Academic Science Journal

Continuous User Verification in Cloud Storage Services based on Deep Learning

Burhan Al-Bayati

Computer Science Department, Science College, Diyala University, Diyala, Iraq

burhan@uodiyala.edu.iq

Received: 1 November 2024 **Accepted: 30 December 2024**

Published: January 2025

DOI: https://dx.doi.org/10.24237/ASJ.03.01.950A

Abstract

Cloud storage services have become a new paradigm for storing data user over the Internet. Users can access these services using a simple authentication login. This has led to increase the potential attacks and vital customer information being misused. Behavior profiling technique has been successfully investigated as an additional intelligent security measures for continuous verification users after the simple login. A credible accuracy has been achieved when applying the technique in a various of applications such as telecommunication, credit cards and cloud services to detect and monitor misuse. To increase the accuracy of making a reliable decision, this paper proposes a system by adopting CNN (Convolutional Neural Network) and LSTM (Long Short-Term Memory) deep learning algorithms and combining two private datasets which are real-life user interactions with the desktop computer and Dropbox Cloud storage. The best experimental results showed an EER (Equal Error Rate) of 3.6% based on adopting CNN deep learning algorithm. This result indicates and encourages the feasibility of using behavioral profiling to protect cloud users from misuse.

Keywords: Continuous User Verification; Transparent verification; users' Behavioural Profiling; Cloud storage Services; Combining dataset.

Introduction

There are more than 3.6 billion cloud customers in the world which is around 47% of the global population (nearly the half of people on the earth utilize cloud services) [1]. Cloud storage services have become widely popular, particularly by offering big storage for their customers (both individuals and companies). They can easily access these services from anywhere at any time through the Internet. The flexibility of directly accessing information, uploading, downloading, and updating user documents made customers more attractive to join these services. For instance, Dropbox has more than 700 million registered customers, they are uploading billions of documents in daily usage [2].

There is no doubt about the scalability, accessibility, flexibility, efficiency, simplicity, and pay as you go that are offered by cloud storage services to their customers to access their data easily. Therefore, users still have a vital concern about security issues, which is how they can keep their data that is stored distantly in these services from illegal access, hackers can get access and abuse the services by stealing customers' login credentials. Many popular cloud computing service providers have been targeted in numerous incidents of sensitive customer information being misused. For instance, the Cloud Security Alliance reports that in 2014 and 2015, a number of security incidents that affected TalkTalk, a British telecom provider, caused in the expose of personally identifiable data of approximately 4 million of TalkTalk's users [3]. The cloud computing platform Microsoft Azure experienced serious difficulties that led to a catastrophic collapse and an outage of the service for a period of 22 hours, during which time 45% of users' data was lost [4]. In July of 2012, Dropbox was the target of a cyberattack. The usernames and passwords of a large number of customers were taken from third-party websites. Hackers successfully gained access to client accounts and exploited the data they contained [5]. In 2014, more than 20,000 passwords to Apple iCloud accounts were obtained, which led to the exposure of users' personal images, particularly those of celebrities, on the internet. This compromised a large number of Apple iCloud accounts [6]. In 2016, an attack was launched against Google's Gmail service, which resulted in the theft of about 272 million passwords and email addresses [7].

More recently, in telecommunications cloud services in August 2024, two significant data breaches were recorded, one in March, 60 million records accessed by cybercriminals. Moreover, in the same year, in August 40 million records were compromised [8]. Additionally, in April 2024, 2.9 billion records were affected by a potential cyberattack that exposed sensitive data of customers of cloud-based data storage across the U.K. Canada, and the U.S. [9].

