ISSN 2320 - 2602

Volume 9 No.6, June 2020

International Journal of Advances in Computer Science and Technology

Available Online at http://www.warse.org/IJACST/static/pdf/file/ijacst04962020.pdf

https://doi.org/10.30534/ijacst/2020/04962020

User Authentication for Smart Home using IoT Devices

Abraham John[1], Akash R[2], Ajay Manuel[3], Dibin Joseph[4], Syamamol T[5]

[1] Mangalam College of Engineering, India, abrahamjohn891gmail.com

[2] Mangalam College of Engineering, India, rakash75@gmail.com

[3] Mangalam College of Engineering, India, ajaymanuel198@gmail.com

[4] Mangalam College of Engineering, India, dibinjoseph71@gmail.com

[5] Mangalam College of Engineering, India, syamamol.t@mangalam.in

ABSTRACT

A smart home is one that provides its home owners comfort, security, energy, efficiency and convenience at all times, regardless of whether anyone is home. End-user devices, such as mobile phones and tablets, have become essential tools for accessing smart homes. Although mobile phones are equipped with different means of authentication such as fingerprint readers, face lock, etc., these methods are only employed at time of access. User Authentication for Smart Home Networks Based on the Usage of IoT Device. It presents a continuous user authentication model based on behavioral features extracted from user interactions with IoT devices. If and when an unauthorized access occurs, permission to access the system would be denied. Consecutively an alert message would be sent to the emergency phone number provided during time of user registration.

Key words: Machine Learning, IoT, Smart home, classification algorithm.

1.INTRODUCTION

The technological advancement has created remarkable change, which is also reflected in the smart home as one of the Internet of Things (IoT) applications. However, this growth has brought an increased number of security threats. Many end user devices from which access to smart home devices are achieved, including mobile phones and tablets, are currently equipped with authentication techniques, such as fingerprint readers, face locks, etc. Since these methods are only employed at the time of access, many resulting problems should be taken into consideration. Firstly, after initial authentication, many users avoid setting short access sessions; as a result, after the entry point, devices are vulnerable to use by intruders leading to unauthorized access. Secondly, traditional authentication credentials could be observed by external users, as the devices are susceptible to loss or theft. Hence, it is a key factor and it is important to consider the continuous authentication techniques that continuously examine the access situation and the legitimacy of the communicating party. This can be achieved by utilizing behavior-based authentication methods. Behavior-based authentication methods are built on an assumption that users have unique behaviors, like browsing particular pages and using specific apps. The advantage of considering behavior-based authentication is that information such as apps, device resources and network information generated while interacting with mobile devices, can be retrieved and employed in the background for the purpose of continuous authentication [1]. In addition , the enhanced computation capabilities of smart home hubs, for example , Samsung Smart Things and Wink home , can run continuous authentication procedures while accessing home devices. Here presents a method that authenticates users to smart home devices at the starting point of access , as well as during the access session, without requiring further action. The proposed approach utilizes unique user app usage events on the mobile phone. In consequence, the contribution of this study is a ML authentication model utilizing access behavior to apps on mobile devices to authenticate users while utilizing shared apps at the same daily intervals.

2.RELATED WORKS

Much information can be retrieved while interacting with mobile devices and applied for continuous user authentication such as device resources and apps, as well as network information generated from apps. The study in “Into the Wild: Studying Real User Activity Patterns to Guide Power Optimizations for Mobile Architectures”, finds that there is a high correlation between the power consumption of the used device and the usage patterns of users [1]. Furthermore, as demonstrated in “Continuous Authentication on Mobile Devices Using Power Consumption, Touch Gestures and Physical Movement of Users”, user behavior relies on the power consumption of the accessed apps [4]. In addition, utilizing built-in battery voltage sensors, the authors in “Design and Implementation of a Contextual-Based Continuous Authentication Framework for Smart Homes”, present a power model construction approach that monitors power consumption per app on user devices [8]. However, the works above mentioned do not utilize app access patterns for users’ authentication and it is difficult to model the power consumption only for specific apps as there are apps running in the background. Considering network-based information, the authors in “Fraud Detection Concepts: Final Report”, report on building user behavior profiles, including phone calls and network migration patterns over service provider networks [12]. The research in “Behaviour Profiling for Transparent Authentication for Mobile Devices”, shows that users can be differentiated and anomalies can be detected based on users’ interactions with their mobile apps utilizing text messages and calling behavior, hence enabling continuous user authentication [13]. According to the assumption that users of mobile phones tend to utilize apps in different locations at different times, host-based behavior profiling is described by the same authors. The study in “Implicit Authentication Through Learning User Behavior”, proposes an anomaly-based detection approach based on monitoring users’ actions [6].

