



# Risk-Driven Behavioral Biometric-based One-Shot-cum-Continuous User Authentication Scheme

Attallah Buriro<sup>1</sup> · Sandeep Gupta<sup>1,2</sup> · Artsiom Yautsiukhin<sup>2</sup> · Bruno Crispo<sup>1</sup>

Received: 19 June 2019 / Revised: 27 January 2021 / Accepted: 3 March 2021 / Published online: 7 April 2021  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

The paper presents a risk-driven behavioral biometric-based user authentication scheme for smartphones. Our scheme delivers one-shot-cum-continuous authentication, thus not only authenticates users at the start of the application sign-in process but also, throughout the active user session. The scheme leverages the widely used PIN/password-based authentication technology by giving flexibility to users to enter any random 8-digit alphanumeric text, instead of pre-configured PIN/Passwords. Internally, the scheme exploits two behavioral biometric traits, i.e., touch-timing-differences of the entered strokes and the hand-movement gesture recorded during the *random text* entry, to authenticate users. And, for the entire user session, the scheme *continuously* authenticates the user by computing the risk-score every time the user initiates a sensitive activity. If the risk-score is higher than the predefined threshold, the current user session terminates. Afterward, the scheme requests the user to re-authenticate. Thus, our scheme serves three main objectives: Firstly, it offers users the flexibility to enter an 8 – *digit* random alphanumeric text as their secret enhancing the usability of PIN/password-based schemes. Secondly, it strengthens the security of PIN/password-based schemes as verification decision is not binary, and mimicking the invisible touch-timings and hand-movements simultaneously, could be extremely difficult as our security analysis determined. Lastly, the scheme does not require any dedicated device (e.g., a smart token for OTP generation) for 2-factor authentication. The results obtained on 11,400 user-samples (collected by 3 days *in-the-wild* testing) and user-experience responses (received from the *Software Usability Scale*<sup>4</sup> survey) of 95 testers demonstrate our scheme as an accurate and acceptable user authentication scheme.

**Keywords** Behavioral biometrics · Risk-driven authentication · Human-computer interaction · Smartphones

## 1 Introduction

Smart devices offer a large number of security-sensitive applications, such as mobile banking app, mobile commerce app, on-demand ride-booking app, social networking app, to their users enabling anytime, anywhere access to them. Commonly, these applications have deployed PIN/password-based user authentication schemes to secure access despite numerous security and usability issues present in such schemes [1]. Some of these applications

have deployed 2-factor authentication schemes by introducing one-time-passcodes (OTP), smart-tokens, verification-over-the-call, etc., to address some security issues, however, they too do not deliver a comprehensive risk assessment of the active user session but degrade usability in particular [2].

From the security perspective, PIN/password-based schemes are vulnerable to guessing [3], smudge [4], shoulder-surfing [3, 5], dictionary-based [6] attacks. Similarly, from the usability perspective, users face difficulty to manage numerous PINs/passwords [7] and complex passwords add cognitive load on users [8, 9]. Additionally, it is not easy to employ PIN/password-based schemes for continuous user authentication without affecting the user experience [10]. Further, it is worth mentioning that these schemes do not necessarily authenticate the users, but authorize anyone who enters the correct PIN/password [11]. Thus, it becomes requisite to redesign the PIN/password-based authentication mechanism to overcome their inherent shortcomings.

✉ Sandeep Gupta  
sandeep.gupta@unitn.it

<sup>1</sup> Department of Information Engineering and Computer Science (DISI), University of Trento, Trento, Italy

<sup>2</sup> Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche, Pisa, Italy

In this paper, we propose a risk-driven behavioral biometric-based one-shot-cum-continuous user authentication scheme. Our scheme supplements the existing PIN/password-based authentication schemes with two behavioral biometric traits to enhance their usability and security, i.e., users do not require to remember their PINs, or passwords and authentication decision is not simply a binary comparison. Then, throughout the active user session, the scheme continuously performs risk-assessment to eliminate the dependency on any dedicated devices (e.g., smart token) that are typically, required to generate *One Time Password* (OTP) to finish critical operations.

The proposed system consists of two independent modules, i.e., User Authentication (UA) module and Risk Assessment (RA) module that works in tandem. User Authentication (UA) module creates unique-identification-signature by exploiting the touch-timing-differences, and hand-movement action collected during the course of a 8 – *digit random text* entry by users. The UA module grants access - if both behavioral biometric traits of users match with their stored template. After the successful sign-in, Risk Assessment (RA) module *continuously* tracks client-attributes, such as IMEI number, MAC address, IP address, transaction value, etc., to perform risk assessment throughout the entire user session. The RA module computes the risk-score in terms of the cumulative deviation of client-attributes, every time users initiate a critical operation. If the risk-score is higher than the predefined value, the users' current session is terminated, immediately and UA module prompts for re-authentication.

In brief, our main contributions in this paper are:

- The proposal of a bimodal behavioral biometric-based one-shot-cum-continuous user authentication scheme that authenticates users based on *how* they enter the text instead of *what* they enter, thus strengthen username/password-based schemes.
- The introduction of a novel risk-assessment mechanism that *continuously* determines the need of user re-authentication during the active user session, by computing cumulative deviation of client-attributes.
- The validation of our proposed scheme on a dataset collected *in-the-wild* from 95 testers in three different activities, i.e., *sitting, standing, and walking*.
- The usability evaluation of our scheme by conducting a *System Usability Scale*<sup>1</sup> survey.

**Paper organization** The rest of the paper is organized as the following: Section 2 discusses security, privacy, and usability criteria for an authentication scheme together with classification models and risk-based authentication system

design. Section 3 discusses the threat model, the working of our proposed scheme, and architecture of our system. In Section 4, we discuss the methodology used to design our one-shot-cum-continuous authentication scheme. Section 5 presents the obtained results. Sections 6 and 7 present the security and usability analysis of our proposed system. Section 8 surveys the related approaches proposed over the years for user authentication. Finally, in Section 9, we conclude the paper with a summary of the work and the possible future dimensions.

## 2 Background

In this section, we briefly discuss security, privacy, and usability criteria for designing an authentication scheme. Afterward, an overview to classification model design and risk-based authentication system is presented.

### 2.1 Security, Privacy, and Usability Criteria

While designing a user authentication scheme, the attributes - *security, privacy, and usability* emerged out to be orthogonal to each other [12]. Studies have shown that none of the available authentication schemes can satisfy these three attributes, simultaneously [13]. For instance, PIN/password or smart-token-based schemes do not affect users' privacy, but they have several security and usability issues. Whereas, biometric-based schemes can fulfill security and usability criteria, but affect the privacy of a user.

The foremost security criterion in designing an authentication scheme is the selection of modalities that can reliably verify users. Biometrics naturally fits this purpose as it can establish the identity of individuals based on their distinctive physical or behavioral traits [14]. Moreover, biometric authentication systems are not binary in execution like conventional authentication systems [15]. Biometric-based authentication schemes are more resilient to attacks that are essential for security-sensitive systems [16]. Behavioral biometric-based schemes can function without active input, user cooperation, or knowledge that the underlying authentication engine is acquiring unlike conventional systems [17].

Recent privacy regulation laws like General Data Protection Regulation (GDPR) [18] or California Consumer Privacy Act (CCPA) [19] set guidelines for using and storing biometric data. GDPR enlisted biometric data that allow or confirm the unique identification of an individual as a special category of personal data under Art. 9 [20]. Therefore, conformance for users' privacy to be addressed by incorporating adequate measures (e.g., template protection and template storage location) [21].

<sup>1</sup><https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>

We also recommend hardware-level encryption on client devices to establish trust between users and businesses as a part of the privacy-first approach to employ behavioral biometric-based schemes for users' authentication.

Biometric-based authentication schemes have shown higher usability acceptance in contrast to knowledge- and token-based authentication schemes, thus, meeting usability criteria [22]. Further, they do not add cognitive load on users unlike conventional authentication schemes [23]. Usability evaluation can determine user experiences, challenges, skills, and attitudes in designing a particular authentication scheme.

Many security-sensitive sectors, e.g., banking and finance, transport, smart-offices, etc., enforce user authentication to maintaining and safeguarding themselves from adversaries. And, at the same time, consolidate the security, privacy, and usability criteria of their systems. Thus, a trade-off between security, privacy, and usability is a viable option for designing biometric-based authentication schemes. Eventually, usable security can overcome the inadvertent (or even deliberate) undermining of security by end-users.

## 2.2 Classification Model Design

The machine-learning (ML) enables to derive a precise mapping function to design appropriate classification models as per the use cases [12]. The classification model can learn from the labeled dataset (training data) to predict labels of new data, which is termed as supervised learning. The training process continues until the model achieves the desired level of accuracy on the training data set. Mathematically, the classification process can be defined using Eq. 1.

$$R_A = \begin{cases} M & \text{if } CM(C_I, C_T) \text{ is } \geq T \\ NM & \text{if } CM(C_I, C_T) \text{ is } < T \end{cases} \quad (1)$$

Where,  $CM$  is the classification model that receives the claimant's input ( $C_I$ ) and claimant's templates ( $C_T$ ) to measure the similarity between them for a given threshold ( $T$ ). The authentication result ( $R_A$ ) is set to `Match` ( $M$ ) or `Not Match` ( $NM$ ) according to the score obtained from classification model [24].

Typically, classification models can be divided into multi-class classification and one-class classification to address various user authentication scenarios. Multi-class classification models are best suited for multiple user authentication scenarios. Smart applications (apps), such as online banking, online-shopping, ride-booking, are used by multiple users. Therefore, multi-class classifiers can be exploited to classify more than two classes (users). However, it is expected that classes must be mutually

exclusive to each other and each new instance belongs to one of the classes. Whereas, the one-class classification model is suitable for scenarios like user authentication for accessing single-handled devices. The main purpose of a one-class classification is to detect an anomaly or a state other than the one characteristically shown by the target class (legitimate user). Therefore, information regarding other classes (illegitimate users) is not required while training a one-class classification model. This model is often called outlier (or novelty) detection.

We recommend following guidelines prescribed in [25] to implement and evaluate biometric-based authentication schemes under design for high-quality outcomes and wider user acceptance.

