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

Attaullah Buriro[1][0000−0003−2723−2410],
Sandeep Gupta[1][0000−0001−9220−7700] (sandeep.gupta@unitn.it),
Artsiom Yautsiukhin[2], and Bruno Crispo[1][0000−0002−1252−8465]

[1] Department of Information Engineering & Computer Science (DISI), University of Trento, Italy
[2] Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche, Pisa, Italy

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 application sign-in process but also, throughout the entire 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. Moreover, throughout 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 a $8-digit$ 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. 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*[3] 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 [1,2]. 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 [2]. Thus, it becomes requisite to redesign the PIN/password-based authentication mechanism to overcome their inherent shortcomings.

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*[3] survey.

Paper organization: The rest of the paper is organized as the following: Section 2 discusses the threat model, the working of our proposed scheme, and architecture of our system. In Section 3, we discuss the methodology used to design our one-shot-cum-continuous authentication scheme. Section 4 presents the obtained results. In Section 5, we assess the usability of our proposed system. Section 6 surveys the related approaches proposed over the years for user authentication. Finally, in Section 7, we conclude the paper with a summary of the work and the possible future dimensions.

2 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 it's system architecture.

[3] `https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html`

2.1 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 [10,11] 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.

2.2 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.

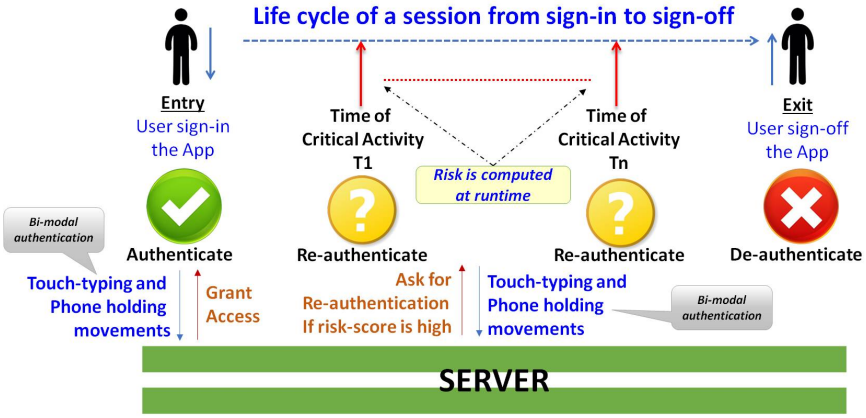


Fig. 1. Our one-shot-cum-continuous authentication scheme model [**?**].

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.

2.3 System Architecture

The system adopts a client-server architecture [12] as shown in Figure 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.

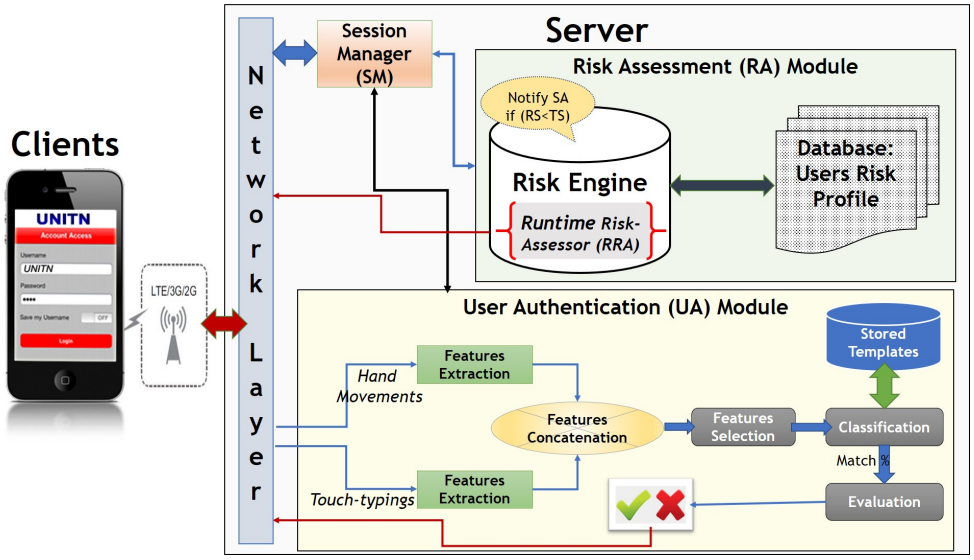


Fig. 2. System Architecture

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 3.2. 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 3.6, 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.

