This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2719706, IEEE Access

Social Behaviometrics for Personalized Devices in the Internet of Things Era

Fazel Anjomshoa, *Student Member, IEEE,* Moayad Aloqaily, *Member, IEEE,* Burak Kantarci, *Senior Member, IEEE,* Melike Erol-Kantarci, *Senior Member, IEEE,* and Stephanie Schuckers, *Senior Member, IEEE,*

***Abstract*—As the integration of smart mobile devices to the Internet of Things (IoT) applications is becoming widespread, mobile device usage, interactions with other devices, and mobility patterns of users carry significant amount of information about the daily routines of the users who are in possession of these devices. This rich set of data, if observed over a time period, can be used to effectively verify a user. In previous works, verification of users on personalized electronic devices via biometric properties such as fingerprint, iris has been successfully employed to increase security of access. However, with the integration of social networks with the IoT infrastructure and their popularity on smart handheld devices, identification based on behavior over social networks is emerging as a novel concept. In this paper, we propose an intelligent add-on for the smart devices to enable continuous verification of users. In the experiments, we use data from built-in sensors and usage statistics of five different social networking applications on mobile devices. The collected feature set is aggregated over time and analyzed using machine learning techniques. We show that when smart devices are equipped with continuous verification intelligence, it is possible to verify users with less than 10% false rejection probabilities, and the users can keep using the devices with no interruption for biometric authentication 90% of the time. In the case of anomalous behavioral patterns, the proposed system can verify genuine users with up to 97% success ratio using an aggregated behavior pattern on five different social network applications.**

***Index Terms*—Internet of Things, continuous verification, mobile crowdsensing, smart cities, intelligent systems**

I. Introduction

Smart mobile devices like smartphones and tablets are equipped with various built-in sensors like GPS, camera, accelerometer, gyroscope and microphone among others. As the popularity and widespread use of these devices continue to increase, they appear to be strong candidates for being integrated to Internet of Things(IoT)-driven sensing applications [1]. In [2], [3], the main components of IoT are highlighted as follows: 1- hardware that consists of sensors and 2- middleware to provide communication between different components, 3) processing of data, 4) storage of data. Computing resources such as processors, memory and data storage have become smaller that can be embedded on wearable and hand-held devices [4].

F. Anjomshoa and S. Schuckers are with the Department of Electrical and Computer Engineering, Clarkson University, Potsdam, NY, 13699 USA (e-mail:{anjomsm, sschucke}@clarkson.edu).
M. Aloqaily is with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada (e-mail:{moayadaloqaily}@cunet.carleton.ca).
B. Kantarci and M. Erol-Kantarci are with the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON, Canada. The two authors also with the ECE Department of Clarkson University on courtesy appointments.
E-mail: {burak.kantarci, melike.erolkantarci}@uOttawa.ca

With the rapid growth of smartphone and personalized mobile device usage, and along with the recent advances in Internet of Things (IoT) where tremendous number of devices are interconnected, continuous authentication on personalized devices has become possible. Smartphones with various types of sensors have the potential for continuous monitoring of phenomena like road condition for smart transportation, public safety and emergency preparedness [5], [6]. With the widespread adoption of IoT devices, their use as a base for user verification is expected to grow [7]. On the other hand, effective incentives are needed in order for the users to provide the resources in their smart mobile devices as a service. The incentives to improve participation and integration of the personalized devices to the IoT environment can be either monetary or non-monetary. However, recruitment of the users is expected to be performed implicitly [1].

The advent of mobile computing and communications made web-based social networking services available through applications on portable smart devices such as phones, tablets and watches. With millions of portable devices in circulation and an immense attachment to everyday life, the popularity of social network services (SNS) have been continuously increasing [8]. Today approximately seven out of ten people in the U.S. use social networking services [9]. According to Ericsson's report, mobile applications for social networking produce high volumes of data that can be augmented with analytics for the betterment of various services [10]. In [11], the data types are classified under 6 different categories: i) Service data, ii) Disclosed data, iii) Entrusted data, iv) Incidental data, v) Behavioral data, vi) Derived data. Most users have regular behavioral patterns that are learnable, which can ultimately be used for continuous recognition of behavioral signatures, in [12]. On the basis of this presumption, we studied the behavioral patterns on smart mobile devices by focusing on mobile social network platforms to investigate identifying users in continuous fashion and verify the smartphones' owners. To this end, we propose a mobile behaviometric framework that assesses users' social activity, and introduce sociability metrics to generate signatures of users' activities. Traditional biometrics-based user identification relies on the uniquely personalized features such as fingerprint [13], iris [14], or face [15], [16], [17] and performs pattern recognition on these features to allow access to a user or a group. This type of identification is usually one-time and requires repeated interaction for validation of identities. On the other hand, personal devices are often in possession of a single user where continuous authentication of that user becomes more practical.

Behaviometrics refers to the behavior of a user which corresponds to mobile users' activities on smart mobile devices. Behavioral characteristics of mobile users can allow continuous authentication of a user on a personal device. Today's smart devices are equipped with various mobile applications. Among these, social networking apps have become an important part of daily life. Hence, these apps push massive amount of data to and from users to the servers of social network providers. The rich data set provided by the social networks can be mined to identify various types of relations. The online student-supervisor collaboration information derived in [18] or the friend ranking application developed in [19] are just to name a few.

User identification and authentication is an indispensable and basic requirement of preventing privacy leakage towards secure and trustworthy systems. Currently, the most common identification mechanisms in smartphones are passwords or pin codes which are not secure enough and require the user to remember the access codes for each device. In some smart phones, fingerprint and face detection are also integrated to ease the burden of remembering passwords and to increase the level of security. Zhang et al [20] categorized two types of biometric identification: physiological identification and behavioral identification. The former includes facial, voice and fingerprint recognition which are mostly device-dependent mechanisms and require costly processing units. On the other hand, continuous identification which is based on behavioral traits is non-intrusive and is based on human habitual patterns like typing [21]–[23], walking [24], [25], social interactions and communication. Sultana et al [26] defined social behavior biometrics as identification of a user in different social settings through interaction and communication patterns. The social setting can involve either online or offline environments where the former leads to the cyber world, and the latter denotes the physical world. In online settings interactions emerge from blogs, social networks and access to the Internet. Behaviometrics [27] is a recently emerging concept for identification and it promises to provide a cost effective alternative without compromising security.

In this paper, we propose an intelligent system to ensure device level security of mobile smart devices particularly to avoid identity spoofing when they are recruited for IoT-driven sensing. The proposed scheme is based on online behaviometrics of mobile users collected via smartphones, and extracts features from smartphone sensors and users' social network interactions. Real data traces were collected over several months and are used for the evaluation of the proposed approach. The feature set used in the paper includes location of users, their data usage, number of sessions in different time granularities and session durations for five different social networking platforms through the mobile device. Preliminary version of this work was presented in [28] focusing on disruption of users in continuous verification on mobile devices in the presence of spoofing attempts. This paper significantly extends the work in [28], by evaluating the performance of the proposed framework under normal and anomalous conditions with different classifiers. Furthermore, an extensive empirical analysis of the contextual weighing factors is also presented in detail. The proposed framework monitors the user data, trains the classifier and identifies users with high accuracy. The results show that identifying users with less than 10% false rejection for original user traces is possible.Thus, under various test scenarios, the proposed behaviometric approach can provide continuous authentication 90% of the time without the need to undergo additional biometric identification. In the presence of anomalous behavioral patterns, the proposed system can identify genuine users with up to 97% success ratio using an aggregated behavior pattern on five different social network applications. It is worth mentioning that as the participants shared almost the same profile, i.e., they were all graduate students in the same college, in some situations, the results showed similar signatures between users and this led to a slight increase in False Acceptance Rate (FAR) at the end. The paper is organized as follows. Section II presents the related work on behavioral identification. Section III provides the detail of the proposed system for identifying users based on online behaviometrics. Section IV provides performance evaluation and Section VI concludes the work and gives future directions.

II. Related Work

The idea of merging IoT and social network phenomena under the concept of Social Internet of Things (SIoT) has been emerging [29]. This convergence has many privileges including network navigability, service scalability and increased in level of trustworthiness by connecting the objects that interact on frequent basis [29]–[31].

To the best of our knowledge, Holmquist et al. [32] had initially proposed the idea of establishing relationship between smart objects which now translates into socialization of smart objects. In [33] the idea of establishing social networks by using IoT concepts was conceptually reviewed. In [34] the behavior of mobile nodes in IoT system by using social networks was studied. In [12], the behavioral patterns on various social network platforms was studied to investigate identification of users in continuous fashion and verification of smartphones' owners who contribute to participatory sensing campaigns in IoT contexts. To this end, in [12], the authors proposed a mobile behaviometric framework that assesses users' social activity, and introduced sociability metrics to generate signatures of users' activities. To the best of our knowledge this is the first research article which proposes continuous identification of users on mobile devices within the social IoT paradigm. The traditional identification schemes on mobile phones use pin codes, passwords, fingerprints or iris recognition. Pin codes and passwords have well-known vulnerabilities as mentioned previously [35]. Alternatively widely used biometric identification schemes (fingerprint, iris, face. etc.) are more secure and hard to compromise. However, they require extra hardware on devices as mentioned previously by several researchers [36]–[39].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2719706, IEEE Access

Sultana et al [14] categorized biometric-based authentication schemes into two groups: 1) Physiological biometrics such as fingerprint, facial recognition, iris and so on, and 2) behavioral biometrics which are based on human habitual signature including walking [13], handwriting, keystroke dynamics [40] and social networking. Continuous identification is based on behavioral patterns of users which advances existing identification mechanisms to a more secure, easier and non-intrusive fashion. Implicit authentication methods, which are based on observing user behavior through multiple sources such as SMS, phone calls, browser history, location, gestural patterns on touch screens and other kind of behavioral information; have become the seat of attention [41]–[47].

Having a large variety of different applications on smart phones or tablets has resulted in users' interacting with their smart devices frequently by revealing their personalized patterns. This fact has stimulated the researchers to mine users' interactions with smart devices as a source of user verification and identification. Yampolskiy et al [15] categorize behavioral biometrics into five different classes as follows: 1) Authorship based biometrics, 2) human computer interaction (HCI)-based biometrics, 3) indirect HCI-based biometrics, 4) motor-skills biometrics and 5) purely behavioral biometrics. In particular, the popularity of social networks yields users to generate large amount of data generated by mobile IoT devices. There are various research efforts in the area of mining social network induced information. Chen et al [48] address the social network traits like scam or finding the stem of rumors [49]. Sultana et al [50], [51] discuss the possibility of using behavioral patterns on social platforms for user identification. Lathia et al [52] proposed a mobile sensing framework for behavioral change interventions, named UBhave. In collaboration with Universities of Cambridge, Birmingham, Southampton, Oxford, and University College London, UBhave, a large digital behavioral change intervention (DBCI) framework, aims to be correlated with Online Social Networking sites (OSNs) in order to have better assessment of participants' social activity to recruit users. Mehrotra et al [53] built an automated context stream middleware based on OSNs to analyze and process users' behavior and interests. Yet, verification with real traces and verification success have not been evaluated comprehensively. In this paper, real traces that were collected over several months are used and machine learning (ML) techniques are applied to verify mobile users.