## 2.3 Risk-based Authentication

Risk-based authentication utilizes the concept of explicit, implicit and continuous authentication mechanisms to make user verification process frictionless as much as possible. Such that, the system can tailor the authentication schemes according to the risk calculated at runtime, to maximize the security and usability requirements [26].

Generally, the existing risk-based authentication system uses a risk-score to estimate the risk associated with user's activities including the sign-in attempt, in a user session [27]. A user-session can be characterized by using historical and contextual attributes such as transactions pattern, user's geographic location, access-time, IMEI number, MAC and IP address of registered devices, user's typing speed and so on, collectively can be defined as *client-attributes*.

Risk-based user authentication can be applied from two different perspectives, i.e., proactive or reactive. When applied proactively, risk-based authentication can be implemented as continuous authentication spanning across the entire user session. The goal is to anticipate the genesis of potential attacks, failures, or any kind of security issues, and to enforce the appropriate response plans. In contrast, reactive risk-based authentication complementary to the proactive risk-based authentication, in which, some risks might be accepted without being eliminated. However, if any risk is detected during the session the re-authentication may be imposed.

## 3 Risk-driven Bimodal Behavioral Biometric-based User Authentication Scheme

This section presents the assumed threat model. Followed by, the working of our one-shot-cum-continuous authentication system and its system architecture.

### 3.1 Threat Model

We considered physical attacks, where (i) the adversary accidentally finds an unlocked smartphone, (ii) the adversary is a friend or colleague (who possibly knowing user’s PIN/Passwords), and (iii) the adversary records users while they interact with their smartphones. Eventually, the adversary exploits the weaknesses of PIN/password-based authentication schemes to gain access to sensitive resources (data and applications) residing on users’ smartphones.

Prior studies [10, 28] also indicated that the above-discussed scenarios are quite apparent, as users use their smartphones at commons places like offices, homes, meeting rooms, or streets, which may give opportunities to adversaries to target their smartphones, easily. As a consequence, smartphone users can be a victim of monetary frauds, identity thefts, or similar unfavorable incidents.

### 3.2 How Our Scheme Works?

Figure 1 illustrates the model of our one-shot-cum-continuous authentication scheme explaining how it addresses security and usability issues in existing user/password-based, and 2-factor authentication schemes.

The scheme enables users to enter any random 8 – digit alphanumeric text to access the application to enhance the usability of existing PIN/Password-based one-shot authentication schemes. Further, the scheme verifies the users’ identity based on timing differences between the entered keystrokes and their hand-movement in 3 dimensional space instead of just a binary comparison, to enhances security.

After the successful sign-in, the scheme *continuously* monitors client-attributes and computes the risk-score at the instant users initiate critical activities. Based on the risk score, it permits users to perform that activity, otherwise, scheme prompts for re-authentication. Thus, our

scheme is capable of detecting any anomalies in the users’ usage pattern throughout the life-cycle of a typical user session and apparently, 2-factor authentication can be safely disregarded.

### 3.3 System Architecture

The system adopts a client-server architecture [2] as shown in Fig. 2. The client consists of a data acquisition (DA) modules that can be added to existing smartphone applications, seamlessly. The DA collects the two behavioral biometric traits along with client-attributes and transfers the encrypted data to the server at runtime for further processing.

The server includes two independent modules, i.e., the User Authentication (UA) and the Risk Assessment (RA) module. The UA module performs user authentication based on features extracted from touch-typing and hand-movements behavioral traits, as explained in Section 4.2. The RA module, using the Runtime-Risk-Assessor (RRA) inside the Risk Engine (RE), computes the risk score at run time, as explained in Section 4.6, each time a critical operation is performed. The RE then, notifies the Session Manager (SM) if the computed risk score is higher than the predefined threshold. Afterward, the SM sends the command to the UA module for re-authentication.

## 4 Methodology

In this section, we explain the steps taken to design and validate the proposed authentication scheme.

### 4.1 Data Collection

We develop a prototype application (app) that can be installed on any Android devices having OS version 4.4.x or higher. To conduct our experiment, we collaborated with

Figure 1 Our one-shot-cum-continuous authentication scheme model.

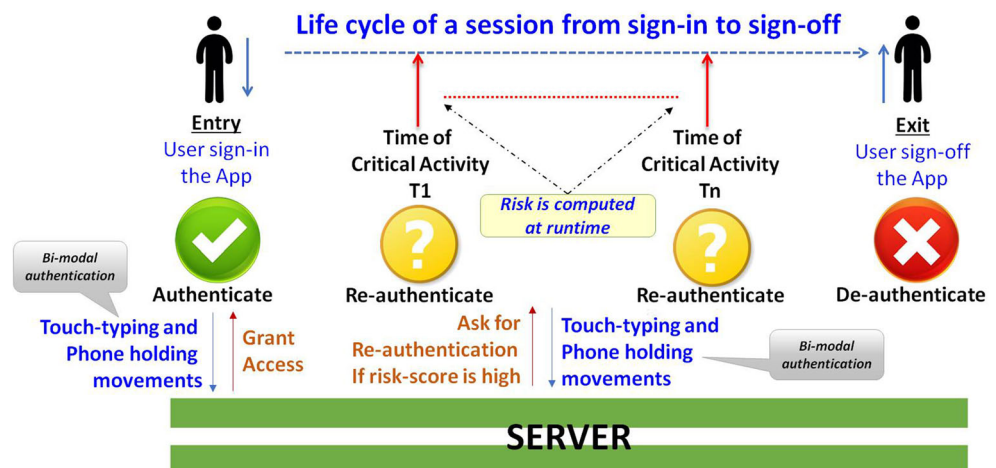
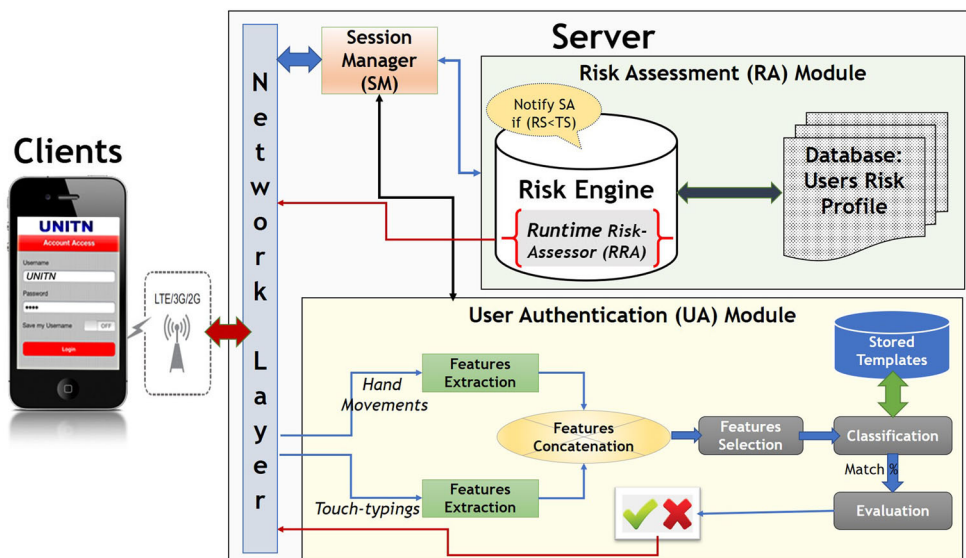


Figure 2 System architecture.



UBERTESTERS<sup>2</sup> - a crowdsourcing software testing platform. Testers were certified quality assurance engineers or experienced software developers and they were rewarded on an hourly basis. Complete instructions to use our prototype application, the installation/uninstallation procedure and the user consent were provided to testers. Each tester signed the consent form before they download and install our application.

The app enables testers to perform the experiment for approximately, an hour that spans over 3 days with 1 session per day, i.e., 3 sessions in 3 days. During each training session, testers can interact with the app for 15 minutes in 3 different activities, i.e., *sitting, standing, and walking*. On the third day, the testers can also test the app with 30 testing samples in any activity of their choice. Afterward, the testers performed the SUS survey, and they filled their demographic information presented in Appendix A.

We recruited 100 testers conduct the experiment. Each tester tested our prototype application on their own smartphones under the real-life conditions. However, we discard the data from 5 testers for reasons like their smartphones did not support the required sensors or Internet connectivity was too slow to transfer the data in real-time to our server. Table 1 summarizes the demographics of testers selected to participate in our experiment.

Overall, we collected 11,400 samples with 120 samples from each tester (30 samples in each of the 3 different training activity and 30 samples during testing) and received 95 SUS responses in this experiment. Thus, we evaluated our scheme on a collected dataset of 95 users having a total of 11,400 samples.

### 4.2 Feature Extraction

We used the touchscreen sensor and seven 3-dimensional motion sensors (i.e., the accelerometer, the high-pass sensor, the low-pass sensor, the orientation sensor, the gravity sensor, the gyroscope, and the magnetometer) to collect raw data for touch-stroke and hand-movement, respectively [29]. The high-pass and low-pass sensory data is computed mathematically, by applying High-Pass (HP) and Low-Pass (LP) filters as shown in Eqs. 2 and 3.

$$Value_{HP} = Value_{Gravity} \times \alpha + Value_{Accelerometer} \times (1 - \alpha) \tag{2}$$

$$Value_{LP} = Value_{Accelerometer} - Value_{Gravity} \tag{3}$$

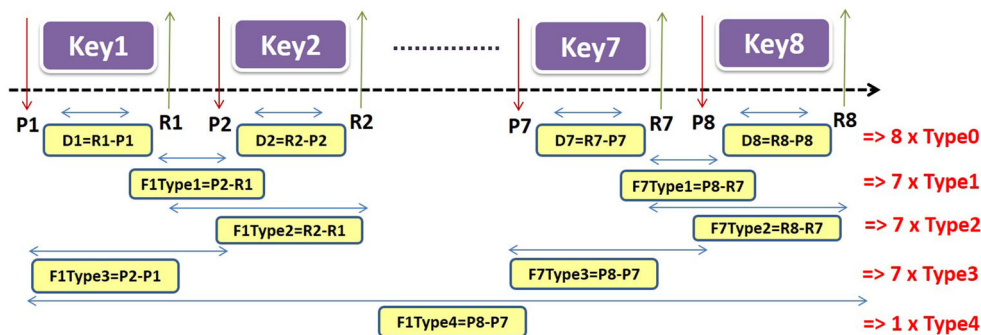
Where,  $Value_{HP}$ ,  $Value_{LP}$ ,  $Value_{Accelerometer}$ , and  $Value_{Gravity}$  represent the value of the high-pass, low-pass,

Table 1 User demographics (M = Male, F = Female, R = Right, L = Left).