3 Methodology

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

3.1 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 UBERTESTERS[4] - a crowdsourcing software testing platform. Testers were certified quality assurance engineers or experienced software developers and they were rewarded on an hourly basis. The 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.

[4] https://ubertesters.com/

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 1.

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.

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

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

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.

3.2 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 [13]. The high-pass and low-pass sensory data is computed mathematically, by applying High-Pass (HP) and Low-Pass (LP) filters as shown in Equation 1 and 2.

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

$$Value_{LP} = Value_{Accelerometer} - Value_{Gravity} \tag{2}$$

Where, $Value_{HP}$, $Value_{LP}$, $Value_{Accelerometer}$, and $Value_{Gravity}$ represent the value of the high-pass, low-pass, accelerometer, and gravity sensor, respectively at a time t. We set α to 0.1 that was determined, empirically.

As shown in Figure 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.

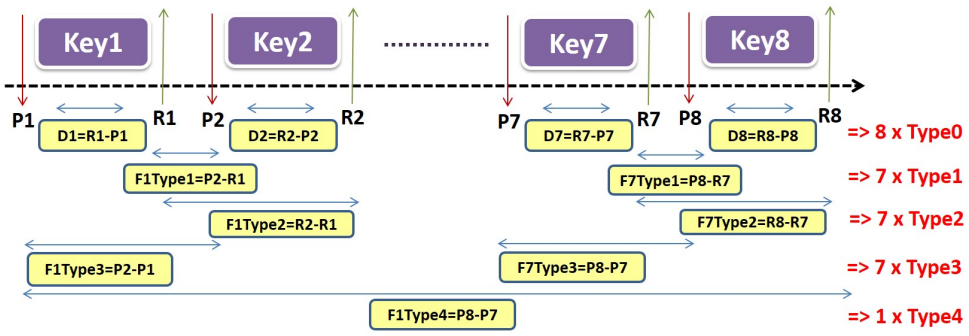


Fig. 3. Touch-typing features for 8-keys entry [14].

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^{th} dimension, Magnitude (M), by using Equation 3.

$$Value_M = \sqrt{(Value_x^2 + Value_y^2 + Value_z^2)} \tag{3}$$

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* [13]. Then, from each data stream, we extracted 4 statistical features, namely Mean (μ), Standard Deviation (σ), Skewness (s), and Kurtosis (k), that gives 16 statistical features per sensor as shown in Table 2.

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

No.	Hand-movement Features			
1-4	μ_X	μ_Y	μ_Z	μ_M
5-8	σ_X	σ_Y	σ_Z	σ_M
9-12	s_X	s_Y	s_Z	s_M
13-16	k_X	k_Y	k_Z	k_M

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 [15].

3.3 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 [16]. This improves the performance of a classifier as processing smaller feature vectors would be computationally faster. We applied Information Gain Attribute Evaluator (IGAE) - a Weka [17] implemented Information Gain

based feature selection scheme. This scheme evaluates the worth of a feature by computing its information gain with respect to the class [18]. We obtained the threshold for feature selection by dividing the number of users (95) by the total number of features (142). The feature with higher weight was picked for further analysis.

3.4 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 [19].

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 [20]. 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 [21]. 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 large datasets, etc. We preferred this classifier because of its effectiveness in the previous studies [22,23]. 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).

3.5 Classifier Training & Testing

We consider remote-user-authentication to access security-sensitive applications on smartphones as a multiclass classification problem. We used PRTools [24], a Matlab-based toolbox, to construct a classification model and validated users in two scenarios, (i) a verifying legitimate user scenario, and (ii) an 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.

3.6 Risk Assessment Model

According to ISO 9000:2015 [25], 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 [26]. 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.

$$CA_i \;=\; VALUE \begin{cases} \forall\, i \in M \\ VALUE \;\geq\; 1 \end{cases} \tag{4}$$

Equation 4 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 $CA_{IMIE} > CA_{AccessTime}$. 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. Structure of User's Client Profile

#	Client-Attributes	Weight of Client-Attributes	$Session_N$	$Session_{N-1}$	...	$Session_2$	$Session_1$	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	$Value_{5N}$	$Value_{5(N-1)}$	...	$Value_{52}$	$Value_{41}$	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

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., $Session_N$ 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 [27], which is determined using Equation 5.

$$
\begin{aligned}
O_i &= \sum_{j=1}^{N-1} [Value_{iN} = Value_{ij}] \; \forall \; i \in M-1, \quad \text{and} \\
FO_i &= \frac{O_i}{N-1} \; \forall \; i \in M-1
\end{aligned}
\tag{5}
$$