Behaviometrics is also an important part of smart environments such as smart homes as user signals and interaction with the homes can be used to reconfigure smart home settings [54]. Application of behaviometrics is not only limited to smart spaces but also used as an effective tool for continuous authentication. For instance, usage behavior patterns on hand-held devices (e.g. gestures on touchscreens) have been considered as continuous authentication solutions which is proposed by Buduru et al [55]. Although these works are relevant, they do not focus on verifying users, they rather search for usage patterns of appliances, lights or consumer devices. Another application in behaviometrics is health-care. CABA [56] is a continuous authentication health monitoring system which uses wearable medical sensors (WMSs). CABA is based on biomedical signal streams named BioAura that are continuously captured by WMSs. Schobel et al [57] addressed the flexibility issues in a mobile healthcare framework while using mobile healthcare applications for collecting patient data.

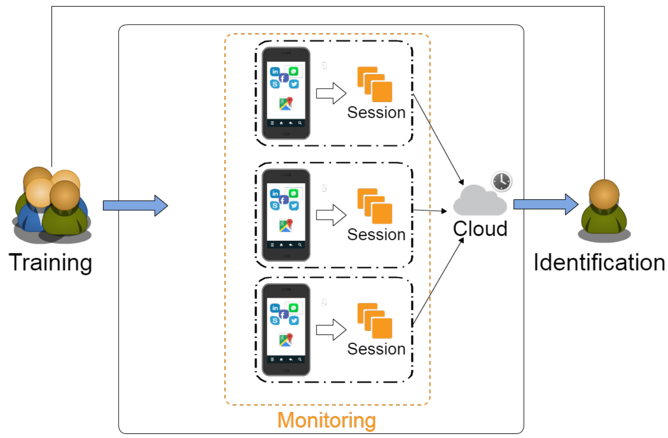


Figure 1. System overview

Having said that behavioral biometrics can be applied in smart environments, smart cities can be considered as another application area [58], [59]. Ziegler [60] provided a comprehensive research on the applicability of adapting behavioral biometrics in smart environments. The author described four possible smart environment applications including smart homes, smart media devices, smart traffic systems and smart health in which implicit identification mechanisms can be applied.

III. System Design

We have developed a front-end application that runs in the background of an Android phone, and monitors users' interactions through a smart device. The application collects data from five popular social network services which are; Facebook, Twitter, LinkedIn, Skype and WhatsApp. The collected data is stored in the form of sessions, and each session presents the corresponding user's interaction through the device. Basically each session data includes session ID, application name, the time that the session started, the time that the session ended, the duration of that session, the amount of data used in the session and the initial location where the session started. The amount of data used is the amount of cellular or Wi-Fi data consumed by the social network service application.

A. System Components

To verify the behaviometric signature of users, the following components are required:

1) Data Collection: Mobile user data collected from the device is uploaded to a private cloud-based server. The server stores the raw data from all users in a database. The database is queried for training and verification purposes.

2) User Characterization Model: User characterization is done by extracting a combination of features from both users' interaction over online social network services as well as the built-in sensors of the smartphones. The details of the model are provided in the following sections.

Table I
SYMBOLS LIST AND DESCRIPTION

SYMBOL	DESCRIPTION
$\mathcal{A}$	Social Activity Rate
$\mathcal{SF}$	Sociability Factor
$\mathcal{D}$	Data usage
τ	The number of sessions per day
T_k	$k-th$ activity rate
t	Duration of the activity
u	User u
$\mathcal{U}$	Set of users | $u \in \mathcal{U}$
p	Data point
$\mathcal{P}$	Set of Data points
ins	Instantaneous rate
sh	Short term activity
$overall$	Overall activity
$normal$	Normalized activity
$\mathcal{A}_{ins_i}^{\mathcal{U}_{app_x}^u}$	instantaneous Social activity of user u using application x in a session i
$\mathcal{A}_{sh}^{\mathcal{U}_{app_x}^u}$	Short-term Social activity of user u using application x
$\mathcal{A}_{overall}^{\mathcal{U}_{app_x}^u}$	Overall Social Activity
$\mathcal{A}_{normal}^{\mathcal{U}^u}$	Normalized Social Activity
α	Contextual parameter weight for running average calculation activity rates
β	Contextual parameter weight for running average calculation activity factors
μ	Mean value
σ	Standard deviation

3) Training Strategy: Training strategy builds a profile for each user based on the collected data. Training is performed continuously on a sliding window of data over time. This allows capturing naturally altering patterns of user behavior.

4) Verification Strategy: Machine learning is the core of user verification thus, the system is trained with feature sets collected by the front-end application, and user verification is performed based on each interaction through the device. The system components are presented in Figure 1.

B. System Architecture

Figure 2 presents the system architecture that includes main modules and methods, namely monitoring, data collection, normalization, training and verification modules. As mentioned earlier, personalized smart devices are envisioned to be integrated into the IoT ecosystem in order for the IoT applications to recruit those devices in various sensing campaigns by accessing their built-in sensors. In IoT sensing applications, to acquire the sensed data from personalized smart devices in a trustworthy manner, device level security must be ensured. Therefore continuous user verification should be positioned at the core of the personalized device-IoT integration. In the rest of this section, more details on each module are provided.

1) Monitoring Module: This module is an Android application that runs as a background process over the operating system. The application collects and updates user location information every 5 minutes. It monitors access to Facebook, Twitter, Linkedin, Skype and WhatsApp applications. These interactions are recorded in sessions. Each session has a session ID, duration, initial location and the amount of data that is used during the session.

2) Data Collection Module: The data collection module is responsible for storing sessions in a standard format so that they can be analyzed more conveniently. To do that, after the session record is created, it is converted to the JSON format and sent to the private cloud server. The cloud server and the analytics performed over the cloud are illustrated in Figure 2.

3) Normalization Module: Once the session data is transferred to the server, the raw collected data is converted to several metrics of interest. This process is called normalization. In this study, two social verification metrics, namely the social activity rate and sociability factor are defined.

Social Activity Rate: Social activity rate corresponds to the relative amount of data that a user generates when using social networking applications. The absolute data usage of a user is normalized by the data usage of all active users. Social activity rate of a user is a function of the user's short term (daily) and instantaneous social activity rates. Instantaneous social activity rate denotes the data usage by a particular social network application in a single session. Thus, $D_i^{app_x}$ denotes the amount of data from the social network app_x at session i and $t_i^{app_x}$ is the duration of time that the app_x at session i was used. Meanwhile instantaneous social activity rate ($\mathcal{A}_{ins_i}^x$) is formulated as shown in Eq. (1).

$$\mathcal{A}_{ins_i}^{\mathcal{U}_{app_x}^u} = \mathcal{D}_i^{\mathcal{U}_{app_x}^u} / t_i^{\mathcal{U}_{app_x}^u} \tag{1}$$

Eq. (2) formulates user's short term (daily) activity, which denotes the average data usage that is spent on social network app in a session per day.

$$\mathcal{A}_{sh}^{\mathcal{U}_{app_x}^u} = \left(\sum \mathcal{D}_i^{\mathcal{U}_{app_x}^u} / t_i^{\mathcal{U}_{app_x}^u}\right) / \tau_x \tag{2}$$

A weighted sum of consecutive short term social activity rates provide the overall social activity rate ($\mathcal{A}_{overall}^{\mathcal{U}_{app_x}^u}(T_k)$) as shown in Eq. (3).

$$\mathcal{A}_{overall}^{\mathcal{U}_{app_x}^u}(T_k) = \alpha * \mathcal{A}_{sh}^{\mathcal{U}_{app_x}^u}(T_{k-1}) + (1-\alpha) * \mathcal{A}_{sh}^{\mathcal{U}_{app_x}^u}(T_k) \tag{3}$$

The normalized social activity rate ($\mathcal{A}_{normal_i}$) is aggregated overall social activity rates of a user averaged by the maximum social activity rate in the pool of active users as shown in Eq. (4)

$$\mathcal{A}_{normal}^{\mathcal{U}^u} = \sum_{x\in\mathcal{X}} \omega_x \mathcal{A}_{overall}^{\mathcal{U}_{app_x}^u}(T_k) / \operatorname*{argmax}_{u\in\mathcal{U}} \sum_{x\in\mathcal{X}} \omega_x \mathcal{A}_{overall}^{\mathcal{U}_{app_x}^u} \tag{4}$$

Sociability Factor: Sociability of users is not limited to their data consumption but it is also a function of the time they spend on mobile social network applications. Therefore we define the sociability factor metric as another verifier. Similar to the social activity rate, the sociability factor also has instantaneous, short term and global components that ultimately lead to a normalized sociability factor value. Thus, instantaneous sociability factor per app is calculated as the total time that a user spends on a social networking app in a single session as formulated in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2719706, IEEE Access