Parameter	Description
No. of Users	95
Sample Size	<b>Sitting - 2,850</b> (95 × 30) <b>Standing - 2,850</b> (95 × 30) <b>Walking - 2,850</b> (95 × 30) <b>Testing - 2,850</b> (95 × 30)
Devices	Android Smartphones having OS 4.4.x version or above
No. of Sessions	3
Password	8-digit <i>free-text</i>
Gender	75(m), 20(f)
Handedness	89(R), 6(L)
Age Groups	90 (20 – 40), 5 (41 – 60)

<sup>2</sup><https://ubertesters.com/>

**Figure 3** Touch-typing features for 8-keys entry [30].



accelerometer, and gravity sensor, respectively at a time  $t$ . We set  $\alpha$  to 0.1 that was determined, empirically.

As shown in Fig. 3, touch-typing features consist of 8 *Type0* (timing difference between each key release and key press), 7 *Type1* (timing difference a key press and previous key release), 7 *Type2* (timing difference two successive keys release), 7 *Type3* (timing difference two successive keys press), and 1 *Type4* (timing difference between last and first key press). Thus, we extracted 30 touch-typing features from the 8-digit *random-text* entry.

Similarly, a user’s hand-movement is modelled in terms of 3-D data streams, i.e., X, Y and Z, from each motion sensor. In addition, we computed the 4<sup>th</sup> dimension, Magnitude (M), by using Eq. 4.

$$Value_M = \sqrt{(Value_x^2 + Value_y^2 + Value_z^2)} \tag{4}$$

Where,  $Value_M$  is the Magnitude and  $Value_x$ ,  $Value_y$  and  $Value_z$  are the values of X, Y and Z value of a sensor, at a time  $t$ .

We obtained 4 data streams from each of the seven motion sensors with the delay set at *SENSOR\_DELAY\_GAME* [29]. Then, from each data stream, we extracted 4 statistical features, namely Mean ( $\mu$ ), Standard Deviation ( $\sigma$ ), Skewness ( $s$ ), and Kurtosis ( $k$ ), that gives 16 statistical features per sensor as shown in Table 2.

Finally, we concatenate 30 touch-stroke features and 112 hand-movements features to create a feature vector of size 142. Here, we prefer to choose the feature level fusion over the sensor level fusion because sensory data could have inconsistent and/or unusable data that may affect classifiers accuracy [31].

**Table 2** Statistical features per sensor for a hand-movement behavior.

No.	Hand-movement Features			
1-4	$\mu_x$	$\mu_y$	$\mu_z$	$\mu_M$
5-8	$\sigma_x$	$\sigma_y$	$\sigma_z$	$\sigma_M$
9-12	$s_x$	$s_y$	$s_z$	$s_M$
13-16	$k_x$	$k_y$	$k_z$	$k_M$

### 4.3 Feature Selection

The primary purpose of any feature selection scheme is to filter out the redundant and less productive features to determine the most productive features [15]. This improves the performance of a classifier as processing smaller feature vectors would be computationally faster. We applied Information Gain Attitude Evaluator (IGAE) for feature selection. This scheme evaluates the worth of a feature by computing its information gain with respect to the class. We obtained the threshold for feature selection by dividing the number of users (95) by the total number of features (142). And, top 66% of 142 feature are selected for designing classification model as illustrated in Fig. 4.

### 4.4 Classifier Selection

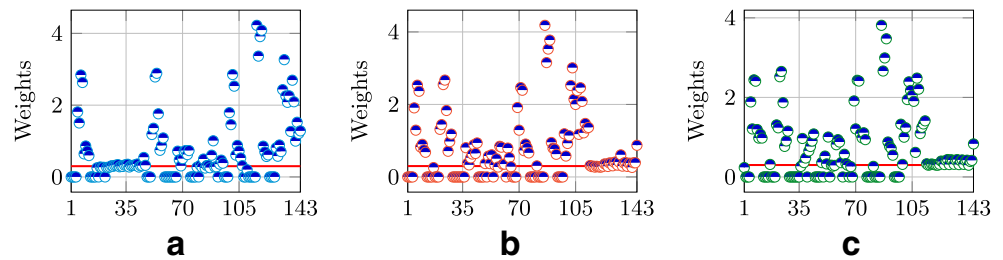
The classifier selection depends on various parameters, such as data size, data characteristics and training time, etc. We selected simple, yet effective state-of-the-art classifiers: Naive Bayes (NB), NeuralNet (NN), and Random Forest (RF) classifiers.

Bayesian classifiers, such as Belief Networks and Naive Bayes employ the probabilistic technique for the classification tasks. The Naive Bayes method starts with a strong but “naïve” assumption that the features are independent of each other. It works perfectly well if this condition holds true. Furthermore, it is widely used because of its super simplicity, faster learning capability, elegance, and robustness [32].

NN classifier belongs to the Artificial Neural Network (ANN) family. These models represent many interconnected network elements designed essentially to classify different patterns. These models have been shown to be quicker and accurate [33]. We used the Levenberg-Marquardt trained feed-forward neural network as the classifier in our analysis.

RF has been considered as an accurate and efficient classifier in recent years [34]. The reasons for their popularity include: (i) its accuracy among the current algorithms even without any optimization, (ii) it generally does not overfit, (iii) it efficiently handles the missing data, and (iv) its effectiveness on small as well as for

**Figure 4** Features vs. Weights plots using Information Gain Attribute Evaluation (IGAE) method.



large datasets, etc. We preferred this classifier because of its effectiveness in the previous studies [35]. RF classifier works on the principle of growing many classification trees and to classify, it puts the query sample down to each of the trees in the forest. Each tree classifies that sample and “vote” for a particular class. The final decision chosen by the forest is based on the higher number of votes (over all the trees in the forest).

### 4.5 Classifier Training & Testing

We consider remote-user-authentication to access security-sensitive applications on smartphones as a multiclass classification problem. We construct a classification model and validated users in two scenarios, (i) a verifying legitimate user scenario, and (ii) a zero-effort attack scenario. We evaluate the classification model by partitioning the dataset into training and testing set. We trained selected classifiers with 5, 10 and 15 samples and used the remaining samples for testing.

### 4.6 Risk Assessment Model

According to ISO 9000:2015 [36], risk is the “effect of uncertainty on objectives” and an effect can be a positive or negative deviation from what is expected. An objective can be strategic, tactical, or operational. Generally, the existing risk-driven authentication system uses a risk-score to estimate the risk associated with the user’s activities including the sign-in attempt, in a typical user session [27]. A user-session can be characterized by using historical and contextual attributes, such as transactions pattern, user’s geographic location, access-time, IMEI number, MAC and IP address of registered devices, the user’s typing speed and so on, collectively can be defined as the *client-attributes*.

The *risk-score* can be computed by determining cumulative *uncertainty* (degree of deviation) associated with each client-attribute. By using a mathematical formula or expression, the degree of deviation can be easily determined to establish a relationship between the present value, and previously recorded values (where the *objectives* achieved successfully) of client-attributes.

In our system, the Risk Engine (RE) configures a client profile of each customer by using contextual and historical data, e.g., transactions patterns, location, access-time, IMEI

number, MAC and IP address of registered devices, operating system, applications installed, and stylometry, etc., as client-attributes.

To create the user’s client profile, *RE* initially assigns a unique weight (natural value) to each client-attribute as per the user’s preferences.

$$CA_i = VALUE \begin{cases} \forall i \in M \\ VALUE \geq 1 \end{cases} \quad (5)$$

Equation 5 describes the weight assignment process to each of the *M* client-attributes. *RE* assigns a higher value to the client-attribute based on the user preference order. For example, if a user has given more importance to *Smartphone IMEI* over *access time* than will be  $CA_{IMEI} > CA_{AccessTime}$ . Two client-attributes can have a common integer value. However, the model can reassign the weights by analyzing the user’s usage pattern, thus, updates the client-profile, automatically.

Table 3 presents the structure of a user’s client-profile. Each row comprises of a client-attribute, its weight, and values of the current session, i.e., *Session<sub>N</sub>* to all the  $N - 1^{th}$  previous sessions. Frequency of Non-occurrence ( $FNO_i$ ) and Impact of Non-occurrence ( $INO_i$ ).

To obtain Frequency of Non-occurrence ( $FNO_i$ ) and Impact of Non-occurrence ( $INO_i$ ), we first calculate Frequency of Occurrence ( $FO_i$ ) as follows:

The Frequency of Occurrence ( $FO_i$ ) is an estimate of how often the current client-attribute value ( $Value_{iN}$ ) has occurred in previous  $N - 1$  sessions [37], which is determined using Eq. 6.

$$O_i = \sum_{j=1}^{N-1} [Value_{iN} = Value_{ij}] \forall i \in M - 1, \quad \text{and}$$

$$FO_i = \frac{O_i}{N - 1} \quad \forall i \in M - 1 \quad (6)$$

Where,  $O_i$  is the occurrence of  $Value_{iN}$  of a  $i^{th}$  client-attribute. The value of  $FO_i$  towards  $\approx 1$  indicates lower risk, whereas towards  $\approx 0$  indicates higher risk.