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

Subsequently, `Frequency of Non-occurrence` (FNO_i) and `Impact of Non-occurrence` (INO_i) are measured at runtime using Equation 6 and Equation 7, respectively.

$$
FNO_i = 1 - FO_i \quad \forall \; i \in M \tag{6}
$$

$$
INO_i = FNO_i \times CA_i \quad \forall \; i \in M \tag{7}
$$

Where, FO_i is defined as the frequency of occurrence, which can be calculated using Equation 5, CA_i is the weight of each client-attribute and M is the number of client-attributes. The value of FNO_i towards ≈ 0 indicates lower risk, whereas towards ≈ 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 ($FO_{location} = \frac{0}{10}$) becomes 0. Therefore, the frequency of its non-occurrence ($FNO_{location}$) becomes 1, which is calculated using Equation 6. As described in Equation 7, 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 Equation 8, 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.

$$
Risk\ Score = \sum_{i=1}^{M} INO_i \tag{8}
$$

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 2.2, offering a user friendly verification mechanism.

4 Results

4.1 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.

4.2 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.

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

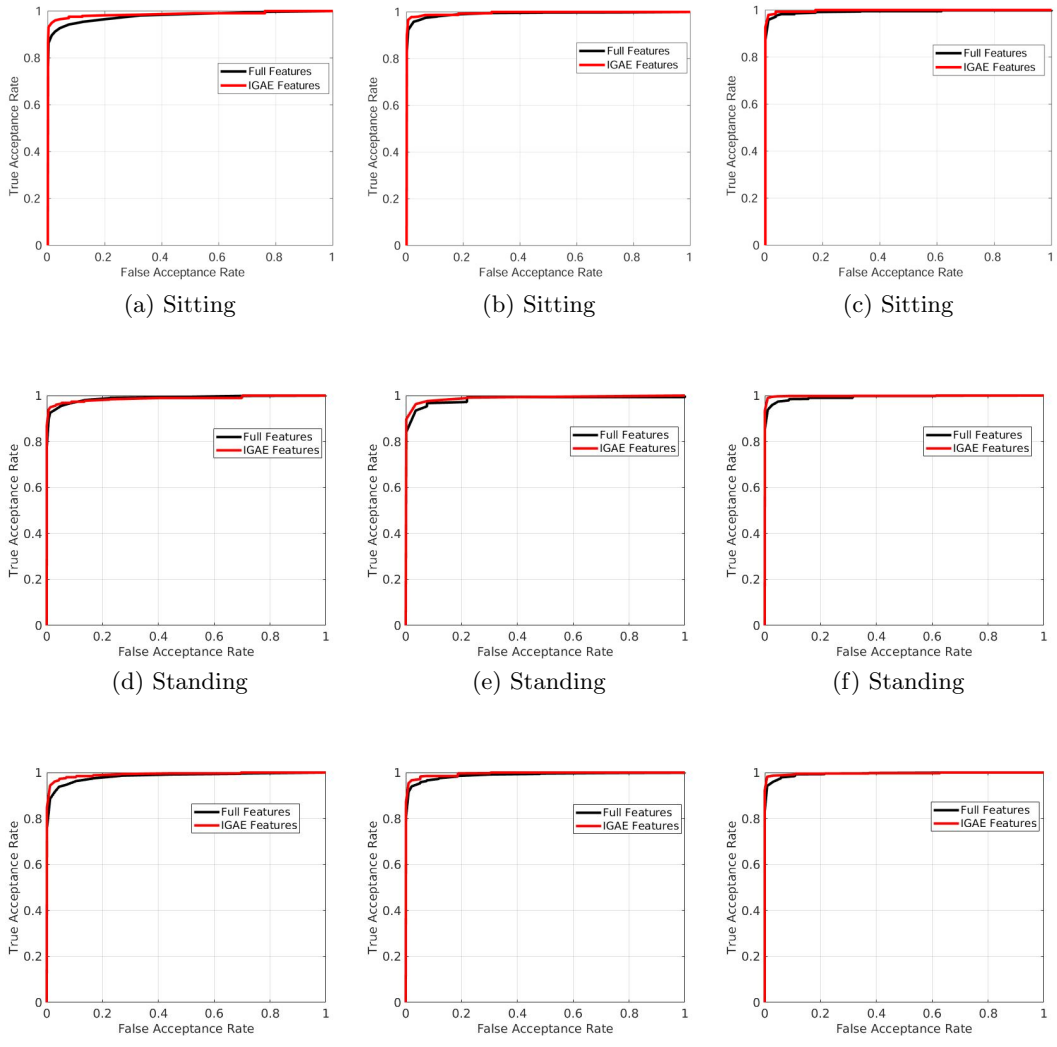
Training Samples	5						10						15					
Activity	*sitting*		*standing*		*walking*		*sitting*		*standing*		*walking*		*sitting*		*standing*		*walking*	
Classifiers	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR
NB	72.72	0.24	79.53	0.18	76.04	0.21	83.66	0.12	84.51	0.11	80.55	0.14	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
RF	**80.51**	**0.17**	**82.91**	**0.17**	**81.38**	**0.19**	**87.87**	**0.09**	**88.72**	**0.12**	**85.31**	**0.16**	**91.79**	**0.04**	**91.58**	**0.08**	**86.95**	**0.13**