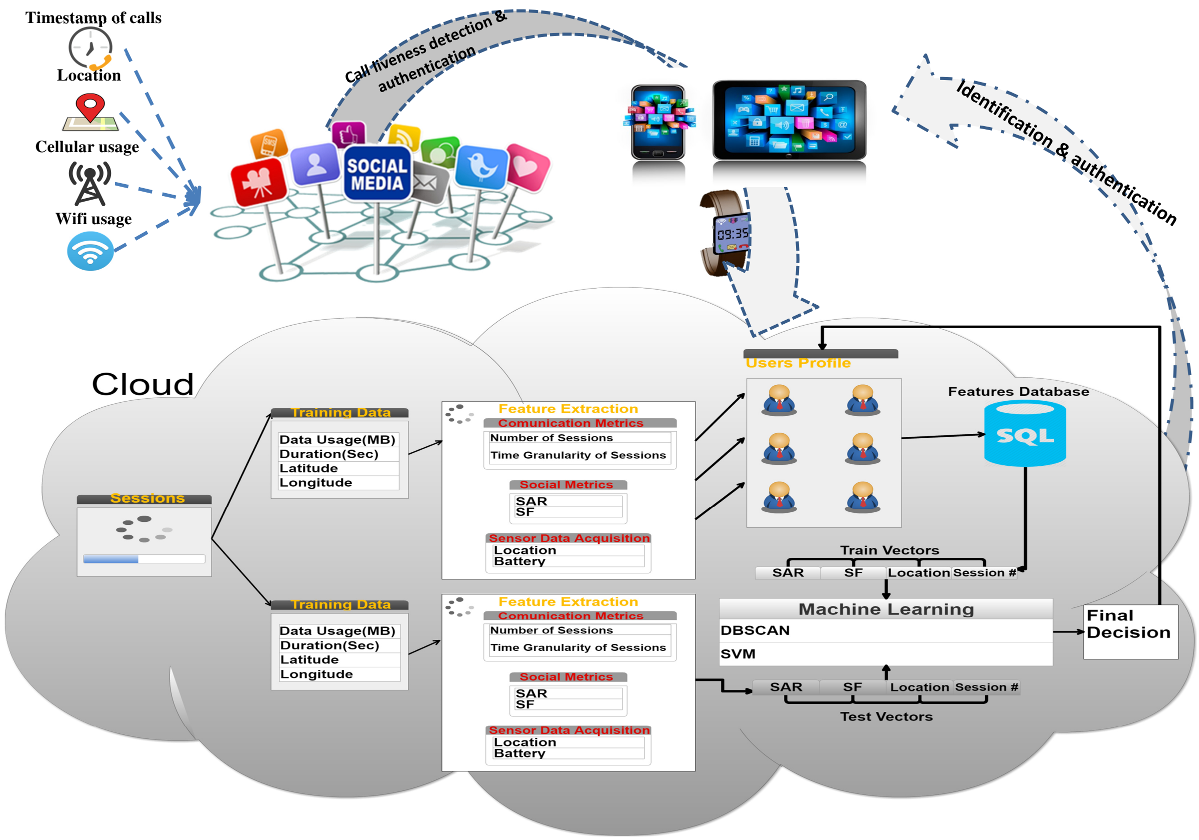


Figure 2. Detailed System Architecture.

Eq. (9). Short term sociability factor ($\mathcal{SF}^{\mathcal{U}^u_{app_x}}_{sh_i}$) is defined as the average time that a user spends on a particular social network app in a session over a short time window, e.g., a day, as formulated in (6) where $t_i^{\mathcal{U}^u_{app_x}}$ stands for duration of session-i of user-u on app_x. As formulated in eq. (7), the overall sociability factor ($\mathcal{SF}^{\mathcal{U}^u_{app_x}}_{overall_i}(T_k)$) is a weighted sum of the short term sociability factors where T_k denotes the $k-th$ short term sociability factor used in the calculation, and β is a weight factor for each mobile social network app. Finally, as expected, the normalized sociability factor ($\mathcal{SF}_{normal_i}$) is the aggregated overall sociability factors of a user scaled by the maximum aggregated sociability factors in the active users pool as shown in Eq. (8).

$$\mathcal{SF}^{\mathcal{U}^u_{app_x}}_{ins_i} = t_i^{\mathcal{U}^u_{app_x}} \tag{5}$$

$$\mathcal{SF}^{\mathcal{U}^u_{app_x}}_{sh} = \left(\sum t_i^{\mathcal{U}^u_{app_x}}\right)/\tau \tag{6}$$

$$\mathcal{SF}^{\mathcal{U}^u_{app_x}}_{overall}(T_k) = \beta * \mathcal{SF}^{\mathcal{U}^u_{app_x}}_{sh}(T_{k-1}) + (1-\beta) * \mathcal{SF}^{\mathcal{U}^u_{app_x}}_{sh}(T_k) \tag{7}$$

$$\mathcal{SF}^{\mathcal{U}^u_{app_x}}_{normal} = \sum_{x\in\mathcal{X}} \omega_x \mathcal{SF}^{\mathcal{U}^u_{app_x}}_{overall}(T_k) / \operatorname*{argmax}_{u\in\mathcal{U}} \sum_{x\in\mathcal{X}} \omega_x \mathcal{SF}^{\mathcal{U}^u_{app_x}}_{overall} \tag{8}$$

4) Training Module: The training module is composed of a learning algorithm that runs in a sliding window of a set of data. The training procedure is based on four factors described in section III-B3; $i)$ social activity rate, $ii)$ sociability factor rate, $iii)$the number of sessions each user has produced per day, and $iv)$ location which is provided by the mobile devices' built-in sensor. In our approach, each user has vectors of feature sets where each vector represents the user behavior throughout a day. The training procedure is updated on a daily basis.

5) Verification Module: To verify user behavior, two different types of learning mechanisms are used, i.e. a supervised learning mechanism and an unsupervised learning mechanism.

Supervised learning mechanism Support Vector Machines (SVMs) are one of the most well-known and effective supervised learning techniques. We used SVM to learn user sociometrics and verify them based on these phenomena. It is worth mentioning that the feature vectors need to be normalized prior to being sent to SVM. In addition, we apply soft normalization which is the output of subtracting each data point from the mean values and scaling by twice the standard deviation as shown in Eq. (9).

$$\forall_{p \in \mathcal{P}} | p = (p - \mu)/\sigma^2 \tag{9}$$

Unsupervised learning mechanism We also used Density-Based clustering of applications with noise (DBSCAN) which is an unsupervised learning technique. DBSCAN groups the data points that are nearest neighbors of each other with the ultimate goal of forming dense regions. Outliers whose nearest neighbors are not close enough, are clustered in low density regions [61].

IV. Performance Evaluation

The performance of the proposed technique is evaluated by using the TrackMaison framework (Track My Activity in Social Networks) proposed in [12] which collects data usage, activity duration, location and usage frequency of project participants on five popular social network applications, namely Facebook, Twitter, LinkedIn, Skype and WhatsApp. The back-end server computes the social activity rate and sociability factor by using the data rates and session duration as formulated in in Eq.(4) and in Eq.(8). The front-end connectivity of the testbed is provided by Android-based tablets that continuously push data collected from 13K sessions in a two-month time window. In this paper, six representative users out of the participant set are chosen. It is worthwhile mentioning that the algorithm filled any missing data points with the mean value up to that point. The results are based on different set of values for α and β in Eq. (3) and Eq. (7). Figures 3 and 4 illustrate users' real data within a period of 3 months approximately.

As mentioned before, social activity rate denotes the amount of data that a user spends on social network applications whereas sociability factor is a function of the duration that a user interacts with their mobile device. The results are based on different set of values for α and β in (3) and (7). Based on the results, the users can be categorized into three groups of highly active, moderately active and least active users. Based on the defined categories, users 2 and 3 are highly active users, 1, 4 and 5 are moderately active users and user 6 is the least active user. Moreover users 4, 5 and 6 spent significantly short time but consumed high volume of data. It can be assumed that these users intend to access multimedia contents like movies and photos rather than regular browsing activity. It is worthwhile noting that connected IoT devices, and mobile applications that run on those devices are prone to security vulnerabilities as a result of unauthorized access [62]. Therefore, this paper does not aim to replace biometric authentication in IoT-integrated platforms or personalized devices but aims to strengthen existing password, fingerprint, face or speech recognition-based authentication by incorporating knowledge based spatiotemporal abstraction. That being said, a performance metric, namely the authentication error probability is defined in order to evaluate the disruption probability in continuous authentication of users on connected mobile devices. In this paper the authentication error probability is cumulative, and it denotes the cases when the device falls back to one of the biometric authentication and liveness detection methods which is proposed by Akhtar et al [63].

Two machine learning (ML) approaches are used, namely Support Vector Machines (SVM) [64] and Density-based spatial clustering of applications with noise (DBSCAN) [65] to authorize user access to mobile devices. SVM is a supervised learning method that basically defines hyperplanes which separate the data into different groups while DBSCAN groups the data points that are nearest neighbors of each other, and aims at forming dense regions. We also present a set of selected users where we randomly injected daily behavioral patterns of other users to each of these users for randomly selected five days after behavioral patterns have been learned. The experiments have been carried out under the following two scenarios:

1) *Normal condition* denotes the scenarios where user identities were not spoofed, and the only possible false alarm in continuous authentication could be the false rejection. This results in the system to fall back to biometric authentication to validate the user, even though it is a legitimate user. The aim is to minimize false rejections.

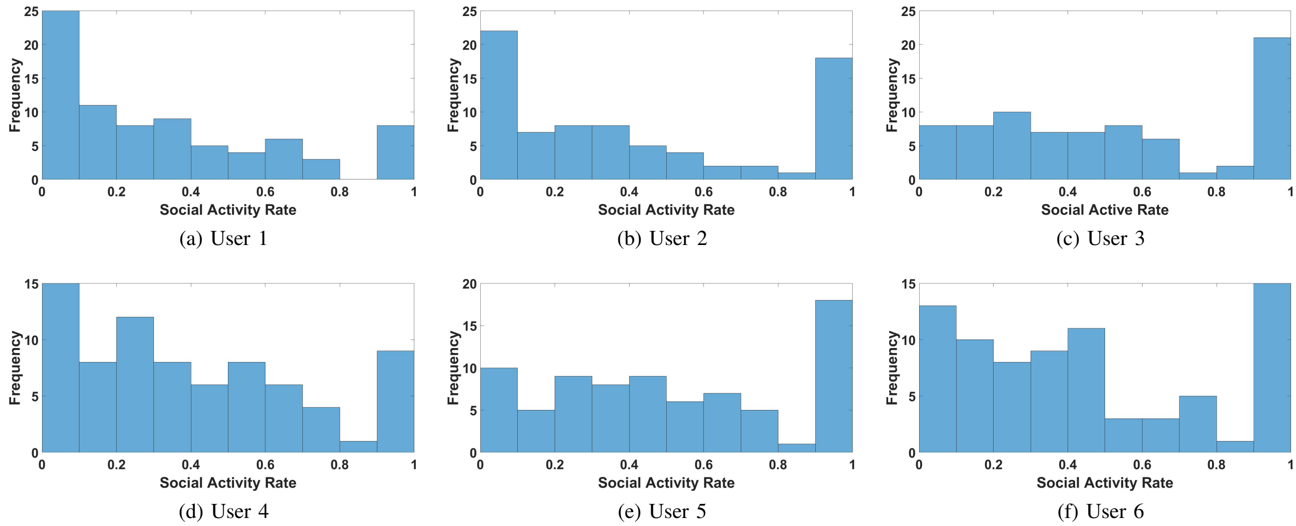
2) *Anomalous condition* denotes the situations where we have created spoofed identities by mapping a randomly selected user's patterns onto the records of a particular user after the continuous authentication system has been trained to verify the social behavioral context of the corresponding user. This could result in false acceptance which is also aimed to be minimized.