Subsequently, Frequency of Non-occurrence ( $FNO_i$ ) and Impact of Non-occurrence ( $INO_i$ ) are measured at runtime using Eqs. 7 and 8, respectively.

$$FNO_i = 1 - FO_i \quad \forall i \in M \quad (7)$$

**Table 3** Structure of user’s client profile.

#	Client-Attributes	Weight of Client-Attributes	Session <sub>N</sub>	Session <sub>N-1</sub>	...	Session <sub>2</sub>	Session <sub>1</sub>	Frequency of Non-occurrence	Impact of Non-occurrence
1	TRANSACTION PATTERN	$CA_1$	$Value_{1N}$	$Value_{1(N-1)}$	...	$Value_{12}$	$Value_{11}$	$FNO_1$	$INO_1$
2	LOCATION	$CA_2$	$Value_{2N}$	$Value_{2(N-1)}$	...	$Value_{22}$	$Value_{21}$	$FNO_2$	$INO_2$
3	ACCESS TIME	$CA_3$	$Value_{3N}$	$Value_{3(N-1)}$	...	$Value_{32}$	$Value_{31}$	$FNO_3$	$INO_3$
4	IMEI NUMBER	$CA_4$	$Value_{4N}$	$Value_{4(N-1)}$	...	$Value_{42}$	$Value_{41}$	$FNO_4$	$INO_4$
5	MAC ADDRESS	$CA_5$	$Value_{5N}$	$Value_{5(N-1)}$	...	$Value_{52}$	$Value_{51}$	$FNO_5$	$INO_5$
6	IP ADDRESS	$CA_6$	$Value_{6N}$	$Value_{6(N-1)}$	...	$Value_{62}$	$Value_{61}$	$FNO_6$	$INO_6$
7	OS VERSION	$CA_7$	$Value_{7N}$	$Value_{7(N-1)}$	...	$Value_{72}$	$Value_{71}$	$FNO_7$	$INO_7$
8	APPS INSTALLED	$CA_8$	$Value_{8N}$	$Value_{8(N-1)}$	...	$Value_{82}$	$Value_{81}$	$FNO_8$	$INO_8$
9	TOUCH-TYPING SPEED	$CA_9$	$Value_{9N}$	$Value_{9(N-1)}$	...	$Value_{92}$	$Value_{91}$	$FNO_9$	$INO_9$
...	...	...	...	...	...	...	...	...	...
M	STYLOMETRY	$CA_M$	$Value_{MN}$	$Value_{M(N-1)}$	...	$Value_{M2}$	$Value_{M1}$	$FNO_M$	$INO_M$

$$INO_i = FNO_i \times CA_i \quad \forall i \in M \tag{8}$$

Where,  $FNO_i$  is defined as the frequency of occurrence, which can be calculated using Eq. 6,  $CA_i$  is the weight of each client-attribute and  $M$  is the number of client-attributes. The value of  $FNO_i$  towards  $\approx 0$  indicates lower risk, whereas towards  $\approx 1$  indicates higher risk.

For example, a customer has accessed her banking app from  $X$  location  $\pm 10KM$  in the previous 10 sessions. But, in the current session, the access location is found to be  $Y$  so the frequency of its occurrence ( $FNO_{location} = \frac{0}{10}$ ) becomes 0. Therefore, the frequency of its non-occurrence ( $FNO_{location}$ ) becomes 1, which is calculated using Eq. 7. As described in Eq. 8, multiply  $FNO_{location}$  with  $CA_{location}$  to calculate  $INO_{location}$ , which gives a positive number. Similarly, the impact of non-occurrence of other client-attributes can be calculated.

Finally, the risk-score is computed using Eq. 9, which can be defined as the sum of all the impact-of-non-occurrence of each client-attribute. Higher the number means higher the risk.

$$Risk\ Score = \sum_{i=1}^M INO_i \tag{9}$$

Where,  $M$  is number of client-attributes.

The risk score is computed and matched with the threshold before any of the critical operations is performed. If the risk-score is higher than the predefined value (e.g., average of the risk-scores in previous  $N - 1$  sessions), re-authentication is exercised leveraging the proposed behavioral biometric-based bimodal authentication scheme.

Thus, our authentication scheme utilizes the concept of one-shot and continuous authentication mechanisms driven by risk assessment, as explained in Section 3.2, offering a user friendly verification mechanism.

## 5 Experimentation Results

### 5.1 Success Metric

We report our achieved results using the following metrics:

- **True Acceptance Rate (TAR):** The rate of correctly accepted attempts of the valid user.
- **False Rejection Rate (FRR):** The rate of falsely rejected attempts of the valid user. It can be estimated by computing  $1 - TAR$ .
- **False Acceptance Rate (FAR):** The rate of falsely accepted attempts of an adversary.
- **True Rejection Rate (TRR):** The rate of correctly rejected attempts of an adversary. It can be estimated by computing  $1 - FAR$ .
- **Receiver Operating Characteristics (ROC):** ROC is the graphical representation of classifier performance.



The curve is typically plotted between TAR on the y-axis and False Acceptance Rate (FAR) on the x-axis. The curve starts from (0,0) and ends at (1,1) coordinates. The curve closer to (0,1) shows the better performance.

### 5.2 Authentication Results

We report the results of all of our chosen classifiers in terms of TAR and FAR, on full features, in Table 4. TAR of all the chosen classifiers increases with the increase in the number of training patterns (see Table 4), i.e., for NB classifier TAR increased from 72.72% (on 5 training samples) to 87.58% (on 15 training samples) in *sitting* activity. NN classifier did not work well possibly because of the limited number of training samples as it generally requires more training samples. RF classifier performed consistently well across all the activities and for the different number of samples. We achieved a TAR of 80.51% (in *sitting*), 82.91% (in *standing*), and 81.38% (in *walking*), on just 5 training samples, and this TAR increased up to 91.79%, 91.58%, and 86.95%, on 15 training samples. The highest achieved TAR by RF is 91.79% (at just 0.04% FAR), on 15 training samples.

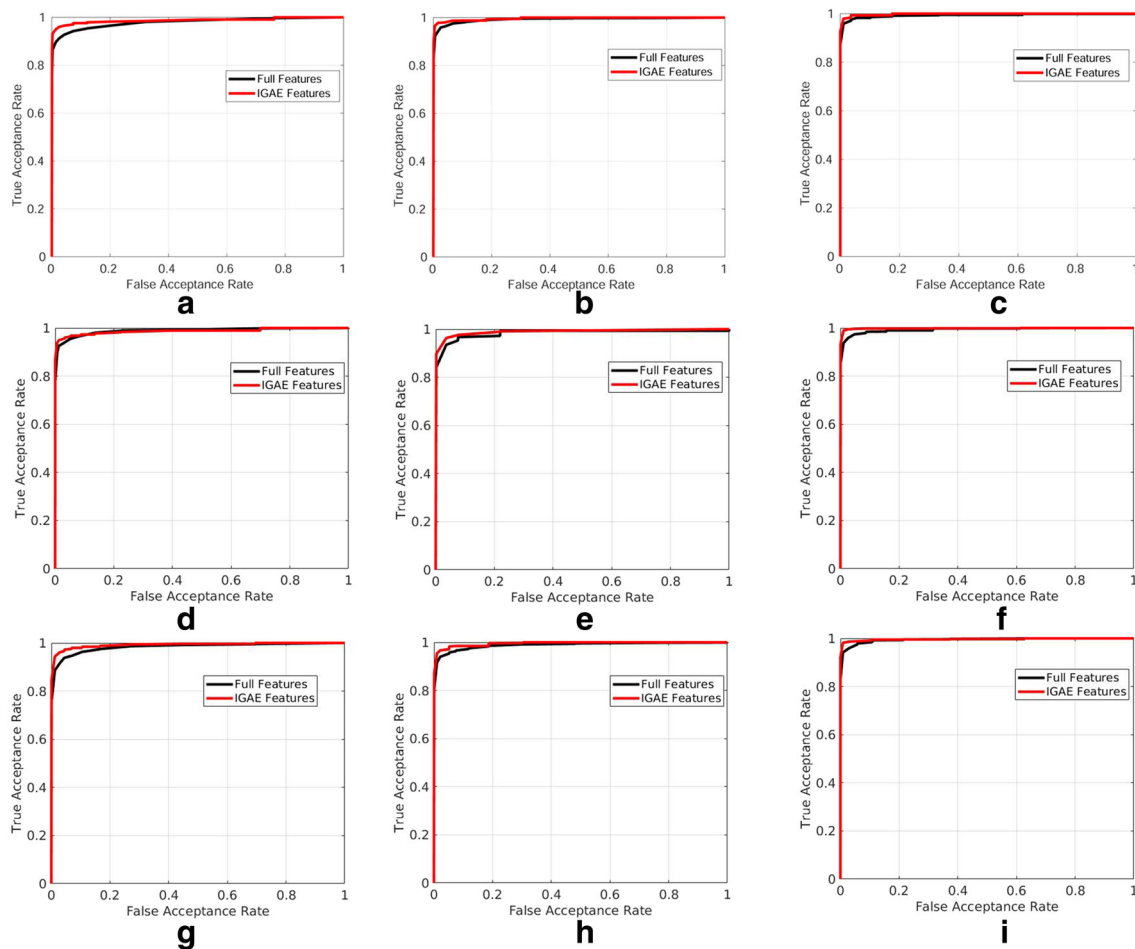
Afterward, we present the results of all the classifiers on IGAE selected features (see Table 5). The results of all the classifiers improved, significantly, over the extracted IGAE features except for NB in *standing* and *walking* activities, over 5 training samples. NN performed comparatively well on the smaller feature vectors. RF classifier improved the authentication results on IGAE features, i.e., from 88.04% to 89.10%, 92.88% to 95.18% and 94.87% to 96.00% for three activities, on 5, 10, and 15 training samples, respectively. It is evident that our scheme is very robust against the zero-effort attacks, i.e., TRR is much higher and FAR is very low.

We also plot the results of RF classifier in terms of ROC curves (see Fig. 5). We show an average ROC of all the users obtained through Vertical Averaging (VA) [38]. In this averaging, the averages of the TAR rates are plotted against the researcher-defined fixed FAR. Due to the space limitations, we illustrate ROC curves for best performing classifier, i.e., for RF, for all the activities and all the training sample scenarios. Figure 5 reflects RF classifier as very productive and accurate classifier throughout.

RF classifiers outperformed both NB and NN classifiers because of its ability to reduce the variances and its most unlikeliness to over-fitting. NB classifier requires Gaussian distributed data, which might not be true in the dataset, hence it failed to address the problem of concept-drift. The NN classifier failed because of the limited number of training samples. It generally requires more training samples to learn well.