3.PROPOSED SYSTEM

In this system, the input as mobile phone which has the access information. The data is collected from real world users by receiving their access logs and performing features selection. The collected data has to be preprocessed by normalizing features, removing noise etc. Behaviour based authentication methods are built up on an assumption that users have unique behaviour patterns. These behaviour patterns can be identified by applying classification algorithms .The authentication can be done based on the accuracy check in these behaviour patterns and classification algorithm is applied by using raspberry pi. The model building includes data pre-processing, identification of the patterns, classification and authentication. These areas can be tested for a better authentication model. If the access to the smart home device is granted or if it’s denied, it will pop up in the build-in mobile application. If an unauthorized access takes place, the system will be blocked and an alert message is sent to the mobile number given at the time of registration. Then the access to the smart home network will be denied.

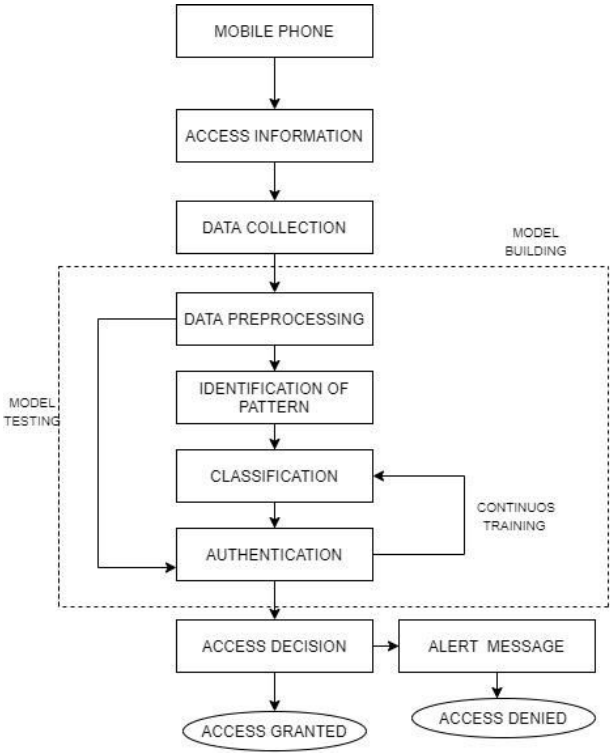


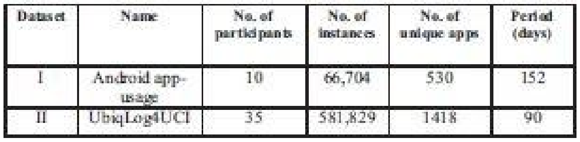
Figure 1: The proposed authentication model

This model presents a user authentication model that first learns users’ app access events on their end devices and produces a pattern for continuous user authentication using classification algorithms. In the implemented smart home framework, users interact with smart home devices through a smart home hub and data are encrypted during transmission.

Dataset

For evaluating the proposed approach, two public real-world datasets, collected in an uncontrolled manner from real users, are utilized. The first dataset is Android app-usage , was collected from users over a period of five months, and publicly available for only ten users. The second dataset is the UbiqLog4UCI , which was collected from 35 users over a period of more than three months. Table 1 presents both datasets regarding the number of users, instances, and accessed apps, as well as app access duration.

Table 1: The Utilized Dataset.



Dataset	Name	No. of participants	No. of instances	No. of unique apps	Period (days)
I	Android app-usage	10	66,704	530	152
II	UbiqLog4UCI	35	581,829	1418	90

Dataset I description

This dataset comprises management activity and network trace data. The management activity data comprises history information of apps download, update, and uninstalls. The network trace data comprises the access network, such as Wi-Fi or cellular networks, the apps' access duration of the running apps as well as the generated traffic while accessing these apps. Although data of only 10 users, which is publicly available, is utilized in this work, the number of home users is commonly smaller.