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

Training Samples	5						10						15					
Activity	*sitting*		*standing*		*walking*		*sitting*		*standing*		*walking*		*sitting*		*standing*		*walking*	
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
RF	**89.10**	**0.10**	**89.26**	**0.09**	**88.04**	**0.10**	**95.18**	**0.04**	**93.64**	**0.04**	**92.88**	**0.05**	**96.00**	**0.01**	**95.92**	**0.02**	**94.87**	**0.02**

Then, 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.

Since RF classifier performed pretty well on both full and IGAE features in all the activities, we also show the distribution of TAR (per user) for *sitting, standing, walking* activities,



(a) Sitting (b) Sitting (c) Sitting

(d) Standing (e) Standing (f) Standing

(g) Walking (h) Walking (i) Walking

Fig. 4. 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)

obtained on just 5 training samples, in B. 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 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 Figure 6 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 Figure 6a) 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 figures 6b and 6c).

We also show the results of RF classifier in terms of ROC curves (see Figure 4). We show an average ROC of all the users obtained through Vertical Averaging (VA) [28]. 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 4 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 overfitting. 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.

5 Usability Analysis

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

5.1 Methodology

System Usability Scale (SUS) [29] is considered as a standard tool to record user experience related to the usability of a system and has been extensively used in the context of smartphone user authentication [30,31,32]. 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 [31,30]. 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.

5.2 Responses

Figure 5 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 [33]. 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.

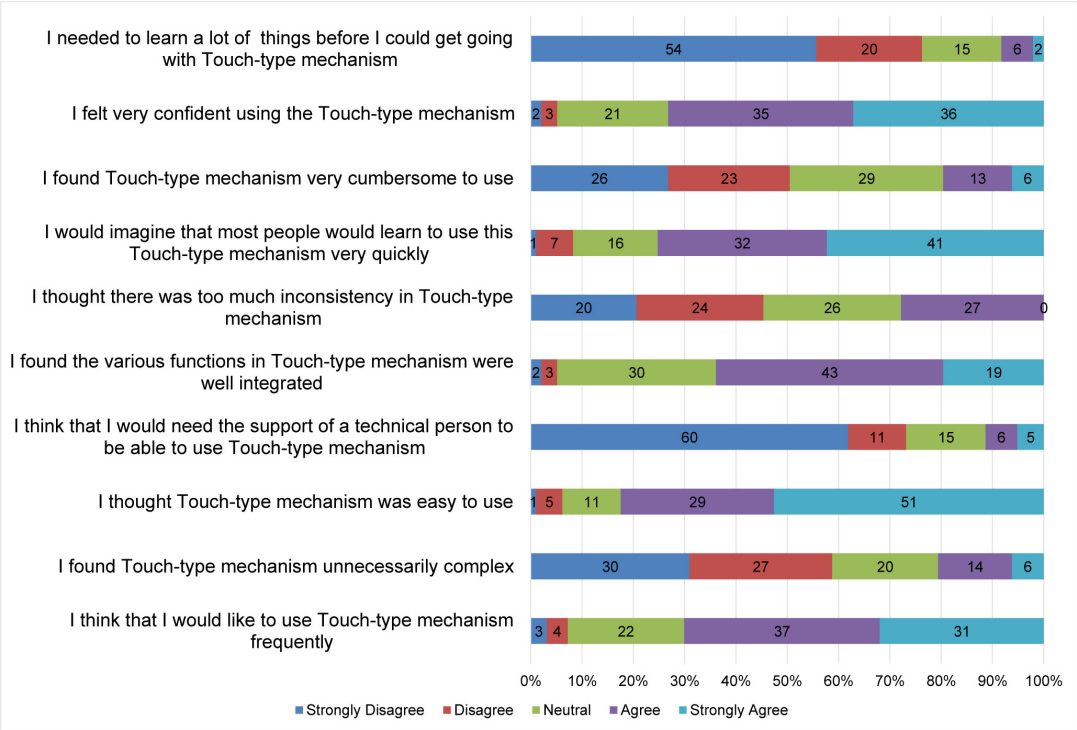


Fig. 5. SUS questionnaire and Users responses

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.