**Table 4** Results of different classifiers (averaged over all 95 users) on full features in (%)

Activity	5						10						15					
	sitting		standing		walking		sitting		standing		walking		sitting		standing		walking	
	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR
NB	72.72	0.24	79.53	0.18	76.04	0.21	83.66	0.12	84.51	0.11	80.55	0.14	87.58	0.07	88.35	0.06	83.72	0.08
NN	57.81	0.37	56.88	0.38	55.03	0.47	63.61	0.27	68.07	0.23	25.27	0.82	70.53	0.16	73.47	0.13	27.24	0.76
<b>RF</b>	<b>80.51</b>	<b>0.17</b>	<b>82.91</b>	<b>0.17</b>	<b>81.38</b>	<b>0.19</b>	<b>87.87</b>	<b>0.09</b>	<b>88.72</b>	<b>0.12</b>	<b>85.31</b>	<b>0.16</b>	<b>91.79</b>	<b>0.04</b>	<b>91.58</b>	<b>0.08</b>	<b>86.95</b>	<b>0.13</b>



**Figure 5** The ROC curves of RF classifier on full and IGAE features for different activities, i.e., (i) Sitting (a - c), (ii) Standing (d - f), and (iii) Walking (g - i).

### 5.3 Discussion on Results

A comparison between Tables 4 and 5 exhibits that classifiers performance enhanced on selected features. Thus, the feature selection process improves the system's accuracy and improve the performance (as the processing of smaller feature vector may incur less computational cost and processing, making decision time comparatively shorter). Overall, the average training (up to 15 samples) and testing time for our proposed scheme is  $\approx 3.0$  to 5.5 seconds and  $\approx 200$  to 300 milliseconds, respectively. However, these timings data are required to be benchmark in more standard lab setting environment.

Further, Appendix C shows the distribution of TAR (per user) for *sitting*, *standing*, *walking* activities, obtained on just 5 training samples for RF classifier performed since it performs better with both the full and IGAE features in all the activities. Due to space limitations, we show such comparison for 5-samples training scenario, only. It is worth reminding that, in this scenario, the classifier was trained on the first 5 samples and tested with the remaining 25 samples

of the same user to obtain TAR and the process was repeated for each user. It is evident from Fig. 7 that the TAR for most of the users increased on IGAE features, i.e., all 25 samples of 41 users were correctly accepted compared to just 13, on full features (see Fig. 8a) in *sitting* activity. Similarly, for *standing* and *walking* activities, 44, and 38 users were correctly accepted (with 100% TAR), compared to 25 and 23, respectively (see Figs. 8b and 8c).

## 6 Security Analysis

We performed additional experiments to replicate random attack and mimic attack scenarios for evaluating the robustness of our proposed system.

### 6.1 Mimic Attack

We recruited 8 testers to carry out the mimic attack. Each one of the 8 testers trained the prototype application installed on a smartphone, which is closely observed by the remaining 7

**Table 5** Results of different classifiers (averaged over all 95 users) on IGAE features in (%)

Training Samples	5						10						15					
	sitting		standing		walking		sitting		standing		walking		sitting		standing		walking	
	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR
Activity																		
Classifiers	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR
NB	79.17	0.22	78.77	0.22	73.56	0.28	85.11	0.16	86.86	0.16	76.99	0.25	88.88	0.14	87.72	0.13	80.14	0.21
NN	77.26	0.23	77.81	0.23	73.64	0.28	84.51	0.17	80.80	0.21	76.49	0.26	85.89	0.14	84.35	0.16	80.77	0.20
<b>RF</b>	<b>89.10</b>	<b>0.10</b>	<b>89.26</b>	<b>0.09</b>	<b>88.04</b>	<b>0.10</b>	<b>95.18</b>	<b>0.04</b>	<b>93.64</b>	<b>0.04</b>	<b>92.88</b>	<b>0.05</b>	<b>96.00</b>	<b>0.01</b>	<b>95.92</b>	<b>0.02</b>	<b>94.87</b>	<b>0.02</b>

testers to learn the holding and typing patterns. In each tester’s training session 30 observations (10 per 3 postures, i.e., sitting, standing, and walking) are collected. Then, the remaining 7 testers tried 10 times to carry out the mimic attack.

A multi-class classification model for 8 testers is generated by using RF classifier with 30 training samples per class, i.e., a total of 240 training samples. Subsequently, we tested this classification model with 8 sets of 70 mimic attack samples collected from the remaining 7 testers (10 samples per tester), labeling each set from 1 to 8.

Table 6 presents the result for each class in terms of True Acceptance (*TA*) and False Rejection (*FR*). Higher the *FR* better the robustness of the system. Thus, the overall robustness of the proposed system against mimic attack is 98.75%.

### 6.2 Random Attack

To carry out the random attack, we asked each of the 8 testers to test the application 10 times in any of the 3 postures. Then, we tested the classification model robustness with 70 random attack samples (excluding the samples of a legitimate user) 8 times by assigning labels from 1 to 8.

Table 7 presents the random attack results for each class in terms of *TA* and *FR*. None of the 240 random attack attempts were successful. Thus, the overall robustness of the proposed system against random attack is 100%.

## 7 Usability Analysis

Secure yet usable user authentication mechanism is a prerequisite to balance between security and usability goals. This section presents a detailed usability evaluation of our proposed scheme.

### 7.1 Methodology

System Usability Scale (SUS) [39] is considered as a standard tool to record user experience related to the usability of a system and has been extensively used in

**Table 6** Mimic attack results.

Class	True Acceptance	False Rejection	Robustness(%)
1	0	70	100
2	0	70	100
3	0	70	100
4	1	69	98.57
5	0	70	100
6	0	70	100
7	2	68	97.14
8	0	70	100
3		237	<b>98.75</b>

**Table 7** Random attack results.

Set	True Acceptance	False Rejection	Robustness(%)
1	0	70	100
2	0	70	100
3	0	70	100
4	0	70	100
5	0	70	100
6	0	70	100
7	0	70	100
8	0	70	100
	0	240	<b>100</b>

the context of smartphone user authentication [40, 41]. The user’s response to each question is recorded on a 5-point scale ranging from “Strongly Disagree” to “Strongly Agree”. The output is computed as a score between 0 - 100. The higher the score more usable the system.

We replaced the word “system” with “mechanism” in the SUS questionnaire as done in the previous studies [40, 42]. We added an open, subjective but optional question (“Do you have any feedback you like to share with us?”), as question 11, to get the participant’s feedback on our scheme.

### 7.2 Responses

Figure 6 illustrate the SUS questionnaire and the collected responses from all the 95 participants.

Overall, our scheme achieves the SUS score of  $\approx 73$  which is significantly above the standard average score of 68 [43]. As per the recorded feedback, the majority of the users looked satisfied describing our proposed scheme as a simple, extremely convenient, user-friendly and intuitive. In response to question 3, i.e., “I thought Touch-type mechanism was easy to use”, 80 users ( $\approx 81\%$ ) agreed or

strongly agreed with the point that our scheme is easy to use in contrast to just 6 ( $\approx 5\%$ ) who disagreed or strongly disagreed. Similarly in response to Question 10, i.e., “I needed to learn a lot of things before I could get going with Touch-type mechanism” 74 users ( $\approx 75\%$ ) were disagreed or strongly disagreed in contrast to just 8 (9%) who agreed or strongly agreed to consider our scheme as difficult and would require to learn the scheme.

We also received some negative responses related mainly to the number of digits (8) and the number of training samples. Most of the testers suggested using less number of samples, i.e., 5 (46.5%), 10 (22.7%) as setting up a PIN or registering the face requires less training. We are agreed to the suggestion of less number of samples and also to reduce the number of digits. The same scheme, if reduced to 4, could be used for smartphone unlocking. However, reducing the number of digits is not viable in social networking and mobile banking scenarios, as their existing app require 8-digit fixed alphanumeric passcode.

Overall most of the testers seem comfortable and confident about our scheme mainly because of the flexibility of typing any combination of 8-digit text. Experimental results confirm our scheme as usable, practical and would be widely acceptable.

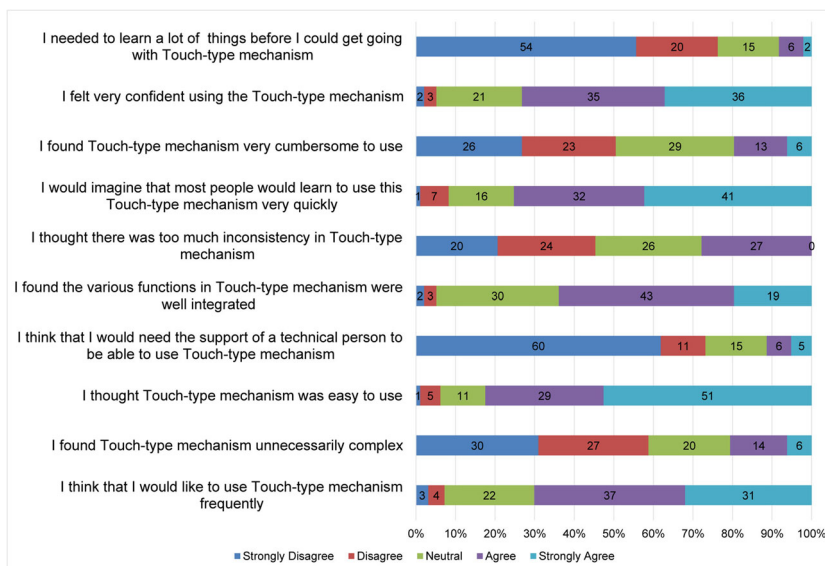
### 8 Related Work

In this section, we present the most relevant schemes proposed over the years.

#### 8.1 Behavioral Biometric-based User Authentication

Behavioral biometrics offers a simple way to implement a frictionless user authentication schemes, which are suitable

**Figure 6** SUS questionnaire and Users responses.



for continuous authentication. This is possible due to the advantages associated with behavioral biometrics: 1) transparent collection, 2) no special hardware requirements, and 3) cost effective deployment [44].