From the above incidents, even when security controls have been implemented and dedicated security teams have been assigned, it is still possible for cybercriminals to access critical information that is stored in cloud services. Because of this, more security measures are required to safeguard cloud storage services from being hacked and exploited. In this research, a unique continuous identity authentication system for protecting the data of users of cloud storage services is proposed. The system would function invisibly to identify any instances of unwanted access. Users' identities can be continuously and transparently evaluated through the use of behavioural profiling while they are interacting with cloud storage services. Through the creation of user behaviour profiles, it is possible to recognise individuals based on the manner in which they engage with the aforementioned services. Therefore, the actions of the present user (such as the time at which the service was initially opened, for example) are compared with a template for an earlier user. These templates are produced by applying a machine learning technique, such as neural networks, to historical usage data in order to derive new configurations. The outcome of the comparison will indicate whether or not the currently logged-in user is valid.

The remaining structure of the paper is organized as follows: Section 2 introduces a literature review of using user behavioural profiling to identify abnormal usage within a variety of technologies. Section 3 describes the experimental methodology. Section 4 presented a number of the comprehensive practical experiments that were studied and examined the ability to use user behavioural profiling of cloud storage services (e.g., Dropbox). Section 5 discusses the effect of the experimental outcomes. The conclusion and future work of the proposed study are introduced in Section 6.

Related Works

From security perspectives, a number of studies have investigated behavioural profiling to identify abnormal usage. Some of these studies focused on fraud and Intrusion Detection (ID) for telecommunication and credit cards, while the authentication aspect was implemented with personal computer systems, web browsers and modern mobile phones.

Early studies on mobile services invoked mainly fraud and ID (Intrusion Detection) through building user behavioural profiling templates during user interactions with these services. Telephony and mobility activities are used to build the user's templates, which are utilised later to identify anomalous usage, see [10] [11] [12] [13] [14] [15] as well as [16, 17]. These studies focused on the accuracy (DR: Detection rate) and FAR (False Accepting Rate). The detection rate was between 80% and 90%, whereas the false accepting rate was between 3% and 50%. To detect abnormal usage, it should focus on two factors which are the false accepting rate and False Rejecting Rate (FRR), then calculate the Equal Error Rate (EER) in order to show the robustness of the system from both sides. Therefore, these studies can be considered as missing some important factors to evaluate the proposed system.

However, the new studies have focused on verifying user usage transparently based on modelling application usage to detect a misuse in devices [18] [19] [20] [21] [22] [23]. The studies gathered a hug amount of information about such as: GPS locations, emails, website visits, and calendar activities during users' interactions with applications of these devices. From these users' interactions, accurate behavioural profiles are created to increase the performance of the security system for the applications of the devices. The researchers got an accuracy of EER (Equal Error Rate) between 3% and 13% approximately. Another researchers created user behaviour profiles from personal computer usage and log files of websites to identify any abnormal access to their devices [17] [24, 25]. Users' templates are built based on a number of features, including applications used, access time of computer files, names of websites that have been visited, number of pages, starting time, and duration time of sessions. The accuracy result of EER was approximately 7.1%. From perspective of cloud computing services, a number of studies investigated user behavioural profiles to detect illegitimate usage [26], [27], and [28].

The two nearest studies related to the proposed study are: One focused on collecting users' interactions with the Dropbox cloud service. Another study constructed on building users' templates usage during user access to cloud infrastructure as a service (IaaS). The real dataset was collected for the Dropbox experiment, and a number of machine learning algorithms were implemented. The best accuracy result of this study was 5.8% of EER. While the second experiment of IaaS was a mirror dataset, which is not real [5].

As demonstrated by the state of the art, detecting the anomaly usage within various technologies such as mobile phones, and personal computers based on behavioural profiling has been applied successfully. This can improve the security level of any system to verify unauthorised usage or misuse. As the accuracy of the system is vitally important to increase the level of the correct decision, the author will combine two datasets using two selected a deep learning algorithms in order to increase the accuracy of the existing study.

Experimental Methodology

The main target of this work is to understand what degree using user behaviour profiling can be contributed to identify users that deal with the storage of cloud services whether they are legitimate or not. This can also lead to provide a basis understanding of the system that potential misuse might have occurred. Therefore, various factors are examined through a number of practical experiments investigated on users of cloud storage services in order to check the impact of these factors on the performance of the algorithms of deep learning. These include:

- Investigate two deep learning algorithms to understand the impact of these algorithms on the performance of the suggested system.
- Examine each dataset separately using the algorithm that provided better results in order to record the performance of each dataset.
- Explore the impact of combining the features of the two given datasets on the accuracy of the system.
- The effect of the data volume for training and testing on the accuracy of a decision of the proposed model.