6 Related Work

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

6.1 Behavioral Biometric-based User Authentication

Behavioral biometrics offers a simple way to implement a frictionless user authentication schemes, which are suitable for continuous authentication [2]. This is possible due to the advantages associated with behavioral biometrics: 1) transparent collection, 2) no special hardware requirements, and 3) cost effective deployment [34].

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.

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 dynamics refers to user profiling based on touch patterns (i.e., touch duration and direction, etc.) on the touchscreen. The touchscreen allows the user to interact with the smartphone by touching different locations on the screen. Touch-biometrics have been proposed for both one-shot and continuous user authentication on smartphones.

The touch-based scheme [35] 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. [36] 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 Gesture Sequence Based Authentication.

A study by Frank et al. [37] 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. [38] 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 multi-touch 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 [39] 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.

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 [40], general hand-movement [41,42,43], special hand-movement (while entering PIN, password) [18,44,45], and hand-movement (how a user moves the phone to place or answer a call [46,22] and profiled gesture models [42], etc.

The study by Shi et al. [43] 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 [41] 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. [42] 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.

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., [47,23,30].

Giuffrida et al., [47], 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. [44,30] 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 [44] 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 [44], 20 [47], and 30 [30]. 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 [44] - replicating authentication only on smartphones [30,45].

6.2 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 [48].

Risk-based authentication approaches based on basic communication information [49], 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 [50], 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 [51] is designed for PCs and laptops. Whereas, IBM's Tivoli Federated Identity Manager [52] 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. [53] 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 [51,54,55]. However, the existing systems apply simplistic risk management models or ad-hoc rule-based techniques, which prove to be ineffective for risk assessment [56]. Furthermore, they mainly rely on knowledge-based authentication mechanisms such as `username/password`, or multi-factor authentication (e.g., OTP, token generator) [2], which affects the usability of a system adversely.

7 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 ≈ 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 also perform security analysis, i.e., system robustness against common attacks such as mimic, shoulder surfing, replay attack, and performance evaluation, i.e., power consumption, computational constraints, i.e., CPU and memory overhead, the sample-acquisition- and decision-making time.

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 675320.

References

1. Statista, "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/`, 2018. online web resource.
2. S. Gupta, A. Buriro, and B. Crispo, "Demystifying authentication concepts in smartphones: Ways and types to secure access," *Mobile Information Systems*, vol. 2018, 2018.
3. C. Katsini, M. Belk, C. Fidas, N. Avouris, and G. Samaras, "Security and usability in knowledge-based user authentication: A review," in *Proceedings of the 20th Pan-Hellenic Conference on Informatics*, p. 63, ACM, 2016.