Behavioral data, such as gait, grip, swipe, pick-up, touch, and voice can be collected, unobtrusively, due to the availability of sensors, particularly accelerometer, gyroscope, magnetometer, proximity sensor, soft keyboards, touch screens and microphone in smartphones and have become widely researched subject these days.

In this section, we survey various behavioral authentication schemes proposed for user authentication over the years. Our emphasis will be on the (i) novel behaviors, (ii) the work which uses smartphone sensory data and/or (iii) which require minimal user effort.

### 8.1.1 Keystroke/Touch based authentication:

The concept of augmenting keystroke/touch-based behavioral biometrics to PIN or password is predicated on the understanding that users need a better way to prove their identities. The musculoskeletal structure in human produces unique finger movements resulting in distinguishable keystrokes or touch-points which can be utilized in anchoring an extra layer of security for user authentication.

Touch-biometrics have been studied for both one-shot and continuous user authentication on smartphones. Touch dynamics refers to user profiling based on touch patterns (i.e., touch duration and direction, etc.) on a touchscreen. A touchscreen enables users to interact with the smartphone by touching different locations on the screen. In [45], a two-tier user authentication solution involving passphrases and keystroke dynamics to increase both usability and security. The scheme exploits the use of passphrases as the first tier of authentication augmented with a keystroke authentication algorithm as the second tier of authentication.

The touch-based scheme [46] leverages different touch features: X and Y coordinates, touch-pressure, the size of touch and the time offset, generated from different slide operations to identify a user. Authors report 77% accuracy (with 19% FRR and 21% FAR) using DTW as the classifier over a dataset of 48 participants. Feng et al. [47] presented a finger-gesture based authentication system (called as FAST) in addition to the digital gloves. Every touch gestures include 53 features: X & Y coordinates, the direction of finger motion, the pressure at each sample touch-point, and the distance between multi-touch points. Digital gloves add angular values from X, Y and Z direction in addition to roll, pitch, and yaw values. FAST achieved a FAR of 4.66% and FRR of 0.13% on a dataset of 40 users using Decision tree, Random Forest and Bayes net classifiers.

A study by Frank et al. [48] also explores the touchscreen gestures for continuous smartphone user authentication.

This mechanism exploits the very common navigational movements (e.g., horizontal/vertical strokes) and shows their efficacy to authenticate the real user. This study achieves an EER of 0%, 2 – 3% and <4%, respectively, in intra-session, inter-session and authentication tests after one week of enrollment using KNN classifier and SVM - with Gaussian Radial Basis Function (RBF) kernel, on a dataset of 41 testers.

Sae-Bae et al. [49] exploit single and multitouch gestures for user authentication on touch-sensitive devices, i.e., smartphones and tablets. On a dataset of 34 participants, they report an average EER of 7.88% using a single instance of multitouch gesture and an EER of 1.58% with a combination of three gestures (static counter-clockwise rotation, closed and opened, with all five fingertips). Authentication solution [50] profiles simple touch actions, i.e., keystroke, sliding, pinch, and handwriting and continuously authenticates the smartphone user. The scheme leverages multiple features related to coordinates, pressure, size, etc., and achieves the lowest EER of 0.75% for sliding gesture and for all other action types, lower than 10% with SVM classifier using RBF kernel.

### 8.1.2 Sensors/motion based authentication:

In addition to the touch-based solutions, researchers have also exploited smartphone's built-in physical 3-dimensional sensors, such as accelerometer, gyroscope, orientation, etc., to profile phone movements, for smartphone user authentication. The data from these sensors is used to identify users from their walking patterns [51], general hand-movement [52–54], special hand-movement (while entering PIN, password) [55, 56], and hand-movement (how a user moves the phone to place or answer a call [57] and profiled gesture models [53], etc.

The study by Shi et al. [54] presents a multi-sensor-based approach to passively identify a real user. Their system incorporates the accelerometer, touch screen, voice and location data for user authentication. They achieve around 97% TPR, using the Naive Bayes as the classifier, from their dataset of 7 users (three females and four males). The study [52] explores the role of three sensors: accelerometer, orientation, and compass in addition to the touch gestures towards continuous user authentication. This transparent mechanism profiles finger movements with classical touch-based features and interprets the sensed data as different gestures. It then trains the SVM classifier on those gestures and performs authentication tasks. The paper reports as high as 95.78% accuracy on a database of 75 users.

The study by Zhu et al. [53] proposes a mobile framework model *Sensec* based on the accelerometer, orientation, gyroscope, and magnetometer, to construct a user gesture profile. The model then continuously computes the sureness score and keep the user sign-in.

By concatenating X, Y, Z values from these sensors, they identify a valid user with 75% accuracy and an adversary with an accuracy of 71.3% (with 13.1% FAR) on their collected dataset of 20 users. However, the study required a user to follow a script and collects the sensory data for the entire duration of that interaction.

### 8.1.3 Sensor-enhanced touch-typing based authentication:

Our scheme is a bimodal system which leverages the timing-differences from the entered 8-digit “text-independent” secret and the hand-movements while the user enters the text to sign-in to the security-sensitive apps, we compare our work with the closely related works proposed in the literature, i.e., [35, 58, 59].

Giuffrida et al., [58], proposed sensor enhanced fix-text scheme for user authentication on Android smartphones. They reported 4.97% EER on fixed-text passwords and 0.08% on sensor data on a dataset of 20 users. Later, Buriro et al. [42] modeled sensory readings as hold behavior and introduced free-text secret the user needs to enter or writes on the touchscreen. They achieved 1% EER on a dataset of 12 users for touch-typing [55] and  $\approx 95\%$  TAR at 3.1% FAR on the dataset of 30 users.

The papers discussed here implemented a behavioral biometric-based authentication scheme performed in in-the-lab supervised settings, and their analysis was based on a small number of users, e.g., just 12 [55], 20 [58], and 30 [42]. We evaluated our scheme on a comparatively larger dataset of 95 users collected in-the-wild. Since the number of users in previous studies was less and data was collected in in-lab settings, it is difficult to examine how their achieved error would have varied if the number of users was more and data was collected in-the-wild. Also, we evaluated our data by applying multi-class classification to replicate a server-based remote client authentication with the risk-based authentication mechanism. However, the papers discussed here evaluated their data either using one class or binary class classification approaches - replicating authentication only on smartphones.

## 8.2 Risk-based Authentication Schemes

Most of the systems deploying risk-based authentication approaches typically generate a risk profile for each of the users. Based on the risk score, the complexity of the challenge is determined to authenticate the user, i.e., a higher risk score leads to stronger authentication, whereas a risk score below the threshold means minimal or no authentication requirement [60].

Risk-based authentication approaches based on basic communication information [61], such as the source-destination IP addresses, or frequency of transactions,

performed by a user on her devices to determine risk, are easily exploitable. According to Traore [62], such systems could be exploited by polling or cloning users’ devices. Then, the same settings can be replicated on different machines to access their systems by attackers.

Cognitive fraud detection system by IBM Trusteer [63] is designed for PCs and laptops. Whereas, IBM’s Tivoli Federated Identity Manager [64] is designed for web platform based on policy rules that determine the access request to be allowed, denied, or challenged at run-time. However, these are limited to static devices only, e.g., a personal computer and laptops, etc.

Sepczuk et al. [65] designed the remote-services for authentication management, which can be registered by the user either manually or automatically. Manual registration requires users to fill a form describing their day-to-day activities, e.g., what they do between 9 a.m. to 5 p.m.? or which network they use at home or workplace. Whereas, automatic data gathering configures the system to collect contextual data, spontaneously. However, the solution may be subjected to insider attacks and lacks transparency, as service providers could misuse user contextual data, i.e., they are aware of an individual’s day-to-day activities.

Generally, the contextual or historical data or both, to generate a risk profile of a user, is considered more suitable for risk-based authentication approaches [66, 67]. However, the existing systems apply simplistic risk management models or ad-hoc rule-based techniques, which prove to be ineffective for risk assessment [68]. Furthermore, they mainly rely on knowledge-based authentication mechanisms such as `username/password`, or multi-factor authentication (e.g., OTP, token generator) [11], which affects the usability of a system adversely.

## 9 Conclusions & Future Work

The proposed one-shot-cum-continuous user authentication scheme is a simple, effective, and user-friendly solution for smartphone security-sensitive applications (e.g., social networking app, online mobile banking app, etc.). The scheme can be seamlessly integrated into the existing PIN/password-based authentication schemes to enhance their usability and security. Flexibility to access an application by entering any random 8-digit alphanumeric text makes the sign-in process very convenient for smartphones users. At the same time, mimicking invisible, and inherently secure natural human behaviors simultaneously can be an onerous job for attackers.

With RF classifier, we obtained 96% TAR (at the cost of 0.01% FAR) in *sitting* activity for 15 samples training-set with selected features, whereas 95.92% and 94.87% TAR is achieved in *standing* and *walking* activity, respectively. Our

scheme obtained a SUS score of  $\approx 73$  out of 100 that can be considered positive feedback.

We will further improve and fine-tune our prototype for wider user-acceptability. In future, we will evaluate

performance, i.e., power consumption, computational constraints, i.e., CPU and memory overhead, the sample-acquisition- and decision-making time, in standard lab environment.