Microsoft One Drive, Google Drive, IDrive, and Dropbox are the best examples of popular cloud storage services. However, it is difficult to get or collect users' activity from cloud storage providers because of the security and privacy concerns. Additionally, to the best of the author's knowledge, there is no dataset of user activities on cloud storage that is available publicly. Dropbox has been selected as cloud storage as a service in order to collect a real dataset. The main reason choosing Dropbox to collect the dataset is that it is one of the wide cloud storage services. Additionally, the more important reason is that it can simply access users' interactions logs. Therefore, from the account of each user, it is possible to collect the users' usage by downloading the user's historical activities. Also, the dataset was collected from 20 participants during 6 months' usage which contains totally about 70,481 users' interactions. This dataset has the following descriptions:

- Timestamp of each interaction (for example minutes, hours, and days).
- Type of documents (such as doc, bmp, and pdf).
- Type of each interaction (such as upload, move, rename, and delete).

The second dataset was collected from the personal computer of the same participants over the same period of time by installing software that works in the background of the computer OS in a transparent manner. The total users' usage of this dataset contains 90,595 users' interactions. The dataset contains the following information:

- Timestamp of each interaction (such as minutes, hours, and days).
- Name of application/URL (such as Word, MatLab, Google).
- Event (focus or not focus).

Therefore, from these two datasets, they can be combined into one dataset. Tables 1, 2, and 3 show an example of both datasets and how they can be combined in one dataset.

Table 1: User interactions with Dropbox

Second	Minute	Hour	Day	File Type	Activity
20	9	8	3	docx	Edit
33	10	8	3	pdf	Upload
20	18	8	3	jpg	Delete

Table 2: User interactions with personal computer

Second	Minute	Hour	Day	Apps/URLs	Event
12	9	8	3	Excel	Focus
30	13	8	3	Matlab	Lost focus
44	20	8	3	Google	Focus

Table 3: Combine the two datasets

Second	Minute	Hour	Day	File Type/ Apps/URLs	Activity/Event
12	9	8	3	Excel	Focus
20	9	8	3	docx	Edit
30	10	8	3	pdf	Upload
33	13	8	3	Matlab	Lost focus
20	18	8	3	jpg	Delete
44	20	8	3	Google	Focus

In order to make the above feature acceptable by the deep learning algorithms, each symbolic-valued attribute (such as file type, application name, or URL) was changed to the numerical attributes, which took a range between 0 and 1 [29].

The user's dataset was separated into two stages: the first one was used to generate a user's template for training; whereas the remaining dataset (test stage) was implemented to examine the accuracy of classifiers. This work mainly focused on identifying the normal usage and impostor. So, the type of research problem is a 2-class problem. Due to the impostor dataset being unavailable, one user will act as a legitimate user, while all other remaining users will act as impostor users. This procedure will be repeated for all users in order to make sure that all customers have the same chance to detect which one is the authorised user or not. This needs to compute the three important factors, which are: FAR (False Acceptance Rate), FRR (False Reject Rate), and the average of these two factors will be the EER (Error Equal Rate). The EER

will be utilised as a key factor to measure the accuracy of the proposed system. It means at which point the FAR and FRR will be equal.

The first experiment examined how the accuracy result of the system is affected by applying the different deep learning methods. The outcomes of this practical experiment will identify the optimal method. The two deep learning algorithms selected are: Long short-term memory (LSTM) and A convolutional neural network (CNN). These two algorithms are examined on the first dataset (Dropbox dataset) with 80/20 splitting for the training and testing data.

