDAMN: an intrusion detection architecture for mobile networks. IEEE Journal on Selected Areas in Communications, 15(7), pp.1373-1380.

Shi, E., Niu, Y., Jakobsson, M. and Chow, R., 2010, October. Implicit Authentication through Learning User Behavior. In ISC(Vol. 6531, pp. 99-113).

Sola, J. and Sevilla, J., 1997. Importance of input data normalization for the application of neural networks to complex industrial problems. IEEE Transactions on Nuclear Science, 44(3), pp.1464-1468.

Subudhi, S. and Panigrahi, S., 2015. Quarter-Sphere support vector machine for fraud detection in mobile telecommunication networks. Procedia Computer Science, 48, pp.353-359.

Sun, B., Yu, F., Wu, K. and Leung, V., 2004, October. Mobility-based anomaly detection in cellular mobile networks. In Proceedings of the 3rd ACM workshop on Wireless security(pp. 61-69). ACM.

Sun, B., Chen, Z., Wang, R., Yu, F. and Leung, V.C., 2006, January. Towards adaptive anomaly detection in cellular mobile networks. In Consumer Communications and Networking Conference, 2006. CCNC 2006. 3rd IEEE (Vol. 2, pp. 666-670).

IEEE.

Wu X, Kumar V, Ross QJ, Ghosh J, Yang Q, Motoda H, McLachlan GJ, Ng A, Liu B, Yu PS, Zhou ZH, Steinbach M, Hand DJ and Steinberg D., 2008. Top 10 algorithms in data mining. Knowledge and Information Systems. London: Springer.

Yang, Y., 2010. Web user behavioral profiling for user identification. Decision Support Systems 49(3): 261–271.

Yazji, S., Chen, X., Dick, R.P. and Scheuermann, P., 2009, July. Implicit User Re-authentication for Mobile Devices. In UIC (pp. 325-339).

Yazji, S., Dick, R.P., Scheuermann, P. and Trajcevski, G., 2011, September. Protecting Private Data on Mobile Systems based on Spatio-temporal Analysis. In PECCS (pp. 114-123).

Yazji, S., Scheuermann, P., Dick, R.P., Trajcevski, G. and Jin, R., 2014. Efficient location aware intrusion detection to protect mobile devices. Personal and Ubiquitous Computing, 18(1), pp.143-162.