## Appendix A: Comparison of user authentication schemes

Study	Features	Evaluation	Participants	Performance
[46]	Touch features X and Y coordinates, touch-pressure, the size of touch and the time offset.	DTW	48	FRR: 19% and FAR: 21%
[47]	X & Y coordinates, the direction of finger motion, the pressure at each sample touch-point, and the distance between multi-touch points. Digital gloves add angular values from X, Y and Z direction in addition to roll, pitch, and yaw values.	Decision tree, Random Forest and Bayes net classifier.	40	FRR: 0.13% and FAR: 4.66%
[48]	Touchscreen gestures like navigational movements. (e.g., horizontal/vertical strokes)	KNN classifier and SVM with Gaussian Radial Basis Function (RBF) kernel	41	EER: <4%
[49]	Single and multi touch gestures by combining static counter-clockwise rotation, closed and opened gesture with all five fingertips).	Ppairwise distance calculation and score calculation	34	EER: 7.88% (Single), 1.58% (Combined)
[50]	Simple touch actions, i.e., keystroke, sliding, pinch, and handwriting to extract features like coordinates, pressure, size, etc.	SVM classifier using RBF kernel	30	EER: 0.75% (Sliding gesture)
[54]	Extracts finger movements and touch features using accelerometer, touch screen, voice and location data	Naive Bayes classifier	7	TPR: 95.78%
[52]	Extracted touch positions, touch pressure, touch area, moving direction, distance, duration, average moving direction and curvature, average curvature distance, average pressure, average touch-area, max-area portion, min-area portion.	SVM classifier	75	Accuracy: 95.78%
[53]	Construct feature vectors from X, Y, Z values acquired from sensors and clustered them into V classes using K-means algorithm.	Continuous n-gram language model	20	Accuracy: 75%
[58]	Extract statistical features for touch dynamics from the raw data acquired from the sensors.	Distance metrics: Euclidean, Euclidean normed, Manhattan, Manhattan scaled.	20	EER: 4.97% (fixed-text passwords), 0.08% (sensor data)
[63]	Static, contextual, and analytically calculated attributes	Provides policy rules that determine whether an access request must be permitted, denied, or challenged.	-	Calculates a risk score based on multiple weighted attributes
[67]	User location and contextual data are associated with different risk assessments and accordingly user authentication was applied.	Risk-aware Authentication as per the user location.	-	CORMORANT Framework
[66]	Acquire language, color depth, screen resolution, timezone, platform, plugins, etc. IP address range, time of access, geolocation, request headers, etc.	Adaptive and dynamic context fingerprinting based on Hoeffding trees	-	SmartAuth continuously assess the risk of fraudulent activities during long-lived user authenticated session.

## Appendix B: Demographic Questionnaire

1. What is your gender?

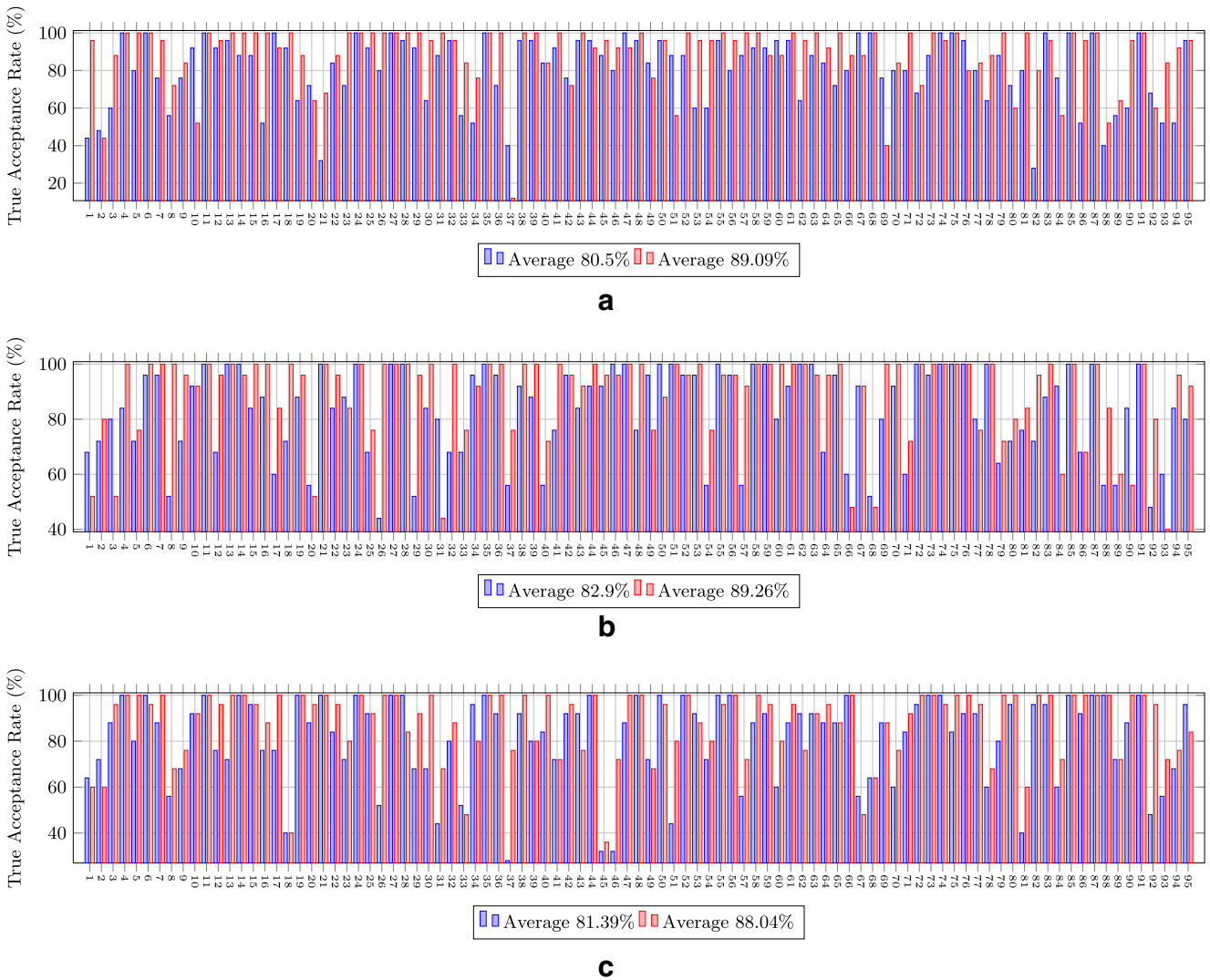
- Male
- Female
- I don't want to disclose

- How old you are?

- $\leq$  than 20 years.
- $> 20$  years and  $\leq 40$  years.
- $> 40$  years and  $\leq 60$  years.

- > than 60 years.
- I don't want to disclose
- Tell us about your nationality.
- \_\_\_\_\_
- I don't want to disclose
- Which hand(s) do you use for interacting with your smartphone?
  - Right
  - Left
  - Both
  - I don't want to disclose

**Appendix C: TAR comparison of RF classifier for individual users in 3 activities**



**Figure 7** Comparison of RF classifier performance (TAR) on 5-sample training over full and IGAE features.



**Acknowledgments** This project has received funding from the European Union’s Horizon 2020 research and innovation programme DS 2018-2019-2020 as part of the E-Corridor project ([www.e-corridor.eu](http://www.e-corridor.eu)) under grant agreement No 883135.

## References

- Statista (2018). What authentication methods do you usually use when logging in to your main bank? <https://www.statista.com/statistics/786638/online-banking-authentication-security-methods-usage-united-kingdom/>. online web resource.
- Gupta, S., Buriro, A., Crispo, B. (2019). Driverauth: Behavioral biometric-based driver authentication mechanism for on-demand ride and ridesharing infrastructure. *ICT Express*, 5(1), 16–20.
- Katsini, C., Belk, M., Fidas, C., Avouris, N., Samaras, G. (2016). Security and usability in knowledge-based user authentication: A review. In *Proceedings of the 20th Pan-Hellenic conference on informatics* (p. 63): ACM.
- Aviv, A.J., Gibson, K.L., Mossop, E., Blaze, M., Smith, J.M. (2010). Smudge attacks on smartphone touch screens. *Woot*, 10, 1–7.
- Ye, G., Tang, Z., Fang, D., Chen, X., Kim, K.I., Taylor, B., Wang, Z. (2017). Cracking android pattern lock in five attempts. In *Proceedings 2017 network and distributed system security symposium 2017 (NDSS’17)*.
- CAPEC-Release1.6 (2016). Common attack pattern enumeration and classification. online web resource.
- Bhattachali, T., Saeed, K., Chaki, N., Chaki, R. (2014). A survey of security and privacy issues for biometrics based remote authentication in cloud. In *Proceeding of IFIP International conference on computer information systems and industrial management* (pp. 112–121): Springer.
- Zhang-Kennedy, L., Chiasson, S., van Oorschot, P. (2016). Revisiting password rules: facilitating human management of passwords. In *Proceedings of APWG symposium on electronic crime research (eCrime)* (pp. 1–10): IEEE.
- Komanduri, S., Shay, R., Kelley, P.G., Mazurek, M.L., Bauer, L., Christin, N., Cranor, L.F., Egelman, S. (2011). Of passwords and people: measuring the effect of password-composition policies. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 2595–2604): ACM.
- Shila, D.M., & Srivastava, K. (2018). Castra: Seamless and unobtrusive authentication of users to diverse mobile services. *IEEE Internet of Things Journal*, 5(5), 4042–4057.
- Gupta, S., Buriro, A., Crispo, B. (2018). Demystifying authentication concepts in smartphones: Ways and types to secure access. *Mobile Information Systems*, 2018.
- Gupta, S. (2020). Next-generation user authentication schemes for iot applications, Ph.D. dissertation, University of Trento, Italy.
- Halunen, K., Häikiö, J., Vallivaara, V. (2017). Evaluation of user authentication methods in the gadget-free world. *Pervasive and Mobile Computing*, 40, 220–241.
- Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T., Koucheryavy, Y. (2018). Multi-factor authentication: a survey. *Cryptography*, 2(1), 1.
- Gupta, S., Buriro, A., Crispo, B. (2019). Smarthandle: A novel behavioral biometric-based authentication scheme for smart lock systems. In *Proceeding of the 3rd international conference on biometric engineering and applications*: ACM.
- Rui, Z., & Yan, Z. (2019). A survey on biometric authentication: Toward secure and privacy-preserving identification. *IEEE Access*, 7, 5994–6009.
- N. R. Council, W. B. Committee, et al. (2010). *Biometric recognition: challenges and opportunities*. Washington: National Academies Press.
- EU, Principles of the gdpr, [https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/principles-gdpr\\_en](https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/principles-gdpr_en), Accessed on 16-08-2020, online web resource.
- C. Constitution, Ab-375 privacy: personal information: businesses, [https://leginfo.ca.gov/faces/billTextClient.xhtml?bill\\_id=2\\_01720180AB375](https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=2_01720180AB375), Accessed on 16-08-2020, online web resource.
- Krausova, A., & recognition, O.nline.b.ehavior. (2018). Can we consider it biometric data under gdpr. *Masaryk UJL & Tech.*, 12, 161.
- Minaee, S., Abdolrashidi, A., Su, H., Bennamoun, M., Zhang, D. (2019). Biometric recognition using deep learning: A survey. arxiv:1912.00271.
- Neal, T.J., & Woodard, D.L. (2016). Surveying biometric authentication for mobile device security. *J Pattern Recognit Res*, 1, 74–110.
- Still, J.D., Cain, A., Schuster, D. (2017). Human-centered authentication guidelines. *Information & Computer Security*.
- Gupta, S., & Crispo, B. (2019). A perspective study towards biometric-based rider authentication schemes for driverless taxi. In *Proceedings of the international conference on innovation and intelligence for informatics, computing, and technologies (3ICT)*. IEEE (pp. 1–6).
- Buriro, A., Akhtar, Z., Crispo, B., Gupta, S. (2017). Mobile biometrics: Towards a comprehensive evaluation methodology. In *Proceedings of the international carnanhan conference on security technology (ICCST)* (pp. 1–6): IEEE.
- Gupta, S., Buriro, A., Crispo, B. (2019). A risk-driven model to minimize the effects of human factors on smart devices. In *Proceedings of the international workshop on emerging technologies for authorization and authentication* (pp. 156–170): Springer.
- D. Insights Articles, Risk-based authentication: A primer, <https://deloitte.wsj.com/cio/2013/10/30/risk-based-authentication-a-primer/>, 2013, online web resource.
- O. M. S. Project, Owasp mobile security project. accessed: Dec. 2016, [https://www.owasp.org/index.php/OWASP\\_Mobile\\_Security\\_Project](https://www.owasp.org/index.php/OWASP_Mobile_Security_Project), 2016, online web resource.
- Android, Developers guide: Sensorevent, <https://developer.android.com/reference/android/hardware/SensorEvent.html>, 2018, online web resource.
- Buriro, A., Gupta, S., Crispo, B. (2017). Evaluation of motion-based touch-typing biometrics in online financial environments. In *BIOSIG*, Vol. 2017.
- Pires, I., Garcia, N., Pombo, N., Flórez-Revuelta, F. (2016). From data acquisition to data fusion: a comprehensive review and a roadmap for the identification of activities of daily living using mobile devices. *Sensors*, 16(2), 184.
- Han, J., Pei, J., Kamber, M. (2011). *Data mining: concepts and techniques*. Amsterdam: Elsevier.
- Demuth, H.B., Beale, M.H., De Jess, O., Hagan, M.T. (2014). *Neural network design*. Stillwater: Martin Hagan.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Buriro, A., Crispo, B., Gupta, S., Del Frari, F. (2018). Dialerauth: A motion-assisted touch-based smartphone user authentication scheme. In *Proceedings of the eighth ACM conference on data and application security and privacy* (pp. 267–276): ACM.
- ISO9000:2015, Quality management systems — fundamentals and vocabulary, <https://www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:v1:en>, 2015, online web resource.