The second experiment examined each dataset separately using the deep learning algorithm that provides better result from the first experiment in order to check and record the accuracy of each one, which can be utilised to compare later with the result accuracy of the combining of the two datasets (Dropbox and personal computer dataset).

The third experiment focuses on examining the deep learning algorithm m that gives better result with the two datasets, then combining the features of these datasets together and measuring the vital impact on the performance of the proposed system.

The fourth experiment explored the effect of the volume training and testing datasets on the accuracy of the system. The best accuracy of the dataset was selected to check the impact of the volume of the dataset with the best performance of the deep learning algorithm of the first experiment that was implemented. 50/50, 60/40, and 80/20 were investigated as splitting volumes of the given dataset. This will lead to a better understanding about the nature of user behavioural profiles based on the volume of data that each user should need for training data to provide a good level of performance.

Experimental Results

1. First Experiment

As mentioned in the methodology, the first experiment selected two deep learning algorithms (LSTM and CNN). The Dropbox dataset was selected in this experiment with data volume for training and testing (80/20). The overall output performance of these algorithms is shown in Table 4. In general, the outcomes encourage the idea of identifying the authorised user or

unauthorised access to data stored in cloud computing systems, with EERs that are in parallel to the same results in other applications from the previous studies [20] [21].

Table 4: Performance of deep learning algorithms

Deep Learning Method	EER (%)
LSTM	13.1
CNN	11.6

The primary result of deep learning algorithm performances in Table 4 supports the idea that it can be verified whether legal or illegal access to cloud storage services possible. Additionally, the nature of deep learning algorithms has a good impact on the overall accuracy of the suggested system.

2. Second Experiment:

As mentioned in the methodology section, the second experiment examined the performance of each dataset separately (Dropbox and personal computer dataset) in order to record each accuracy for comparison purposes with the next experiment, which might show an improvement in the result. Table 5 below shown the performance of each dataset with the CNN algorithm as selected the best method from the previous experiment. The volume of data was selected for the training and testing stage is (80/20).

Table 5: Performance of two datasets with CNN

User	Dropbox EER(%)	Personal Computer EER(%)
1	0	2.1
2	0	1.3
3	41.2	15.3
4	4.4	5.3
5	32.1	7.3
6	15	7
7	0	1.2
8	1	7.2
9	0	10.2
10	21	0
11	0	8.9
12	0	6.2
13	0	10.5
14	8.3	0
15	16	23
16	1.3	12.3

17	7.6	0
18	40.3	20.6
19	40.4	23.4
20	4.5	0
Average	11.6	8.09

Table 5 shows that each dataset got an acceptable performance of EER, the Dropbox dataset achieved 11.6% of EER, whereas the personal computer dataset achieved 8.09% of EER. The performance of the personal computer dataset achieved a better result than the Dropbox dataset; this is because it might be the number of interactions in personal computers is more than the Dropbox dataset, which might give insight to the classifier algorithm to perform a better decision. These results of both experiments will be compared with the performance of the combination of the two datasets of the next experiment.

3. Third Experiment:

As demonstrated in Table 3, combining the features of the two datasets might lead to improve the performance of the proposed system. Applying this procedure, might increase the level of disgusting among cloud computing users. So, this can give a better insight to the classifier to perform a better performance. Table 6 shows the users' performance of merging the feature of Dropbox and personal computer datasets. The best deep learning algorithm (CNN) was selected as it performed a higher accuracy result than the algorithm LSTM. The splitting data 80/20 was selected for training/testing in this experiment.

Table 6: Performance of combining datasets

User	EER(%)	User	EER(%)
1	0	11	0
2	0	12	0
3	6.8	13	0
4	0	14	0.5
5	7	15	5.8
6	0	16	0.1
7	0	17	0
8	0	18	30
9	0	19	21
10	0	20	2.6
Average 3.6			

Based on the overall average performance, the experiment has shown that combining the features of the two datasets has a significant impact on performance, as illustrated in Table 6. Many users achieved very high accuracy which reached to 100% such as users 1, 2, 4, 6,7, 8, 9, 10, 11, 12, 13, and 17. This means the classifier identified the authored and impostor users correctly without any error. While other users, such as users 18 and 19 their performance improved, the classifier sometimes cannot discriminate between the legal and illegal usage of users of cloud storage services.