Dataset II description

This dataset was collected in the form of records, each of which represents a sample of the observations in the dataset. All participants were university students who own a smartphone and who participated in the data collection without reward. This dataset contains 35 users; however, the data of 30 participants were included, while the data of 5 of them were removed during this study due to errors and missing app usage information. The meaning and the fields of both utilized datasets are shown in Table 2.

Table 2: Utilized features from both Datasets.

	Feature	**Meaning**	**Type**
1	Phone number	The user's identifier	Text
2	Timestamp	The access time and date	Date
3	Apps name	The app's identifier	Text
4	* F-(Cellular & Wi-Fi) access duration	Foreground cellular and Wi-Fi connection duration. This feature is separately provided with the Android app-usage dataset and provided in total with the UbiqLog dataset	Time (ms)
5	* F-(Cellular & Wi-Fi) Traffic	Foreground cellular and Wi-Fi traffic. This feature is utilized only with the Android app-usage dataset as it is not collected	Byte

Datasets analysis

The interaction time in this study is considered as the total access time to the app in the foreground, meaning that the user is continuously interacting with the app during the session. As an overview, Figures show the average daily interaction time of four randomly selected users from each dataset, over a period of 90 consecutive days. The average daily access time ranged from 10.5 minutes to 317.9 minutes per day in Dataset I, and from 75.4 minutes to 392 minutes in Dataset II.

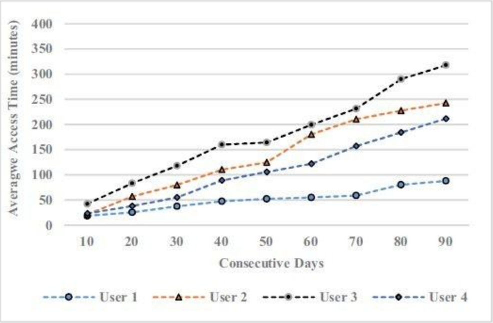


Figure 2: Average access time to apps per day in Dataset I.

Moreover, 30% of the interaction time of the selected apps had an average range between 30 and 180 minutes. Even though there are some variations in the daily access time over the selected period , as Fig. 2 and Fig . 3 show, the difference among users is obvious . Furthermore , the daily average access time per user is unique as compared with the rest of users. Consequently, the total access time can be applied for user authentication .Before training the model , it is important to consider the imbalance in the observations in each of the datasets . For example , in Dataset I, class observations vary from 5.5 % to 16.2%, while in Dataset II, class observations vary from 0.

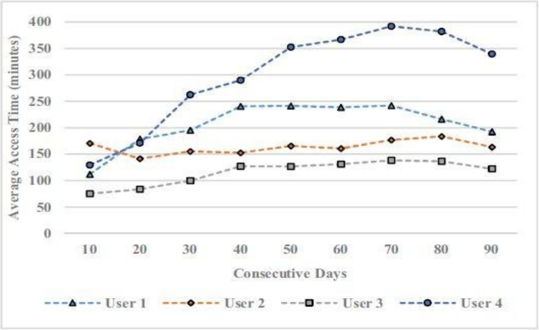


Figure 3: Average access time to apps per day in Dataset II.

Since the focus here is on classifying multi-users rather than a single user, different classification strategies are applied. The accuracy measures for each algorithm are evaluated 5 times, and here six classifiers are applied, namely, KNN, MLP, NC, RF, GBC and AdaB, which provide the best accuracy compared with other algorithms. The classification algorithms are trained over a period of time until reaching no considerable variation in the access pattern of users.

APP SELECTION

Based on the assumption that each user has unique accessing patterns to apps in different locations and at different times, it is rare that two users will have identical access patterns. However, this situation could occur especially in localized environments such as at home or work; for example, two relatives residing in the same house and utilizing the same apps, such as Facebook or WhatsApp, on their smartphones. Accordingly, the model is evaluated in identifying users utilizing shared apps at the same daily intervals. Before evaluating the presented model, it is tested if there is a significant difference between two data samples, all apps and most used apps. For the most used apps, it is considered that apps are accessed twice or more by all users.