A. Experimental results under normal condition

1) Verification by SVM: In this section the results of applying SVM on the dataset are presented. The behavioral data have been collected for 76 days. The system was trained within the first week of data and SVM was set to six different classes corresponding to each user. As mentioned before two normalization techniques are applied, namely the soft normalization and hard normalization to the results of the ML processes. The proposed framework is also improved by dynamically adjusting the contextual parameter of weights α and β in long term sociability signature which is shown in Eq. (3) and Eq. (7). To be able to analyze the impact of the contextual parameters on the performance of the proposed framework, wide range of values have been set in the form of ((α)-(1-α)) for social activity rate, and in the form of ((β)-(1-β)) for sociability factor as follows: 15%-85%, 30%-70%, 50%-50%, 70%-30% and 85%-15% where each set refers to α (β) and 1-α (1-β). For example, 15%-85% means α and β are equal to 15%.

Figure 5 illustrates the results under SVM with soft normalization technique. By applying different values for α and β, it can be concluded that the performance of verification by using SVM has better results when α or β are low. The proposed framework can verify User 1 with 100% under 15%-85% and 50%-50% situations otherwise the verification success ratio is approximately 95% which means that the cumulative authentication error probability for user 1 is 5% from day 7 through 76. These performance metrics for User 2 for all settings is almost the same, which is approximately 83% except for 85%-15% which differentiated on day 70 and ended by 75%. The framework could be able to verify user 3 with 100%

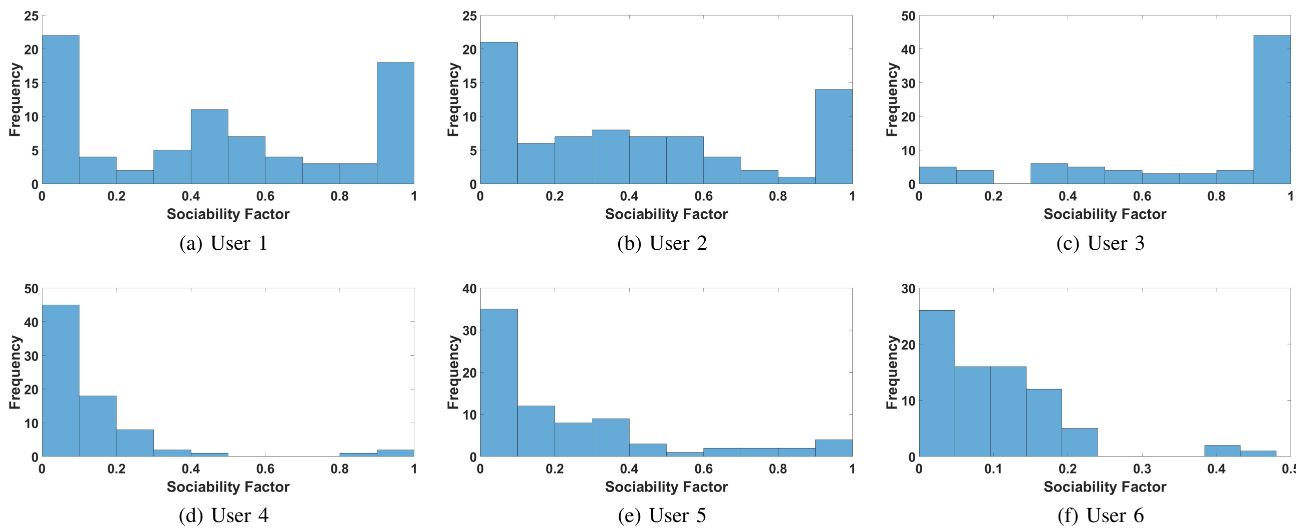
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2719706, IEEE Access



(a) User 1 (b) User 2 (c) User 3

(d) User 4 (e) User 5 (f) User 6

Figure 3. Users' trait for the following representative user profiles (a) User-1, (b) User-2, (c) User-3, (d) User-4, (e) User-5, (f) User-6. Frequency refers to the number of occurrences of social activity rate.



(a) User 1 (b) User 2 (c) User 3

(d) User 4 (e) User 5 (f) User 6

Figure 4. Users' trait for the following representative user profiles (a) User-1, (b) User-2, (c) User-3, (d) User-4, (e) User-5, (f) User-6. Frequency refers to the number of occurrences of sociability factor.

success rate when α and β are 15%. User 4 experiences 100% success rate in verification for all set of configurations. This is mainly because the user's exhibiting a visibly inactive behavior most of the days during the data collection. User 5 shows better performance under 15%-85% and 30%-70% settings. User 6 is more sensitive to each setting. The framework has approximately 100% success rate in verifying User 6 under the case, 15%-85%, and then for the rest of the settings, 50%-50% has the best match for the user verification. To summarize, the proposed framework performs better when the system relies more on long term activity than short term.

B. Verification by DBSCAN

The performance of the proposed framework under the DBSCAN algorithm on the collected dataset was also evaluated. Figure 6 illustrates the results of continuous authentication through an unsupervised approach, namely DBSCAN. Similarly, verification performance is better when long term activity is assigned higher weights. However the best settings for the users under DBSCAN are 15%-85% and also 50%-50%. At the end of the training period, the authentication error probability for all users is below 0.1 % except user 6. This corresponds to the situation when user behavior on social network applications has been verified as an anomaly so the back-end server sends a biometric authentication triggering signal to the front-end device. In Figure 6, for each user, the anomalies marked by the ML-based continuous authentication (which are observed by an increase authentication error probability in the plots) is due to users' having different social behavioral patterns in weekdays and weekends. For instance, user-1 has been found anomalous from day-13 to day-15 with increasing authentication error probability, whereas user-4 has been found anomalous from day 26th to day 29. Over the weekends, most participants follow different profiles, provide less data as they interacted less frequently over social networking applications. Furthermore, the participants' travel patterns change over the weekends, which in turn, affects the feature set of the ML-based authentication framework.

C. Experimental results under anomalous condition

Once the continuous authentication platform has been trained, in order to imitate the situation where identities were spoofed which can be due to exchanging mobile devices between users or stolen devices, we introduce artificial noisy patterns to the social behavioral profile of each user in particular days. The noisy patterns are created by copying a usage pattern on the records that belong to another user.

This scenario mimics the situation where random user pairs were selected to use each other's mobile device for five consecutive days after the platform has been trained.

Each figure illustrates the authentication error probability (AEP) under the proposed system during the 5-day period after a user's behavior has been learned (i.e., converged authentication error probability). The time when the user behavior has been learned also denotes the time when the smartphone can be safely recruited for opportunistic or participatory sensing purposes within the IoT architecture. User is recruited for opportunistic or participatory sensing purposes in the IoT context has to be in an implicit manner. Thus, the authenticity of the smart device user should not undergo biometric authentication frequently. As (10) formulates, AEP_t stands for the disruption probability that results in after biometric authentication has been triggered: The ratio of the cumulative value of false rejections or true rejections (FR and TR) starting from the beginning of training moving to the end of the time of interest (t) to the cumulative value of total acceptances and rejections. Disruption (AEP) affects user experience negatively and may result in de-incentivizing users in participating IoT sensing through their smart mobile devices. On the other hand, false acceptance may lead to reduced trustworthiness of the sensory data acquired through built-in sensors of these devices. Therefore, we also present the false acceptance probability, and the impact of the contextual parameter weights ($\alpha - \beta$) on the number of false acceptances.

$$AEP_t = \frac{\sum_{k=0}^{t} (FR_k + TR_k)}{\sum_{k=0}^{i} (FR_k + FA_k + TR_k + TA_k)} \tag{10}$$

For each user, five days have been selected based on the initial training duration of the continuous authentication platform. The proposed framework is also improved by dynamically adjusting the contextual weight parameters for social activity rate and sociability factor which is shown in Eq. (3) and Eq. (7). Similar to the tests in the previous subsection, in order to be able to analyze the impact of α and β on the performance of the proposed framework, wide range of values have been set in the form of ((α)-(1-α)) for social activity rate, and in the form of ((β)-(1-β)) for sociability factor as follows: 15%-85%, 30%-70%, 50%-50%, 70%-30% and 85%-15% where each set respectively refers to (α) and (1- α). For example, 15%-85% means α and β are set to 15%.

Figures 7-11 show the result of different configuration of α and β each corresponding to different user. In the figures, the y-axis shows the authentication error probability and the x-axis represents the random days that other users' behavioral patterns were injected to simulate identity spoofing. Each plot carries two information, the normal condition (Gray bars) and anomalous condition (Black bars) referring to the results that were collected under identity spoofing scenarios. By providing these two information side by side, the performance of the system in detecting spoofed identities could be highlighted by considering true rejection (TR) and false rejection (FR) results.

Figure 7 illustrates the situation where α and β are equal to 15%. Given this value, the system is able to detect the noisy points with 0% error rate except for user 4 where the system has one FA out of five spoofing attempts. AEP rate for the situation where $\alpha = \beta = 30\%$ increased as shown in Figure 8. Under this setting, users 1, 3, 4 and 6 experience 100% TR but the TR probability for users 2 and 5 drops down to 80% and 20%, respectively. The AEP rate declined when α and β are equal to 50% as shown in Figure 9. In this setting, users 1, 3, 4, and 6 has 100% TR probability. However, User 5 and 2 recorded one FA. Figure 10 shows the situation where α and β are equal to 70%. In this case, users 1, 2, 4 and 5 scored 100% success with TR whereas user 3 recorded one FA and the system could not recognize any of the five spoofing attempts on the device possessed by user 6. The last figure 11 shows the situation where α and β are equal to 85% which indicates the worst results for user verification. In this configuration, users 2, 4 and 6 recorded five FA and user 3 was able to catch only one TR point; whereas users 5 and 1 were able to catch just two TR out of five spoofing attempts.

Based on the discussions on Figures 7-11, the following conclusions can be made: When the contextual weights, α and β, are increased from 15% to 85%, the error rate of the system to verify the genuine users increases. Based on the results presented in the figures above, the contextual weights can be set to 15% and 50% as these lead to the best values with least error (disruption) rate for the proposed system in the detection of genuine smartphone users. By studying the given data, the minimum number of FAs for all users is scored in 15% setting, and the system experiences just one false acceptance out of 30 spoofing attempts which is close to 3% error ratio in verification. This rate increased approximately to 26% through eight FAs out of thirty selected points when α and β are set to 30%. The performance of the system under the setting where $\alpha = \beta = 50\%$ is improved by just experiencing two FAs which means close to 6% error rate. However the system scored 20% error ratio when α and β were set to 70%, and the worst case is experienced when α and β were set to 85% by having 83% error rate with 25 false acceptance points out of 30 spoofing attempts.