4. A. J. Aviv, K. L. Gibson, E. Mossop, M. Blaze, and J. M. Smith, "Smudge attacks on smartphone touch screens.," *Woot*, vol. 10, pp. 1–7, 2010.
5. G. Ye, Z. Tang, D. Fang, X. Chen, K. I. Kim, B. Taylor, and Z. Wang, "Cracking android pattern lock in five attempts," *Proceedings 2017 Network and Distributed System Security Symposium 2017 (NDSS'17)*, 2017.
6. CAPEC-Release1.6, "Common attack pattern enumeration and classification," 2016. online web resource.
7. T. Bhattasali, K. Saeed, N. Chaki, and R. Chaki, "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, 2014.
8. L. Zhang-Kennedy, S. Chiasson, and P. van Oorschot, "Revisiting password rules: facilitating human management of passwords," in *Proceedings of APWG Symposium on Electronic Crime Research (eCrime)*, pp. 1–10, IEEE, 2016.
9. S. Komanduri, R. Shay, P. G. Kelley, M. L. Mazurek, L. Bauer, N. Christin, L. F. Cranor, and S. Egelman, "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, 2011.
10. D. M. Shila and K. Srivastava, "Castra: Seamless and unobtrusive authentication of users to diverse mobile services," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 4042–4057, 2018.
11. O. M. S. Project, "Owasp mobile security project. accessed: Dec. 2016." `https://www.owasp.org/index.php/OWASP_Mobile_Security_Project`, 2016. online web resource.
12. S. Gupta, A. Buriro, and B. Crispo, "Driverauth: Behavioral biometric-based driver authentication mechanism for on-demand ride and ridesharing infrastructure," *ICT Express*, vol. 5, no. 1, pp. 16–20, 2019.
13. Android, "Developers guide: Sensorevent." `https://developer.android.com/reference/android/hardware/SensorEvent.html`, 2018. online web resource.
14. A. Buriro, S. Gupta, and B. Crispo, "Evaluation of motion-based touch-typing biometrics in online financial environments," *BIOSIG 2017*, 2017.
15. I. Pires, N. Garcia, N. Pombo, and F. Flórez-Revuelta, "From data acquisition to data fusion: a comprehensive review and a roadmap for the identification of activities of daily living using mobile devices," *Sensors*, vol. 16, no. 2, p. 184, 2016.
16. S. Gupta, A. Buriro, and B. Crispo, "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, 2019.
17. I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.
18. A. Buriro, B. Crispo, and Y. Zhauniarovich, "Please hold on: Unobtrusive user authentication using smartphone's built-in sensors," in *Proceeding of IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*, pp. 1–8, IEEE, 2017.
19. J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*. Elsevier, 2011.
20. H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural network design*. Martin Hagan, 2014.
21. L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
22. A. Buriro, B. Crispo, F. Del Frari, J. Klardie, and K. Wrona, "Itsme: Multi-modal and unobtrusive behavioural user authentication for smartphones," in *Proceeding of International Conference on Passwords*, pp. 45–61, Springer, 2015.
23. A. Buriro, B. Crispo, S. Gupta, and F. Del Frari, "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, 2018.
24. R. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. De Ridder, D. Tax, and S. Verzakov, "A matlab toolbox for pattern recognition," *PRTools version*, vol. 3, pp. 109–111, 2000.
25. 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.
26. 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.

27. T. Wu, J. Blackhurst, and V. Chidambaram, "A model for inbound supply risk analysis," *Computers in industry*, vol. 57, no. 4, pp. 350–365, 2006.
28. T. Fawcett, "Roc graphs: Notes and practical considerations for researchers," *Machine learning*, vol. 31, no. 1, pp. 1–38, 2004.
29. Usability, "System usability scale (sus)." `https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html`, 2018. online web resource.
30. A. Buriro, B. Crispo, F. DelFrari, and K. Wrona, "Hold and sign: A novel behavioral biometrics for smartphone user authentication," in *Proceeding of IEEE Security and Privacy Workshops (SPW)*, pp. 276–285, IEEE, 2016.
31. S. Trewin, C. Swart, L. Koved, J. Martino, K. Singh, and S. Ben-David, "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, 2012.
32. T. V. Nguyen, N. Sae-Bae, and N. Memon, "Draw-a-pin," *Computers and Security*, vol. 66, no. C, pp. 115–128, 2017.
33. J. Sauro, "Measuring usability with the system usability scale (sus)," 2011.
34. R. Ritchie, D. Rubino, K. Michaluk, and P. Nickinson, "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`, 2013. online web resource.
35. 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 Conference on Human Factors in Computing Systems Proceedings of the SIGCHI*, pp. 987–996, ACM, 2012.
36. T. Feng, Z. Liu, K.-A. Kwon, W. Shi, B. Carbunar, Y. Jiang, and N. Nguyen, "Continuous mobile authentication using touchscreen gestures," in *Proceeding of IEEE Conference on Technologies for Homeland Security (HST)*, pp. 451–456, IEEE, 2012.
37. M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, "Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," *IEEE transactions on information forensics and security*, vol. 8, no. 1, pp. 136–148, 2013.
38. N. Sae-Bae, N. Memon, K. Isbister, and K. Ahmed, "Multitouch gesture-based authentication," *IEEE transactions on information forensics and security*, vol. 9, no. 4, pp. 568–582, 2014.
39. H. Xu, Y. Zhou, and M. R. Lyu, "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, 2014.
40. J. Mantyjarvi, M. Lindholm, E. Vildjiounaite, S.-M. Makela, and H. Ailisto, "Identifying users of portable devices from gait pattern with accelerometers," in *Proceedings ofIEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, pp. ii–973, IEEE, 2005.
41. L. Li, X. Zhao, and G. Xue, "Unobservable re-authentication for smartphones.," in *Proceedings of NDSS*, vol. 56, pp. 57–59, 2013.
42. J. Zhu, P. Wu, X. Wang, and J. Zhang, "Sensec: Mobile security through passive sensing," in *Proceedings of International Conference on Computing, Networking and Communications (ICNC)*, pp. 1128–1133, IEEE, 2013.
43. W. Shi, J. Yang, Y. Jiang, F. Yang, and Y. Xiong, "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, 2011.
44. A. Buriro, B. Crispo, F. Del Frari, and K. Wrona, "Touchstroke: smartphone user authentication based on touch-typing biometrics," in *Proceeding of International Conference on Image Analysis and Processing*, pp. 27–34, Springer, 2015.
45. Z. Sitová, J. Šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, and K. S. Balagani, "Hmog: New behavioral biometric features for continuous authentication of smartphone users," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 5, pp. 877–892, 2016.
46. M. Conti, I. Zachia-Zlatea, and B. Crispo, "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, 2011.
47. C. Giuffrida, K. Majdanik, M. Conti, and H. Bos, "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, 2014.