4. Fourth Experiment:

This experiment examined the effect of the volume of the dataset for training on the accuracy of the proposed system. As mentioned previously, CNN was selected for this experiment because it accomplished a better accuracy than the first experiment. The combined dataset was implemented to examine this factor. The volume of data for training and testing that was utilised in this experiment was set to 50/50, 60/40, and 80/20. Table 7 demonstrates the accuracy result (EER) of all participants across the selected volumes of the dataset.

Table 7: Accuracy based on volume of data

User	EER (%) based on volume of data		
	50/50	60/40	80/20
1	8.5	7,2	0
2	0	0	0
3	7.8	7.7	6.8
4	9.2	10.5	0
5	18.9	5.8	7
6	8	8.2	0
7	0	0	0
8	0	0	0
9	4	0	0
10	4	2.5	0
11	13.9	12	0
12	0	0	0
13	0	0	0
14	3.9	7.6	0.5
15	7.5	8.2	5.8
16	7	5.8	0.1
17	0	0	0
18	48.6	21.7	30

User	EER (%) based on volume of data		
	50/50	60/40	80/20
19	40.8	30.7	21
20	4.7	7.6	2.6
Average	**9.3**	**6.7**	**3.6**

As shown in Table 7, the performance of the CNN algorithm was achieved better when the volume of the dataset was increased in the training phase. Therefore, a maximum volume of samples that achieved better accuracy results than a minmum volume of data as an overall result. 3.6% of EER was the best performance volume of the dataset 80/20 for training and testing, respectively. This experiment agrees with the state of the art that a large volume of data in the training stage can have a better impact on the overall accuracy result. This is logical because more datasets for training mean the classifier will learn more about the behaviour patterns of the users, which can lead to better accuracy.

However, an individual user's perspective, most of the users' performances had a positive change when the volume of data training was increased, such as users 5, 9, 18, and 19. However, some users had a negative impact on the performance especially in volume of data 50/50 to 60/40, such as users 4, 14, and 20; while the users' performance with volume 80/80 had a positive impact across all users. This means some users have activities relatively stable, while other users' activities change over time. In a practical sense, the users who have a more stable behaviour profile would be very helpful to the classifier to make an accurate decision. Other users are needed to build a correct profile by renewing their template regularly in order to include all new users' activities in the users' behavioural profiles.

Discussion

It is clear that from all previous experiments, users of cloud storage services can be distinguished via their activity with a reasonable accuracy have been achieved. Moreover, the result of this study is in aligned with the highest performances that are achieved in the state of the art as shown in Table 8. From the classifier performance perspective, the CNN algorithm achieved a better accuracy than the LSTM algorithm. So, with the combining of the two datasets, the CNN performed at 3.6% EER.

Table 8: Performance of previous studies with the current study

Author(s)	Activity	Performance (%)	Method
[10]	Telephony	DR=80	Genetic Programing method
[11]	Telephony	FRR=3	self-Organizing Map and Probabilistic models
[12]	Telephony	DR=70	Neural Network
[13]	Mobility	DR=81	cumulative probability and Marko properties of trajectories
[14]	Mobility	DR=94	cumulative probability and Marko properties of trajectories
[15]	Telephony	DR=97	SVM
[16]	Telephony, Browsing, SMS, Mobility	DR=95	Probability
[17]	Telephony, Browsing, SMS	DR=98.5,EER=1.6	Bayesian network , RBF, KNN, Random Forest
[18]	Telephony, Application, SMS	EER=13.5, 2.2, 5.4	Neural network
[19]	Application Usage	EER=9.8	Rule base
[20]	Text, App, Web and GPS	EER=3	SVM
[21]	PC	EER= 7	Neural Network (FF-MLP)
[22]	Network event	DR=90, FAR=14, FRR=11	K-Means Clustering
[23]	File Access	FAR=1.1	SVM
[24]	Web Site	DR=91	support-based, lift-based profiling
[25]	Web Site	EER= 24	SVM
[26]	Cloud Infrastructure	-	-
[27]	Cloud Infrastructure	DR=92.7, 85.1, 99.8	Random forest, Neural network and Gradient boosting algorithms
[5]	Cloud Computing Storage (Dropbox)	EER= 5.8	SVM, RF-25 trees, FF MLP Neural Network-65, CART
This Study	Cloud Infrastructure + Cloud Storage (Dropbox)	EER= 3.6	CNN, LSTM Deep Learning