V. OPEN ISSUES AND FUTURE DIRECTIONS

In the experimental results, when the contextual parameter weights are properly set, the probability of FA is reasonably low. However, the impact of applying different machine learning algorithms and the possibility of further reducing FA should be investigated. This article aims at presenting the concept of continuous verification; however feasibility study of various machine learning algorithms on the proposed system is included

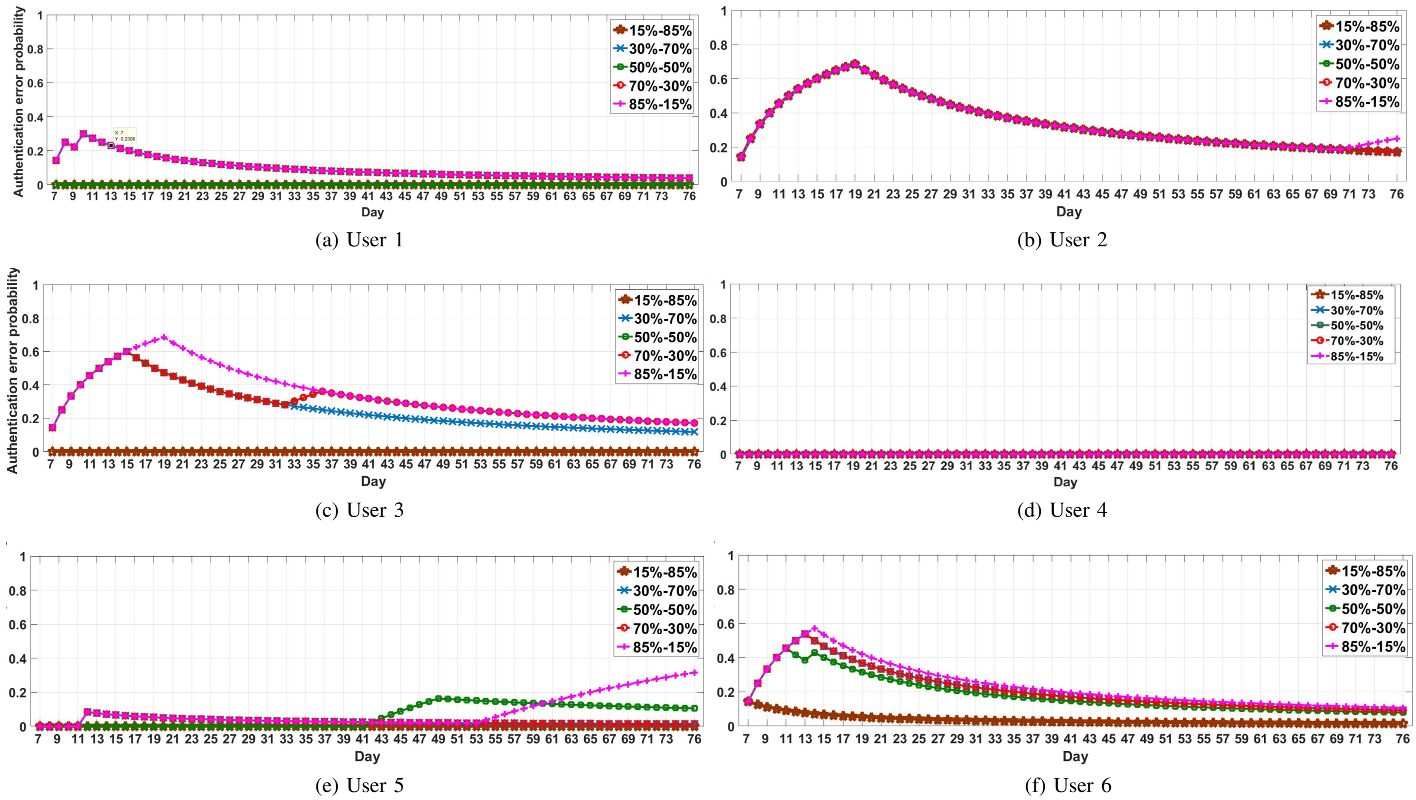


Figure 5. Authentication error probability under SVM with soft-normalization algorithm for the following representative user profiles (a) User-1, (b) User-2, (c) User-3, (d) User-4, (e) User-5, (f) User-6

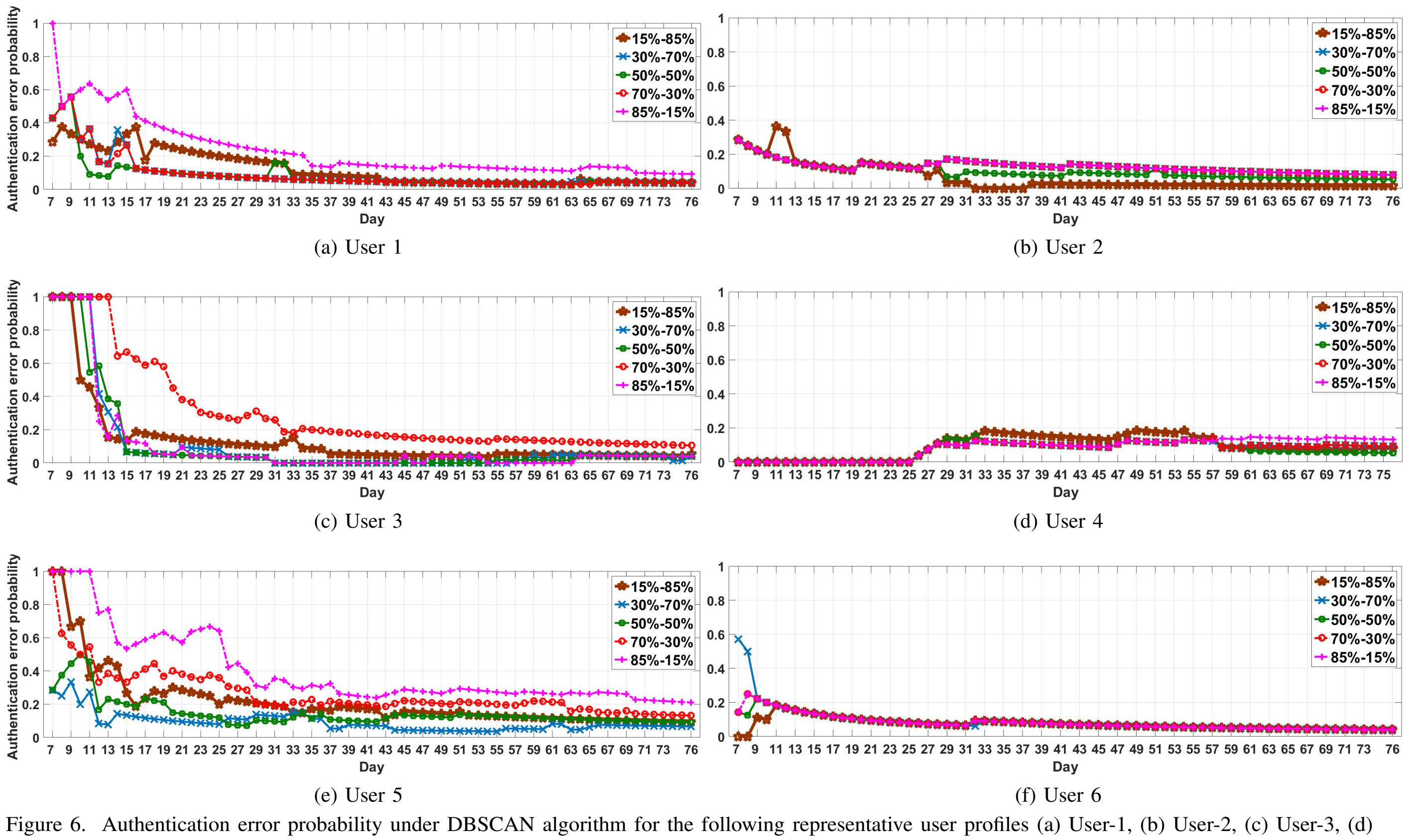


Figure 6. Authentication error probability under DBSCAN algorithm for the following representative user profiles (a) User-1, (b) User-2, (c) User-3, (d) User-4, (e) User-5, (f) User-6

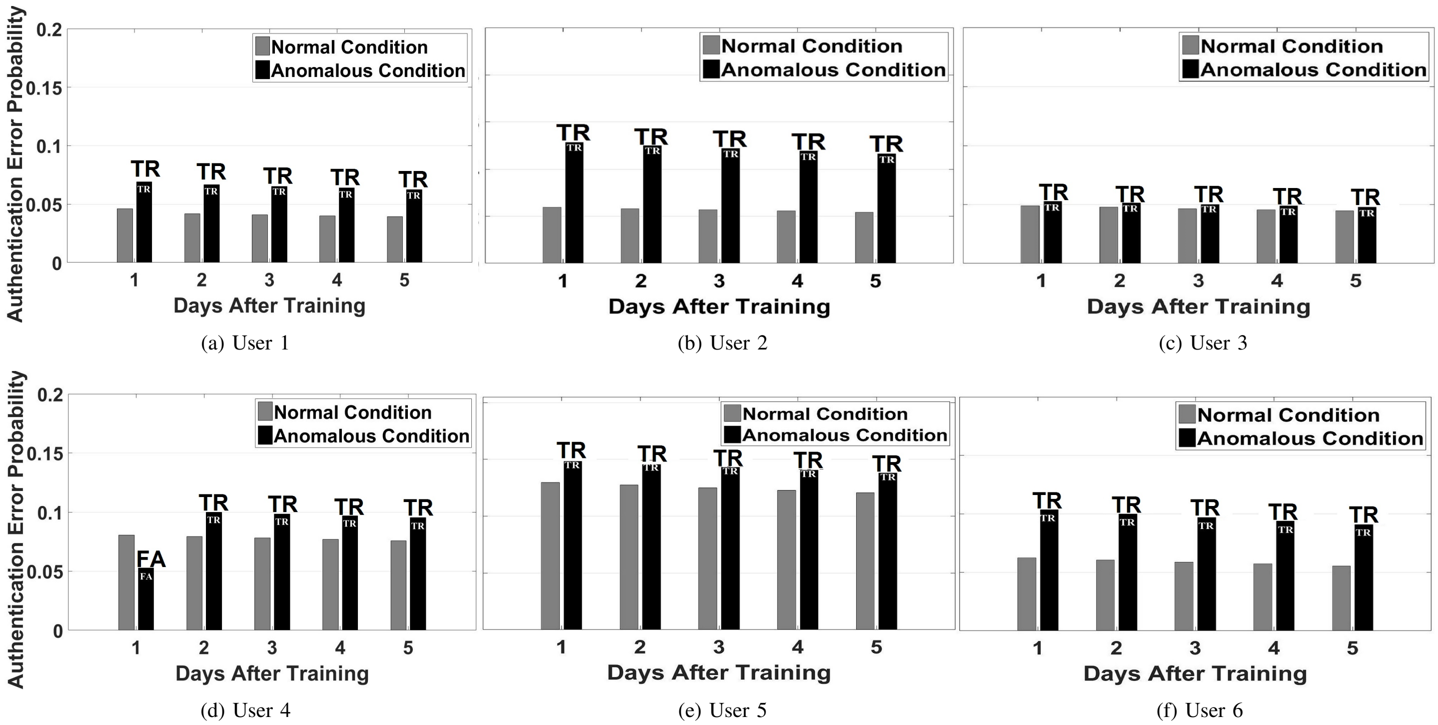


Figure 7. Probability of triggering biometric authentication due to authentication error under DBSCAN with spoofing identities when α and β equal to 15%