4.DISCUSSION

Behavior-based authentication methods are built on the assumption that users have unique behaviors, like browsing particular pages and using specific apps, while interacting with their mobile devices. In this study, it is presented a user authentication model based on apps access events with high TPR and TNR. The proposed method has been evaluated utilizing two datasets, and for this work app events are used over 50 consecutive days. The results show that even when using only shared apps, this model is able to differentiate between users. According to the selected sample group, users spent an average of up to 5 hours a day accessing apps on their smartphones for Dataset I and up to 6.5 hours for Dataset II. However, Dataset I provides higher accuracy levels when accompanying both app access traffic generated along with the time spent accessing these apps. As it is concluded from the utilized datasets in this study, app access patterns are continuously provided. Thus it can be applied for continuous user authentication.

5.RESULT

The results of testing the performance on both datasets with regard to the true negative rate (TNR) and the true positive rate (TPR). The average TPR and TNR based on the one-vs-all classification approach of the proposed model are presented in Fig.
4. From this figure, the RF algorithm achieves the best TPR followed by the KNN for both the datasets and the RF algorithm performs the best TNR for both datasets, followed by the GBC for Dataset I and KNN for Dataset II.

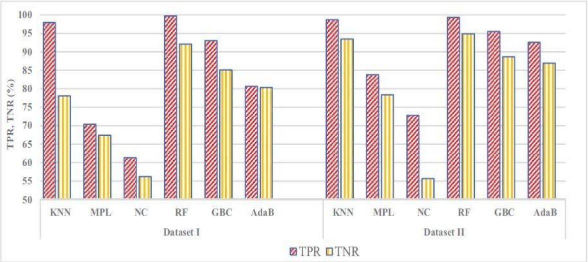


Figure 4: Performance evaluation of the proposed model based on TPR and TNR.

For evaluating the proposed model, the datasets are divided into 30% and 70% for validating and training the model. Training and validation is performed based on two classification methods, one-vs-all and all-vs-all.

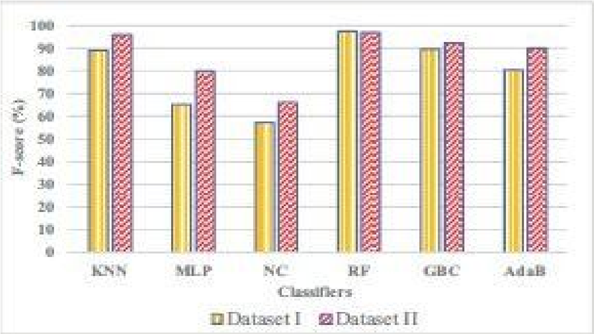


Figure 5: One-vs-all performance of the proposed model.

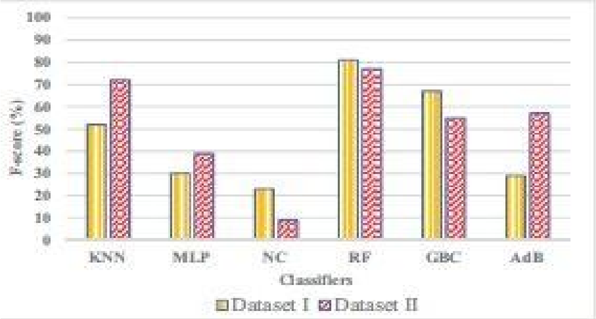


Fig. 6: All-vs-all performance of the proposed model.

6.CONCLUSION

Here we present a user authentication model utilizing app access events based on most used apps among users. The capacity of the presented model, to authenticate users with both high TP and TN rates, is demonstrated and from the results it is concluded that for each user even for the most used apps, the apps access patterns are different. Hence, this model is robust against such kinds of situations. The increased usage of mobile apps on end user devices, such as mobile phones and tablets, will provide the possibility to authenticate users with an appropriate level of accuracy.

The results show that this model is able to differentiate between users even when using only shared apps as compared with the work in. From the results, there are some changes in users' app access patterns, however, the effect on accuracy is minor since utilized apps are the most frequently accessed apps. Moreover, this change is small when including the continuously employed apps among users for a long period of time.