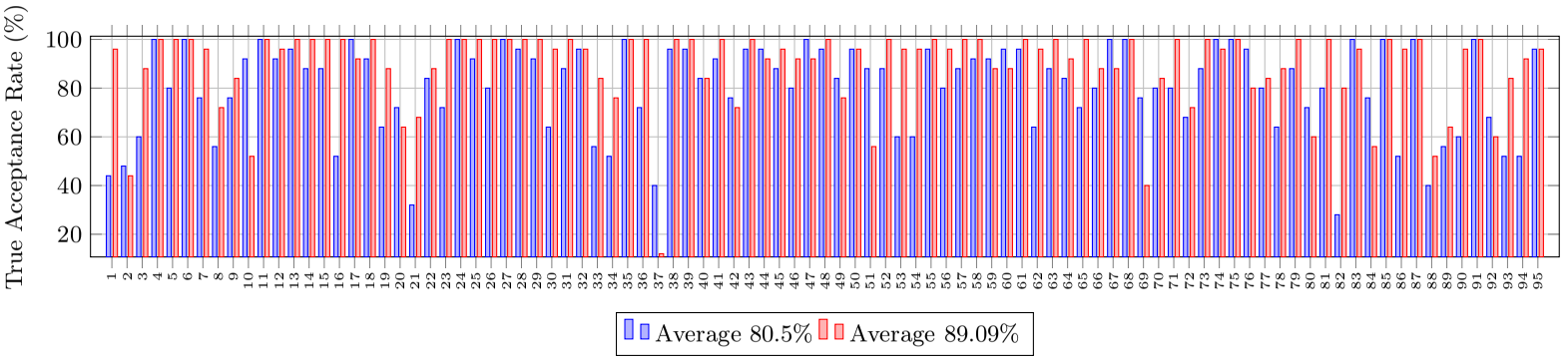
48. B. Schneier, "Risk-based authentication." https://www.schneier.com/blog/archives/2013/11/risk-based_auth.html, 2013. online web resource.
49. M. Butler and R. Butler, "Investigating the possibility to use differentiated authentication based on risk profiling to secure online banking," *Information & Computer Security*, vol. 23, no. 4, pp. 421–434, 2015.
50. I. Traoré and A. A. E. Ahmed, "Introduction to continuous authentication," *Continuous Authentication Using Biometrics: Data, Models, and Metrics: Data, Models, and Metrics*, p. 1, 2011.
51. IBM, "Ibm trusteer." http://www-03.ibm.com/software/products/en/category/advanced-fraud-protection, 2016. online web resource.
52. IBM, "Ibm tivoli federated identity manager." https://www.ibm.com/support/knowledgecenter/en/SSZSXU_6.2.2.7/com.ibm.tivoli.fim.doc_6227/rbaOverview.html, 2016. online web resource.
53. M. Sepczuk and Z. Kotulski, "A new risk-based authentication management model oriented on user's experience," *Computers & Security*, vol. 73, pp. 17–33, 2018.
54. D. Preuveneers and W. Joosen, "Smartauth: dynamic context fingerprinting for continuous user authentication," in *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, pp. 2185–2191, ACM, 2015.
55. D. Hintze, E. Koch, S. Scholz, and R. Mayrhofer, "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, 2016.
56. Y. Y. Haimes, *Risk modeling, assessment, and management*. John Wiley & Sons, 2015.