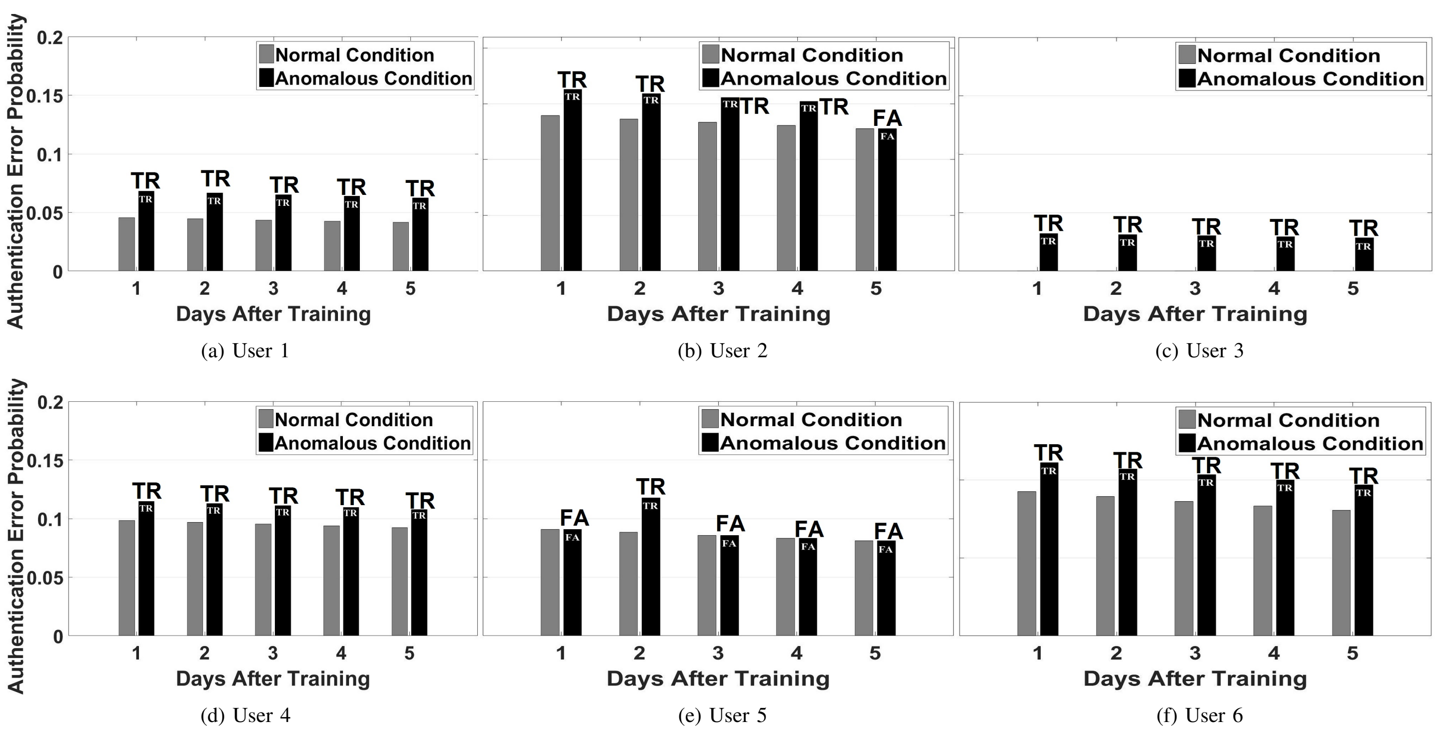
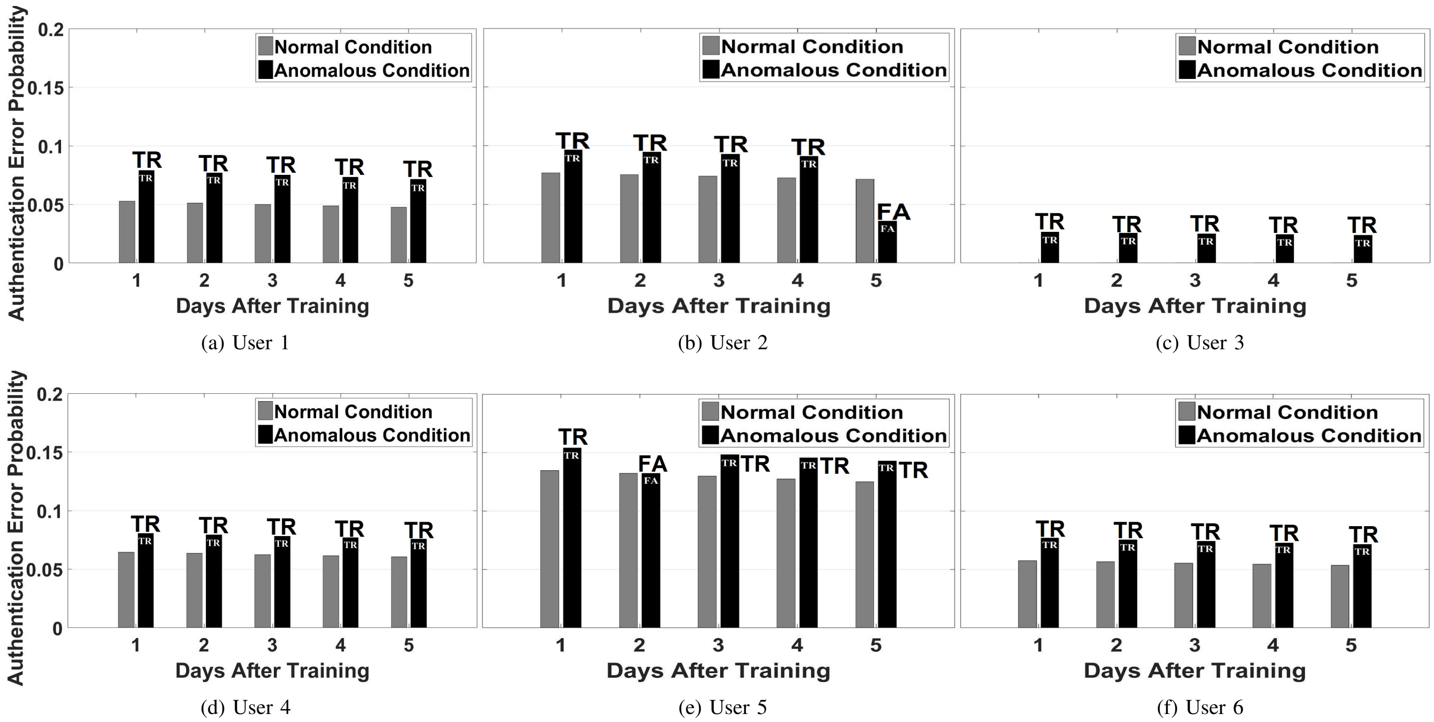


Figure 8. Probability of triggering biometric authentication due to authentication error (TR + FR) under DBSCAN with spoofing identities when α and β equal to 30%

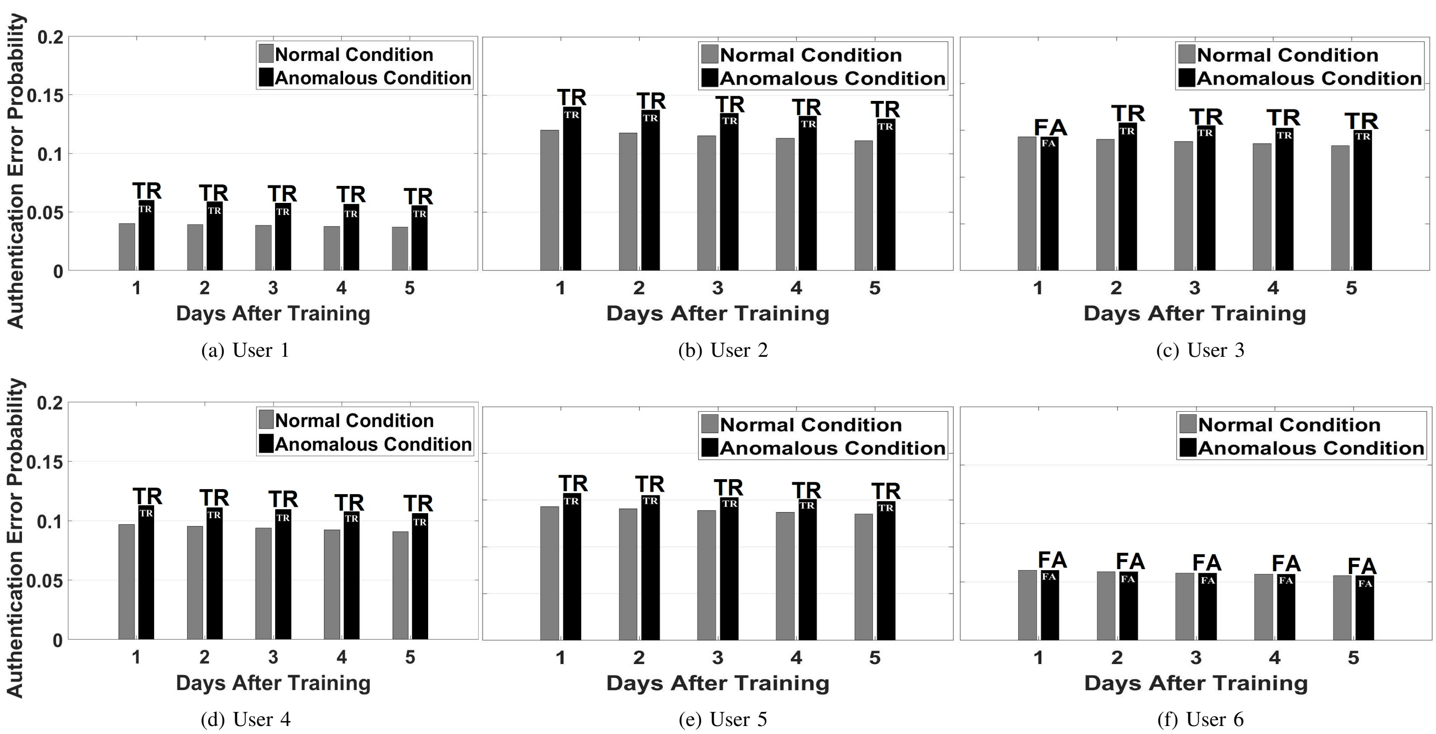
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2719706, IEEE Access