REFERENCES

1. A. Shye, B. Scholbrock, and G. Memik. **Into the Wild: Studying Real User Activity Patterns to Guide Power Optimizations for Mobile Architectures,** *Proceedings of the 42nd Annual IEEE/ACM International Symposium on Microarchitecture.*, pp. 168–178, 2009.
2. L. Abdi and S. Hashemi. **To Combat Multi-Class Imbalanced Problems by Means of Over-Sampling Techniques**, *IEEE Transactions on Knowledge and Data Engineering.,* vol. 28, no. 1, pp. 238–251, 2016.
3. M. F. Amasyali and O. K. Ersoy. **Classifier Ensembles with the Extended Space Forest,** *IEEE Transactions on Knowledge and Data Engineering.,* vol. 26, no. 3, pp. 549–562, 2014.
4. R. Murmuria, A.Stavrou, D.Barbará, and D. Fleck. **Continuous Authentication on Mobile Devices Using Power Consumption, Touch Gestures and Physical Movement of Users,** in International Workshop on Recent Advances in Intrusion Detection, pp. 405–424, 2015.
5. D. Damopoulos, S. A. Menesidou, G. Kambourakis, M.Papadaki, N.Clarke, and S. Gritzalis. **Evaluation of Anomaly-Based IDS for Mobile Devices Using Machine Learning Classifiers,** *Security and Communication Networks,* vol. 5, no. 1, pp. 3–14, 2012.
6. E. Shi, Y.Niu, M.Jakobsson, and R. Chow, **Implicit Authentication Through Learning User Behavior,** *Information Security., Springer,* vol. 6531 LNCS, pp. 99–113, 2011.
7. A. Kalamandeen, E.De Lara, and A. Lamarca. **Ensemble: Cooperative Proximity Based Authentication**, pp. 331–343, 2010.
8. Y. Ashibani, D. Kauling,and Q. H. Mahmoud. **Design and Implementation of a Contextual-Based Continuous Authentication Framework for Smart Homes,** *Applied System Innovation*, vol. 2, no. 1, pp. 1–20, 2019.
9. L. Zhang et al. **Accurate Online Power Estimation and Automatic Battery Behavior Based Power Model Generation for Smartphones**, *Proceedings of the eighth IEEE/ACM/IFIP international conference on Hardware/software codesign and system synthesis.*, pp. 105–114, 2010.
10. A. Buriro, B. Crispo, F. DelFrari, and K. Wrona. **Hold and Sign: A Novel Behavioral Biometrics for Smartphone User Authentication,** *Proceedings - 2016 IEEE Symposium on Security and Privacy Workshops,* SPW 2016, pp. 276–285, 2016.
11. D. Buschek, A. De Luca, and F. Alt, **Improving Accuracy, Applicability and Usability of Keystroke Biometrics on Mobile Touchscreen Devices,** *in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* - CHI '15, pp. 1393–1402, 2015.
12. P. Gosset, **Fraud Detection Concepts: Final Report,** CiteSeer, Doc Ref. AC095/VOD/W22/DS/P/18/1, pp. 1–27, 1998.
13. F. Li, N. Clarke, M. Papadaki, and P. Dowland, **Behaviour Profiling for Transparent Authentication for Mobile Devices,** European Conference on *Cyber Warfare and Security*. Academic Conferences International Limited, pp. 307–315, 2011.
14. U. Mahbub , J. Komulainen , D. Ferreira , and R. Chellappa .**Continuous Authentication of Smart phones Based on Application Usage** , arXiv preprint arXiv:1808.03319, 2018.
15. A. Al-Shookri , G. L. Khor , Y. M. Chan , S. C. Loke, and M. Al-Maskari. **Anomaly Detection: A Survey ,**ACM Computing Surveys (CSUR), vol. 41, no. 3, pp. 1–58, 2009.
16. R. Rawassizadeh , E. Momeni , C. Dobbins , and P . Mirza -babaei ,**Lesson Learned from Collecting Quantified Self Information via Mobile and Wearable Devices ,** Journal of Sensor and Actuator Networks, vol. 4, no. 4, pp. 315–335, 2015.
17. **Android App-Usage Data.** [Online]. Available: http://sei.pku.edu.cn/~liuxzh/appdata/. [Accessed : 15-Jul-2018].