37. Wu, T., Blackhurst, J., Chidambaram, V. (2006). A model for inbound supply risk analysis. *Computers in Industry*, 57(4), 350–365.
38. Fawcett, T., & graphs, R.oc. (2004). Notes and practical considerations for researchers. *Machine Learning*, 31(1), 1–38.
39. Usability, System usability scale (sus), <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>, 2018, online web resource.
40. Trewin, S., Swart, C., Koved, L., Martino, J., Singh, K., Ben-David, S. (2012). Biometric authentication on a mobile device: a study of user effort, error and task disruption. In *Proceedings of the 28th annual computer security applications conference* (pp. 159–168): ACM.
41. Nguyen, T.V., Sae-Bae, N., Memon, N. (2017). Draw-a-pin. *Computers and Security*, 66(C), 115–128.
42. Buriro, A., Crispo, B., DelFrari, F., Wrona, K. (2016). Hold and sign: A novel behavioral biometrics for smartphone user authentication. In *Proceeding of IEEE security and privacy workshops (SPW)* (pp. 276–285): IEEE.
43. Sauro, J. (2011). Measuring usability with the system usability scale (sus).
44. Ritchie, R., Rubino, D., Michaluk, K., Nickinson, P. (2013). The future of authentication: Biometrics, multi-factor, and co-dependency, <https://www.androidcentral.com/talk-mobile/future-authentication-biometrics-multi-factor-and-co-dependency-talk-mobile>. online web resource.
45. Bhana, B., & Flowerday, S. (2020). Passphrase and keystroke dynamics authentication: Usable security. *Computers & Security*, 101925.
46. De Luca, A., Hang, A., Brudy, F., Lindner, C., Hussmann, H. (2012). Touch me once and i know it's you!: implicit authentication based on touch screen patterns. In *Proceedings of conference on human factors in computing systems proceedings of the SIGCHI* (pp. 987–996): ACM.
47. Feng, T., Liu, Z., Kwon, K.-A., Shi, W., Carbanar, B., Jiang, Y., Nguyen, N. (2012). Continuous mobile authentication using touchscreen gestures. In *Proceeding of IEEE conference on technologies for homeland security (HST)*. IEEE (pp. 451–456).
48. Frank, M., Biedert, R., Ma, E., Martinovic, I., Song, D. (2013). Touchalytics: on the applicability of touchscreen input as a behavioral biometric for continuous authentication. *IEEE Transactions on Information Forensics and Security*, 8(1), 136–148.
49. Sae-Bae, N., Memon, N., Isbister, K., Ahmed, K. (2014). Multitouch gesture-based authentication. *IEEE Transactions on Information Forensics and Security*, 9(4), 568–582.
50. Xu, H., Zhou, Y., Lyu, M.R. (2014). Towards continuous and passive authentication via touch biometrics: An experimental study on smartphones. In *Proceedings of symposium on usable privacy and security (SOUPS)*, (Vol. 14 pp. 187–198).
51. Mantyjarvi, J., Lindholm, M., Vildjiounaite, E., Makela, S.-M., Aillisto, H. (2005). Identifying users of portable devices from gait pattern with accelerometers. In *Proceedings of IEEE international conference on acoustics, speech, and signal processing (ICASSP)*, (Vol. 2 pp. ii–973): IEEE.
52. Li, L., Zhao, X., Xue, G. (2013). Unobservable re-authentication for smartphones. In *Proceedings of NDSS*, (Vol. 56 pp. 57–59).
53. Zhu, J., Wu, P., Wang, X., Zhang, J. (2013). Sensec: Mobile security through passive sensing. In *Proceedings of international conference on computing, networking and communications (ICNC)* (pp. 1128–1133): IEEE.
54. Shi, W., Yang, J., Jiang, Y., Yang, F., Xiong, Y. (2011). Senguard: Passive user identification on smartphones using multiple sensors. In *Proceedings of the 7th international conference on wireless and mobile computing, networking and communications (WiMob)* (pp. 141–148): IEEE.
55. Buriro, A., Crispo, B., Del Frari, F., Wrona, K. (2015). Touchstroke: smartphone user authentication based on touch-typing biometrics. In *Proceeding of international conference on image analysis and processing* (pp. 27–34): Springer.
56. Sitová, Z., Šeděnka, J., Yang, Q., Peng, G., Zhou, G., Gasti, P., Balagani, K.S. (2016). Hmog: New behavioral biometric features for continuous authentication of smartphone users. *IEEE Transactions on Information Forensics and Security*, 11(5), 877–892.
57. Conti, M., Zachia-Zlatea, I., Crispo, B. (2011). Mind how you answer me!: transparently authenticating the user of a smartphone when answering or placing a call. In *Proceedings of the 6th ACM symposium on information computer and communications security* (pp. 249–259): ACM.
58. Giuffrida, C., Majdanik, K., Conti, M., Bos, H. (2014). I sensed it was you: authenticating mobile users with sensor-enhanced keystroke dynamics. In *Proceeding of international conference on detection of intrusions and malware, and vulnerability assessment* (pp. 92–111): Springer.
59. Buriro, A., Crispo, B., Eskandri, M., Gupta, S., Mahboob, A., Van Acker, R. (2018). Snap auth: a gesture-based unobtrusive smart-watch user authentication scheme. In *Proceedings of the international workshop on emerging technologies for authorization and authentication* (pp. 30–37): Springer.
60. Schneier, B. (2013). Risk-based authentication, [https://www.schneier.com/blog/archives/2013/11/risk-based\\_auth.html](https://www.schneier.com/blog/archives/2013/11/risk-based_auth.html). online web resource.
61. Butler, M., & Butler, R. (2015). Investigating the possibility to use differentiated authentication based on risk profiling to secure online banking. *Information & Computer Security*, 23(4), 421–434.
62. Traoré, I., & Ahmed, A.A.E. (2011). Introduction to continuous authentication. In *Continuous authentication using biometrics: data, models, and metrics: data, models, and metrics* (p. 1).
63. IBM (2016). Ibm trustee, <http://www-03.ibm.com/software/products/en/category/advanced-fraud-protection>. online web resource.
64. IBM (2016). Ibm tivoli federated identity manager, [https://www.ibm.com/support/knowledgecenter/en/SZSXU\\_6.2.2.7/com.ibm.tivoli.fim.doc\\_6227/rbaOverview.html](https://www.ibm.com/support/knowledgecenter/en/SZSXU_6.2.2.7/com.ibm.tivoli.fim.doc_6227/rbaOverview.html). online web resource.
65. Sepczuk, M., & Kotulski, Z. (2018). A new risk-based authentication management model oriented on user's experience. *Computers & Security*, 73, 17–33.
66. Preuveneers, D., & Joosen, W. (2015). Smartauth: dynamic context fingerprinting for continuous user authentication. In *Proceedings of the 30th annual ACM symposium on applied computing* (pp. 2185–2191): ACM.
67. Hintze, D., Koch, E., Scholz, S., Mayrhofer, R. (2016). Location-based risk assessment for mobile authentication. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: Adjunct* (pp. 85–88): ACM.
68. Haimes, Y.Y. (2015). *Risk modeling, assessment, and management*. Wiley: Hoboken.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.