(a) User 1 (b) User 2 (c) User 3

(d) User 4 (e) User 5 (f) User 6

Figure 9. Probability of triggering biometric authentication due to authentication error under DBSCAN with spoofing identities when α and β equal to 50%

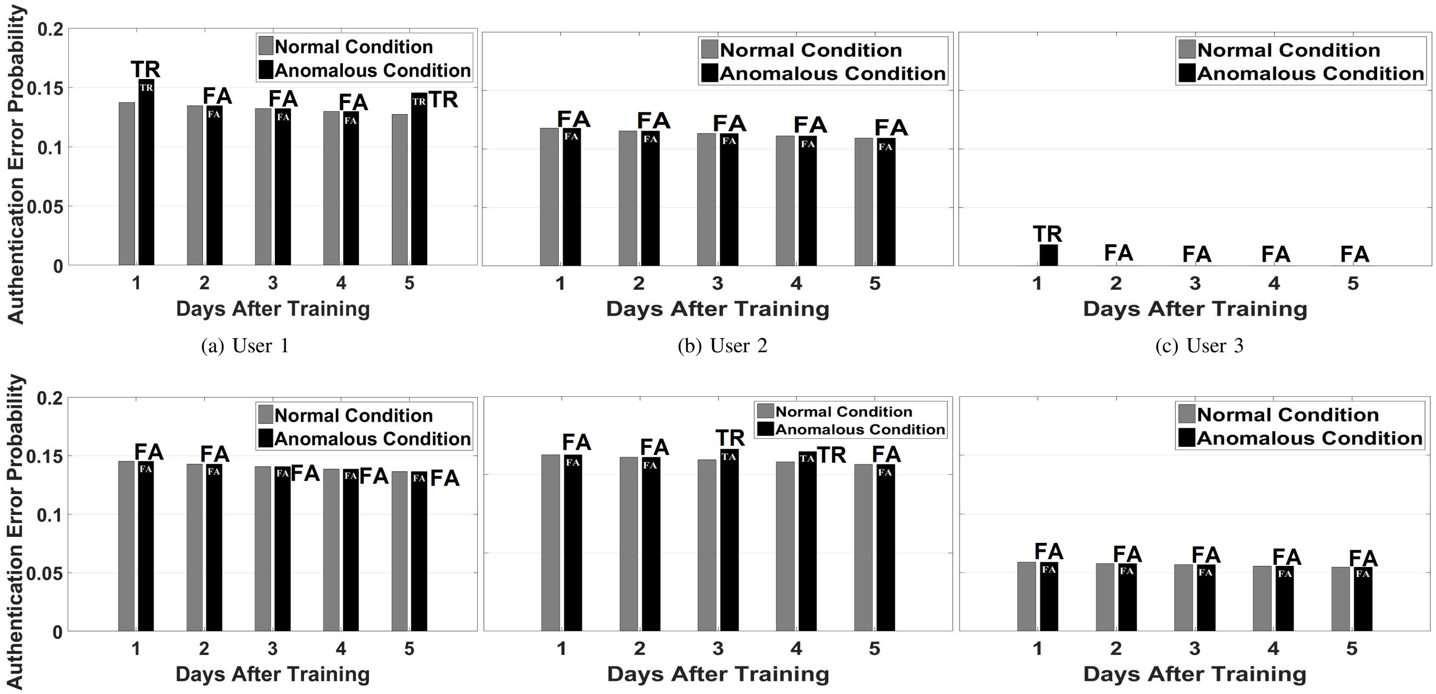


(a) User 1 (b) User 2 (c) User 3

(d) User 4 (e) User 5 (f) User 6

Figure 10. Probability of triggering biometric authentication due to authentication error under DBSCAN with spoofing identities when α and β equal to 70%

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2719706, IEEE Access



(a) User 1 (b) User 2 (c) User 3

(d) User 4 (e) User 5 (f) User 6

Figure 11. Probability of triggering biometric authentication due to authentication error under DBSCAN with spoofing identities when α and β equal to 85%

in our future research agenda. The members of the participant pool in this research demonstrate fairly similar behavior to each other (i.e. mostly graduate students). Inclusion of a heterogeneous participant pool in the experiments is expected to increase the impact of the normalization module of the system, diversify the behavioral clusters and in turn, result in lower FAs. Moreover, including additional applications in the analysis would improve the accuracy. However, improvement in the accuracy would come at the expense of additional computational overhead. Hence, investigating the optimal number of apps to ensure the trade-off between computational complexity and accuracy in continuous verification is another important future research direction.

In addition to all, we are currently extending the feature set and collecting a richer set of data to reduce verification and training duration. In addition, energy-efficiency is an important concern for mobile platforms; therefore energy-efficient continuous verification mechanisms are also being developed within the ongoing research efforts.

VI. CONCLUSION

In Internet of Things (IoT) contexts, smart user devices can be recruited for participatory or opportunistic sensing campaigns. Verification of genuine users of these devices is of paramount importance for the following reasons: High rejection rates may trigger biometric authentication and may de-incentivize users to offer their built-in sensors as a service whereas high false acceptances may result in reduced trustworthiness of the sensory data. With the convergence of IoT and social networks, Social Internet of Things has emerged which has many advantages including network navigability, service scalability and increased trustworthiness of acquired data. This paper has studied continuous verification in SIoT where online behaviometrics of mobile users collected via smart phones is considered by extracting features from smartphone sensors and users' social network interactions. We have presented a continuous verification scheme that uses social behaviometrics collected from a set of users. We have used real traces collected over several months. Those traces are sent to a cloud server and analyzed with two machine learning techniques, namely the Support Vector Machines (SVM) and Density-Based clustering of applications with noise (DBSCAN). Our results show that genuine users can be verified without any disruption 97% of the time whereas the users can keep using the devices 90% of the time without any disruption.

ACKNOWLEDGMENT

This material is based upon works supported by the Center for Identification Technology and Research (CITeR) and the U.S. National Science Foundation (NSF) under Grant Numbers IIP-1068055 and CNS-1464273, and the Natural Sciences and Engineering Research Council of Canada (NSERC) under RGPIN/2017-04032.

REFERENCES

[1] M. Pouryazdan, C. Fiandrino, B. Kantarci, D. Kliazovich, T. Soyata, and P. Bouvry, "Game-theoretic recruitment of sensing service providers for trustworthy cloud-centric internet-of-things (IoT) applications," in *IEEE GLOBECOM Workshops*, 2016.

[2] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer networks*, vol. 54, no. 15, pp. 2787–2805, 2010.

[3] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (iot): A vision, architectural elements, and future directions," *Future generation computer systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[4] K. Gama, L. Touseau, and D. Donsez, "Combining heterogeneous service technologies for building an internet of things middleware," *Computer Communications*, vol. 35, no. 4, pp. 405–417, 2012.

[5] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: current state and future challenges," *IEEE Communications Magazine*, vol. 49, no. 11, 2011.

[6] Y. X. Wazir Zada Khan, M. Y. Aalsalem, and Q. Arshad, "Mobile phone sensing systems: A survey," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 402–427, 2013.

[7] B. Kantarci, M. Erol-Kantarci, and S. Schuckers, "Towards secure cloud-centric internet of biometric things," in *Cloud Networking (CloudNet), 2015 IEEE 4th International Conference on*. IEEE, 2015, pp. 81–83.

[8] A. R. Mohammad, K. Mohiuddin, M. Irfan, and M. Moizuddin, "Cloud the mainstay: growth of social networks in mobile environment," in *Cloud & Ubiquitous Computing & Emerging Technologies (CUBE), 2013 International Conference on*. IEEE, 2013, pp. 14–19.

[9] P. research center. Social media fact sheet. [Online]. Available: http://www.pewinternet.org/fact-sheet/social-media/

[10] "Ericsson mobility report," ITU, 2016. [Online]. Available: http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf

[11] B. Schneier, "A taxonomy of social networking data," *IEEE Security & Privacy*, vol. 8, no. 4, pp. 88–88, 2010.

[12] F. Anjomshoa, M. Catalfamo, D. Hecker, N. Helgeland, A. Rasch, B. Kantarci, M. Erol-Kantarci, and S. Schuckers, "Sociability assessment and identification of smartphone users via behaviormetric software," in *IEEE Symposium on Computers and Communications (ISCC)*, 2016, pp. 1084–1089.

[13] S. B. Pan, D. Moon, Y. Gil, D. Ahn, and Y. Chung, "An ultra-low memory fingerprint matching algorithm and its implementation on a 32-bit smart card," *IEEE transactions on consumer electronics*, vol. 49, no. 2, pp. 453–459, 2003.

[14] S. Thavalengal, P. Bigioi, and P. Corcoran, "Iris authentication in handheld devices-considerations for constraint-free acquisition," *IEEE Transactions on Consumer Electronics*, vol. 61, no. 2, pp. 245–253, 2015.

[15] D.-J. Kim, K.-W. Chung, and K.-S. Hong, "Person authentication using face, teeth and voice modalities for mobile device security," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, pp. 2678–2685, 2010.

[16] Y. Kim, J.-H. Yoo, and K. Choi, "A motion and similarity-based fake detection method for biometric face recognition systems," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 2, pp. 756–762, 2011.

[17] K. Lee and H. Byun, "A new face authentication system for memory-constrained devices," *IEEE Transactions on Consumer Electronics*, vol. 49, no. 4, pp. 1214–1222, 2003.