The second experiment was the primary investigation that examined each dataset separately in order to record the performance of each dataset. This experiment was conducted to compare the performance results of each experiment.

The third experiment evaluated the accuracy of combining two datasets and then comparing the performance results with the accuracy of a single dataset. This experiment showed that combining datasets achieved a better performance (3.6% EER) than each dataset separately. This result is logical because more data will allow the classifier to learn better about each user. As a result, a good performance was achieved by the classifier with a low error rate.

From the outcomes of the fourth experiment it is demonstrated that the data volume for training and testing has an effect upon the average of accuracy. As shown in the experiment, a big of volume training data (that is, 80/20 splitting) achieved better performance with EER an 3.6%. on average. However, from individual users' performance, a number of users achieved lower accuracy results when training data is increased (50/50 and 60/40), or less accuracy with 80/20. For instance, User 18 had the lowest accuracy (EER) compared to all other users' accuracy results. When the volume of data was increased from 50/50 to 60/40 for the training stage, the accuracy also increased, whereas when the data volume of training was increased to 80/20, the accuracy did not improve. When looking at the pattern of daily activities of user 18, it is found that some events/applications of the user did not seem to have consistent usage such as 'Edit' or 'Delete' which did not appear regularly. Some events appeared in the first or last two months, shown below in Figure 1.

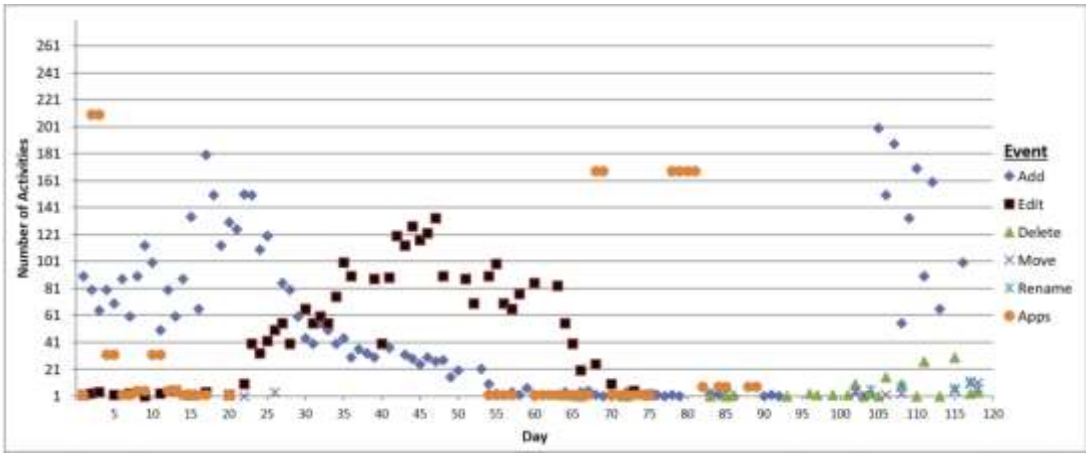

Figure 1: Recorded Activities for User 18

The change in the user's usage can have a negative impact on the accuracy of the deep learning algorithm because user's behaviour is so diverse. Therefore, change in users' behaviour means that users' templates need to be updated regularly to make sure the quality of samples is still good in order to achieve a better level of system performance. However, this is not an easy task because renewal users' templates continually might include impostors' data with the legitimated users' activities. This lead to an authorised user might be accepted by the proposed

system over the time as a legal user. Therefore, this important problem needs to be managed carefully in order to ensure unauthorised activities are not included in these templates.