A Demographic Questionnaire

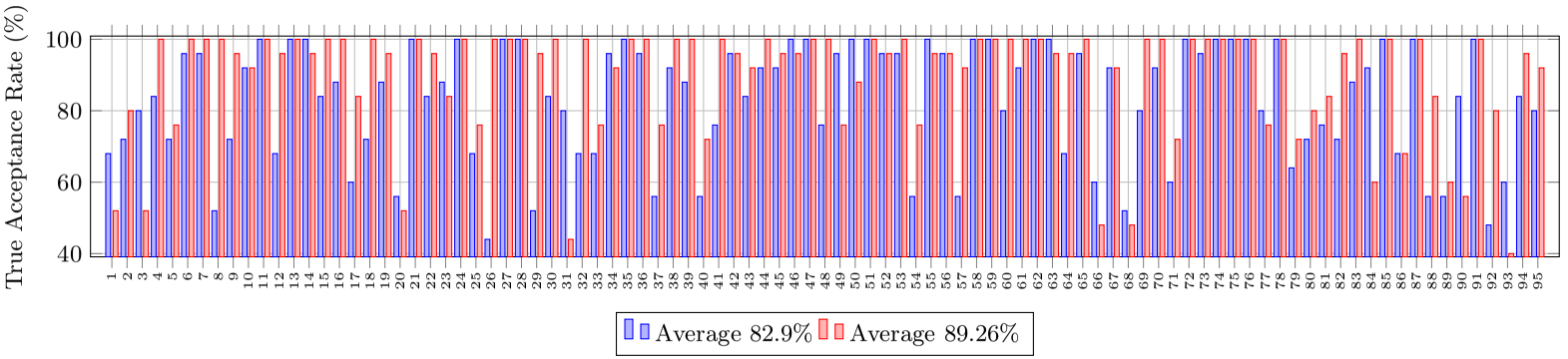
1. What is your gender?
   - Male
   - Female
   - I don't want to disclose
2. How old you are?
   - $\leq$ than 20 years.
   - $>$ 20 years and $\leq$ 40 years.
   - $>$ 40 years and $\leq$ 60 years.
   - $>$ than 60 years.
   - I don't want to disclose
3. Tell us about your nationality.
   - _____
   - I don't want to disclose
4. Which hand(s) do you use for interacting with your smartphone?
   - Right
   - Left
   - Both
   - I don't want to disclose

B TAR comparison of RF classifier for individual users in 3 activities

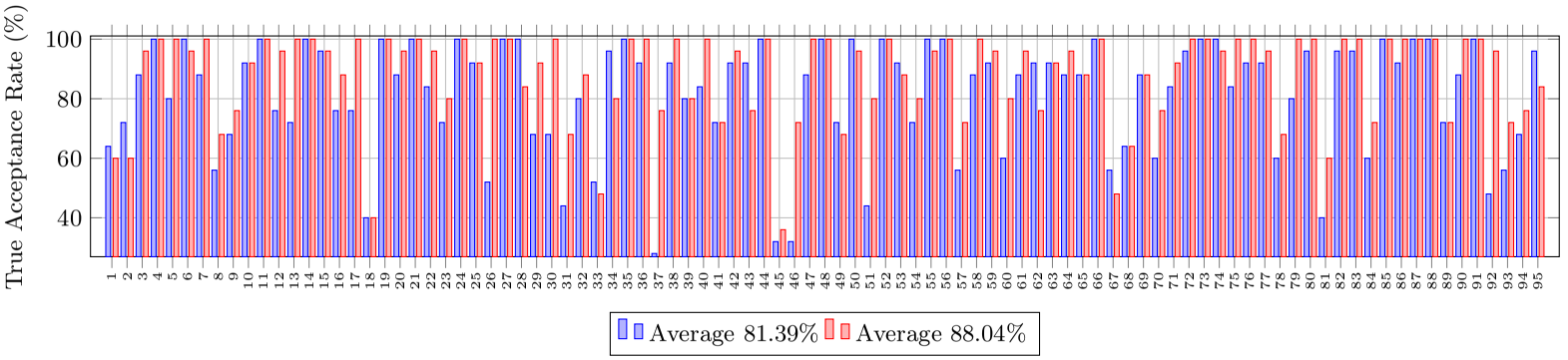
Fig. 6. Comparison of RF classifier performance (TAR) on 5-sample training over full and IGAE features



(a) Sitting Activity



(b) Standing Activity



(c) Walking Activity