[18] Z. Hammook, J. Misic, and V. Misic, "Crawling researchgate. net to measure student/supervisor collaboration," in *2015 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2015, pp. 1–6.

[19] M. I. Akbas, R. N. Avula, M. A. Bassiouni, and D. Turgut, "Social network generation and friend ranking based on mobile phone data," in *Communications (ICC), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1444–1448.

[20] C. Bo, L. Zhang, X.-Y. Li, Q. Huang, and Y. Wang, "Silentsense: silent user identification via touch and movement behavioral biometrics," in *Proceedings of the 19th annual international conference on Mobile computing & networking*. ACM, 2013, pp. 187–190.

[21] A. Messerman, T. Mustafić, S. A. Camtepe, and S. Albayrak, "Continuous and non-intrusive identity verification in real-time environments based on free-text keystroke dynamics," in *Biometrics (IJCB), 2011 International Joint Conference on*. IEEE, 2011, pp. 1–8.

[22] J. Roth, X. Liu, and D. Metaxas, "On continuous user authentication via typing behavior," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4611–4624, 2014.

[23] J. V. Monaco, N. Bakelman, S.-H. Cha, and C. C. Tappert, "Developing a keystroke biometric system for continual authentication of computer users," in *Intelligence and Security Informatics Conference (EISIC), 2012 European*. IEEE, 2012, pp. 210–216.

[24] D. Roggen, M. Wirz, G. Tröster, and D. Helbing, "Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods," *arXiv preprint arXiv:1109.1664*, 2011.

[25] S.-W. Lee and K. Mase, "Recognition of walking behaviors for pedestrian navigation," in *(CCA'01). Proceedings of the IEEE International Conference on Control Applications*, 2001, pp. 1152–1155.

[26] M. Sultana, P. P. Paul, and M. Gavrilova, "A concept of social behavioral biometrics: Motivation, current developments, and future trends," in *International Conf. on Cyberworlds*. IEEE, 2014, pp. 271–278.

[27] R. V. Yampolskiy and V. Govindaraju, "Behavioural biometrics: a survey and classification," *International Journal of Biometrics*, vol. 1, no. 1, pp. 81–113, 2008.

[28] F. Anjomshoa, B. Kantarci, M. Erol-Kantarci, and S. Schuckers, "Detection of spoofed identities on smartphones via. sociability metrics," in *IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 6911–6916.

[29] L. Atzori, A. Iera, G. Morabito, and M. Nitti, "The social internet of things (siot)–when social networks meet the internet of things: Concept, architecture and network characterization," *Computer networks*, vol. 56, no. 16, pp. 3594–3608, 2012.

[30] L. Atzori, A. Iera, and G. Morabito, "From" smart objects" to" social objects": The next evolutionary step of the internet of things," *IEEE Communications Magazine*, vol. 52, no. 1, pp. 97–105, 2014.

[31] A. M. Ortiz, D. Hussein, S. Park, S. N. Han, and N. Crespi, "The cluster between internet of things and social networks: Review and research challenges," *IEEE Internet of Things Journal*, vol. 1, no. 3, pp. 206–215, 2014.

[32] L. E. Holmquist, F. Mattern, B. Schiele, P. Alahuhta, M. Beigl, and H.-W. Gellersen, "Smart-its friends: A technique for users to easily establish connections between smart artefacts," in *international conference on Ubiquitous Computing*. Springer, 2001, pp. 116–122.

[33] L. Ding, P. Shi, and B. Liu, "The clustering of internet, internet of things and social network," in *Knowledge Acquisition and Modeling (KAM), 2010 3rd International Symposium on*. IEEE, 2010, pp. 417–420.

[34] J. An, X. Gui, W. Zhang, and J. Jiang, "Nodes social relations cognition for mobility-aware in the internet of things," in *International Conference on Internet of Things (iThings/CPSCom), 4th International Conference on Cyber, Physical and Social Computing*. IEEE, 2011, pp. 687–691.

[35] H. Zhang and M. Li, "Security vulnerabilities of an remote password authentication scheme with smart card," in *Consumer Electronics, Communications and Networks (CECNet), 2011 International Conference on*. IEEE, 2011, pp. 698–701.

[36] W. Meng, D. S. Wong, S. Furnell, and J. Zhou, "Surveying the development of biometric user authentication on mobile phones," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1268–1293, 2015.

[37] A. Dantcheva, P. Elia, and A. Ross, "What else does your biometric data reveal? A survey on soft biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 3, pp. 441–467, 2016.

[38] A. Poursaberi, J. Vana, S. Mracek, R. Dvora, S. N. Yanushkevich, M. Drahansky, V. P. Shmerko, and M. L. Gavrilova, "Facial biometrics for situational awareness systems," *IET biometrics*, vol. 2, no. 2, pp. 35–47, 2013.

[39] J. Liu-Jimenez, R. Sanchez-Reillo, and B. Fernandez-Saavedra, "Iris biometrics for embedded systems," *IEEE transactions on very large scale integration (vlsi) systems*, vol. 19, no. 2, pp. 274–282, 2011.

[40] H. Lv and W.-Y. Wang, "Biologic verification based on pressure sensor keyboards and classifier fusion techniques," *IEEE Transactions on Consumer Electronics*, vol. 52, no. 3, pp. 1057–1063, 2006.

[41] E. Shi, Y. Niu, M. Jakobsson, and R. Chow, "Implicit authentication through learning user behavior," in *International Conference on Information Security*. Springer, 2010, pp. 99–113.

[42] H. Gascon, S. Uellenbeck, C. Wolf, and K. Rieck, "Continuous authentication on mobile devices by analysis of typing motion behavior." in *Sicherheit*. Citeseer, 2014, pp. 1–12.

[43] H. Khan, A. Atwater, and U. Hengartner, "Itus: an implicit authentication framework for android," in *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 2014, pp. 507–518.

[44] T. Stockinger, "Implicit authentication on mobile devices," in *The Media Informatics Advanced Seminar on Ubiquitous Computing*. Citeseer, 2011.

[45] H. Khan and U. Hengartner, "Towards application-centric implicit authentication on smartphones," in *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications*. ACM, 2014, p. 10.

[46] A. De Luca, A. Hang, F. Brudy, C. Lindner, and H. Hussmann, "Touch me once and i know it's you!: implicit authentication based on touch screen patterns," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 987–996.

[47] T. Feng, Z. Liu, K.-A. Kwon, W. Shi, B. Carbunar, Y. Jiang, and N. Nguyen, "Continuous mobile authentication using touchscreen gestures," in *Homeland Security (HST), 2012 IEEE Conference on Technologies for*. IEEE, 2012, pp. 451–456.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2719706, IEEE Access

[48] X. Chen, R. Chandramouli, and K. P. Subbalakshmi, "Scam detection in twitter," in *Data Mining for Service*. Springer, 2014, pp. 133–150.
[49] A. Louni, A. Santhanakrishnan, and K. Subbalakshmi, "Identification of source of rumors in social networks with incomplete information," *arXiv preprint arXiv:1509.00557*, 2015.
[50] M. Sultana, P. P. Paul, and M. L. Gavrilova, "Online user interaction traits in web-based social biometrics," *Comput Vis. Image Process Intell Syst Multimedia Technol*, pp. 177–190, 2014.
[51] M. Sultana, P. P. Paul, and M. Gavrilova, "Social behavioral biometrics: An emerging trend," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 29, no. 08, p. 1556013, 2015.
[52] N. Lathia, V. Pejovic, K. K. Rachuri, C. Mascolo, M. Musolesi, and P. J. Rentfrow, "Smartphones for large-scale behavior change interventions." *IEEE Pervasive Computing*, vol. 12, no. 3, pp. 66–73, 2013.
[53] A. Mehrotra, V. Pejovic, and M. Musolesi, "Sensocial: a middleware for integrating online social networks and mobile sensing data streams," in *Proceedings of the 15th International Middleware Conference*. ACM, 2014, pp. 205–216.
[54] A. S. Crandall and D. J. Cook, "Behaviometrics for identifying smart home residents," in *Human Aspects in Ambient Intelligence*. Springer, 2013, pp. 55–71.
[55] A. B. Budurusubmi and S. S. Yau, "An effective approach to continuous user authentication for touch screen smart devices," in *IEEE International Conference on Software Quality, Reliability and Security (QRS)*, Aug 2015, pp. 219–226.
[56] A. Mosenia, S. Sur-Kolay, A. Raghunathan, and N. K. Jha, "Caba: Continuous authentication based on bioaura," *IEEE Transactions on Computers*, vol. 66, no. 5, pp. 759–772, 2016.
[57] J. Schobel, R. Pryss, M. Schickler, and M. Reichert, "Towards flexible mobile data collection in healthcare," in *IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS)*, 2016, pp. 181–182.
[58] M. Batty, K. W. Axhausen, F. Giannotti, A. Pozdnoukhov, A. Bazzani, M. Wachowicz, G. Ouzounis, and Y. Portugali, "Smart cities of the future," *The European Physical Journal Special Topics*, vol. 214, no. 1, pp. 481–518, 2012.
[59] R. Khatoun and S. Zeadally, "Smart cities: concepts, architectures, research opportunities," *Communications of the ACM*, vol. 59, no. 8, pp. 46–57, 2016.
[60] C. Ziegler, "Implicit authentication 2.0: Behavioural biometrics in smart environments," in *Human Computer Interaction in the Internet of Things Era*, E. von Zezschwitz et al., Ed. University of Munich, 2015, pp. 100–107.
[61] D. Hristova, M. J. Williams, M. Musolesi, P. Panzarasa, and C. Mascolo, "Measuring urban social diversity using interconnected geo-social networks," in *Proceedings of the 25th International World Wide Web Conference (WWW)*, April 2016.
[62] D. G. Shin and M. S. Jun, "Home iot device certification through speaker recognition," in *17th International Conference on Advanced Communication Technology (ICACT)*, July 2015, pp. 600–603.
[63] Z. Akhtar, C. Micheloni, and G. L. Foresti, "Biometric liveness detection: Challenges and research opportunities," *IEEE Security & Privacy*, vol. 13, no. 5, pp. 63–72, 2015.
[64] A. Bordes, S. Ertekin, J. Weston, and L. Bottou, "Fast kernel classifiers with online and active learning," *Journal of Machine Learning Research*, vol. 6, no. Sep, pp. 1579–1619, 2005.
[65] N. Vaswani, A. R. Chowdhury, and R. Chellappa, "Activity recognition using the dynamics of the configuration of interacting objects," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 2. IEEE, 2003, pp. II–633.