Conclusion and Future work

The experimental results of applying two deep learning algorithms showed the capability to correctly distinguish among users of cloud storage services based on their activities while interacting with the service (Dropbox). Using users' behavioural profiling, an accurate user's template can be created which can help to discriminate between the authorised and unauthorised users' interactions. A better accuracy was achieved by the CNN algorithm especially with combined datasets. Further experiments have shown that the volume of data for training and testing has a significant impact on accuracy. From an individual's perspective, many users achieved a high performance which means the system identified users' interactions fully and correctly without any error. Therefore, the proposed system proved a highly promising solution for applying user behavioural profiling as a second factor that can support identifying the normal and abnormal usage of the users after initial login authentication. This can help and guide the system to detect misuse of cloud storage services in a continuous and transparent way. However, some other users performed with very low accuracy. In this case, the proposed technique would not be suitable as a supporting factor to validate the users.

Future work needs to focus on developing mechanisms for user template renewal in order to continuous updating users' behaviour frequently to ensure new user's behaviour is included. This can lead to improve the system performance and allow the system to make a correct decision.

References

1. Statistics, Amazing Cloud Adoption. "Cloud migration, computing, and more [Electronic resource]." Access mode: https://www. zippia. com/advice/cloudadoption-statistics
2. Dropbox. (2024). Dropbox Statistics (2024). Available: https://www.skillademia.com/statistics/dropbox-statistics

3. C. Alliance, The Treacherous Twelve-Cloud Computing Top Threats in 2016, ed, (2016)
4. D. Chen, H. Zhao, Data security and privacy protection issues in cloud computing, In: 2012 international conference on computer science and electronics engineering, 1, 647-651(2012), IEEE,DOI(https://doi.org/10.1109/ICCSEE.2012.193)
5. B. Al-Bayati, N. Clarke, P. Dowland, F. Li, Continuous identity verification in cloud storage services using behavioural profiling, In: 17th European Conference on Cyber Warfare and Security, 1-10(2018), Academic Conferences and Publishing International Limited.
6. Gupta, Udit, Survey on security issues in file management in cloud computing environment, arXiv preprint arXiv:1505.00729, (2015), DOI(https://doi.org/10.5120/21224-3948)
7. D. J. T. G. Yadron, Hacker collects 272 m email addresses and passwords, some from Gmail| Technology| The Guardian, (2016)
8. G. A. Pimenta Rodrigues, Understanding Data Breach from a Global Perspective: Incident Visualization and Data Protection Law Review, 9(2), 27(2024), DOI(https://doi.org/10.3390/data9020027)
9. N. Lee, Cyberattacks, Prevention, and Countermeasures, In: Counterterrorism and Cybersecurity: Total Information Awareness: Springer, 295-342(2024), DOI(https://doi.org/10.1007/978-3-031-63126-9_10)
10. J. Hall, M. Barbeau, E. Kranakis, Anomaly-based intrusion detection using mobility profiles of public transportation users, In: WiMob'2005), IEEE International Conference on Wireless And Mobile Computing, Networking And Communications, 2005, 2, 17-24(2005),DOI(https://doi.org/10.1109/WIMOB.2005.1512845)
11. C. S. Hilas, S. A. Kazarlis, I. T. Rekanos, P. A. Mastorocostas, A genetic programming approach to telecommunications fraud detection and classification, In: Proc. 2014 Int. Conf. Circuits, Syst. Signal Process. `Commun. Comput, 77-83(2014)

12. F. J. J. o. S. Ogwueleka, Technology, Fraud detection in mobile communications networks using user profiling and classification techniques, 29(3), (2009), DOI(https://doi.org/10.4314/just.v29i3.50052)
13. S. QAYYUN, S. MANSOOR, A. KHALID, Fraudulent Call Detection for Mobile Networks [C], In: Proceedings of 2010 International Conference on Information and Emerging Technologies (ICIET): June, 14-16(2010), DOI(https://doi.org/10.1109/ICIET.2010.5625718)
14. S. Yazji, R. P. Dick, P. Scheuermann, G. Trajcevski, Protecting private data on mobile systems based on spatio–temporal analysis, In: International Conference on Pervasive and Embedded Computing and Communication Systems, 2, 114-123(2011)
15. S. Yazji, P. Scheuermann, R. P. Dick, G. Trajcevski, R. J. P. Jin, U. Computing, Efficient location aware intrusion detection to protect mobile devices, 18, 143-162(2014),DOI(https://doi.org/10.1007/s00779-012-0628-9)
16. S. Subudhi, S. J. P. C. S. Panigrahi, Quarter-sphere support vector machine for fraud detection in mobile telecommunication networks, 48, 353-359(2015), DOI(https://doi.org/10.1016/j.procs.2015.04.193)
17. S. Yazji, X. Chen, R. P. Dick, P. Scheuermann, Implicit user re-authentication for mobile devices, In: International Conference on Ubiquitous Intelligence and Computing, 325-339(2009),DOI(https://doi.org/10.1007/978-3-642-02830-4_25)
18. E. Shi, Y. Niu, M. Jakobsson, R. Chow, Implicit authentication through learning user behavior, In: Information Security: 13th International Conference, ISC 2010, Boca Raton, FL, USA, October 25-28, 2010, Revised Selected Papers 13, 99-113(2011), DOI(https://doi.org/10.1007/978-3-642-18178-8_9)
19. D. Damopoulos, Evaluation of anomaly-based IDS for mobile devices using machine learning classifiers, 5(1), 3-14(2012), DOI(https://doi.org/10.1002/sec.341)
20. F. Li, N. Clarke, M. Papadaki, P. J. I. j. o. i. s. Dowland, Active authentication for mobile devices utilising behaviour profiling, vol. 13, 229-244(2014), DOI(https://doi.org/10.1007/s10207-013-0209-6)

21. L. Fridman, S. Weber, R. Greenstadt, M. J. I. S. J. Kam, Active authentication on mobile devices via stylometry, application usage, web browsing, and GPS location, 11(2), 513-521(2016),DOI(https://doi.org/10.1109/JSYST.2015.2472579)
22. Ö. D. Incel, DAKOTA: Sensor and touch screen-based continuous authentication on a mobile banking application, 9, 38943-38960(2021), DOI(https://doi.org/10.1109/ACCESS.2021.3063424)
23. V. Gattulli, D. Impedovo, G. Pirlo, F. J. S. R. Volpe, Touch events and human activities for continuous authentication via smartphone, 13(1), 10515(2023), DOI(https://doi.org/10.1038/s41598-023-36780-3)
24. M. B. Salem, S. J. Stolfo, Modeling user search behavior for masquerade detection, In: International workshop on recent advances in intrusion detection, 181-200(2011), DOI(https://doi.org/10.1007/978-3-642-23644-0_10)
25. M. Abramson, D. Aha, User authentication from web browsing behavior, In: The Twenty-Sixth International FLAIRS Conference, (2013)
26. D. S. David, Cloud Security Service for Identifying Unauthorized User Behaviour, 70(2), (2022),DOI(10.32604/cmc.2022.020213)
27. L. K. Vashishtha, A. P. Singh, K. J. W. P. C. Chatterjee, HIDM: A hybrid intrusion detection model for cloud based systems, 128(4), 2637-2666(2023)
28. A. Gupta, R. Simon, Enhancing Security in Cloud Computing With Anomaly Detection Using Random Forest, In: 2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), 1-6(2024), IEEE, DOI(https://doi.org/10.1109/ICRITO61523.2024.10522227)
29. J. Sola, J. J. I. T. o. n. s. Sevilla, Importance of input data normalization for the application of neural networks to complex industrial problems, 44(3), 1464-1468(1997)