EasyChair Preprint

№ 7182

Motion-Sequence Authentication System: Guard for Smart Phones

Yuzheng Dong, Yanfeng Zhao, Ziyue Wang, Juan He and Liubin Zhu

December 7, 2021

Motion-sequence Authentication System: guard for smart phones

Yuzheng Dong[1], Yanfeng Zhao[1], Ziyue Wang[1], Juan He[1](✉) and Liubin Zhu[1]

[1] School of Information Science and Technology, Northwest University, Xi'an, Shaanxi, China
hejuan@stumail.nwu.edu.cn

Abstract. In recent years, the mobile privacy protection is becoming increasingly critical due to the popularity of smartphones. The owner needs a "second line of guard" (user behavior authentication) when unlocking methods are attacked. The traditional user authentication approaches either face smudge attacks or can only work on dedicated devices. In this paper, we present a Motion-sequence Authentication System (MAS), an accurate and robust security authentication system that is not limited to expensive phones. MAS（Motion-sequence Authentication System）distinguishes user categories according to the unique characteristics of different user motion sequences. It is a rapid, non-contact and unobtrusive method of user authentication without predefined motions. MAS exploits Markov model to track the behavior of smartphone users, it can achieve real-time user authentication by utilizing the user's short-term interaction with the smartphone. Our experiments in multiple real environments show that MAS can achieve higher than 94% accuracy for authenticating user motion sequences, which fills the gap with motion sequence recognition and provides a way of thinking for the development of human-computer interaction and information security.
Keywords: Human-computer interaction, Behavior recognition, Multi-scenario security authentication, Markov model

1 Introduction

With the rich functionalities and enhanced computing capabilities available on smart phones, users not only store sensitive information, but also use privacy-sensitive applications on smart phones. Consider that the smart phone can be accessed by many people other than the owner, such as be shared with families and friends, be stolen by a thief, or be peeped by an attacker, smart phone owners face increasing risk of privacy leakage [1]. If the smartphone can distinguish the owner and attacker during their accessing, it would be the "second line of guard" for owners.

The existing authentication methods fall short on one or more of the following aspects. The widely adopted PIN/pattern password is inherently vulnerable to shoulder surfing attacks and smudge attacks [2, 10]. Methods of the highest security are using the physiological biometrics of different persons (e.g., fingerprint and face) [14, 17, 18]. However, these solutions require dedicate hardware, such as fingerprinting scanner. A recent alternative is to use the behavior biometrics (e.g., hand waving, gait, touch-based biometrics) [3, 5], which however either incurs high delay (e.g., gait and touch

behavior-based methods have to collect more than 10s motion data for accurate authentication [16]) or impose additional burden to users (i.e., perform a defined motion [3]). Moreover, the smart phone owners may not be willing to take distrust action to reduce permission deliberately before sharing. Besides, the above solutions cannot distinguish the attackers and the friendly users, providing them with the same access permission [4]. However, different from the attacker, the owners may allow the friendly user to access the insensitive applications (such as the mobile games, browser, YouTube, etc.) on the smart phone, but do not want them to access sensitive applications.

Under this circumstance, it would be good for a smart phone to identify who is the current user (the owner, a friendly user or an attacker) instantly and inconspicuously, as well as provide specific and targeted privacy protection and access control automatically for each role. We meet many challenges to this version. First, without relying on a pre-defined distinguishable behavior, we have to identify three kinds of users only based on users interacting with the smart phone. Consider that different people (especially the friendly user and the attacker) exhibit neither consistent nor distinguishing behavior when they interact with the smart phone, it's hard to find features that can constantly identify the users with high accuracy. Second, a great challenge comes from the low-latency requirement. This means that we have to identify the user instantly based on their short-term behaviors, which lasts for short time.

In this paper, we for the first time propose a multi-user authentication system for smart phones, named MAS. MAS is able to instantly identify whether the current user is the owner, a friendly user or an attacker based on only the 2s key behavior, i.e., pick up and unlock the smart phone of the user. The idea behind MAS comes from the following intuitive observations:

i) Different users have the distinctive habit (or motion sequence) when he/she performs the key behavior;

ii) A person has distinctive features when he/she performs each independent motion in the key behavior. Obtaining the abovementioned fine-grained information in real time, however, is a very challenging task. Based on a joint consideration of the user's motion patterns and the order of motions, we provide a model based on Markov chain, which continuously track motions of the smart phone user, and instantly estimate how likely the motions are performed by the smart phone owner, a friendly user, or an attacker.

This paper also provides details of designing and implementing such a system. Specifically, we design a motion segmentation algorithm to detect the transition between two motions from the noisy sensing data. Then we leverage the distinct feature contained in each sub-segment of the unlocking motion, instead of the entire motion, to estimate the probability that the unlocking motion is performed by the smart phone owner himself/herself. At the same time, under the influence of Covid-19, wearing masks have become a necessary means of protection, which limits the application range of face recognition to a certain extent. MAS can identify users independently of facial features, and does not need to take off the mask during detection, so it has great application potential in air defense work of Covid-19 epidemic situation.

Finally, we summarize the following contributions:

i) We observe the distinctive accessing habit of different types of smartphone users, and propose a model based on Markov chain to continuously track motions of the smartphone user, which can identify different users instantly and accurately.

ii) Based on the proposed model, we introduce MAS, the first system which can distinguish whether the current user is the owner or a friendly user or an attacker in real time, and then provides necessary privacy protection and access control based on the identification result.

iii) We implement MAS on several platforms (including Samsung S4, m1 mental, MI 2s, and OPPO r9s). The extensive experiments demonstrate the accuracy and robustness of MAS.

2 Related Work

Parallel work on identification can be classified into three categories: passwords/PINs/patterns, physiological biometrics and behavioral biometrics. However, passwords/PINs/patterns are inherently not security, since malicious users can unlock the smartphone by peeping attacks or smudge attacks. Physiological biometrics need extra hardware and also can be spoofed. Moreover, these solutions fail to achieve a goal of distinguishing multi-users.

The two key technologies in behavioral biometrics domain are feature based and similarity based.

Feature based: Feature based recognition approaches mainly exploit the distinctive features among different behaviors, extracting features of a behavior to establish a classifier [23–25]. For example, GEAT [3] is a gesture-based user authentication system, which extracts behavioral-related feature of predesigned 10 sliding gestures and uses a SVDE classifier. It achieves an average equal error rate of 0.5% to identify legitimate and illegitimate for each gesture. Touchalytics [5] presents a continuous authentication scheme that leverages features of a swipe. An evaluation of Touchalytics shows that with the SVM or the KNN classifier it provides an EER of 4% to distinguish owner and other people. LXG [6] devices following 9 features of users' finger movements during a swipe gesture and uses an SVM classifier to check the current behaviors of user against the owner's, which achieve an equal error rate (EER) of 8%.

They perform a high accuracy under the assumption that different behaviors have distinguishable features, which is not hold for the multi-user identification scenario, where the motions performed by different people are very similar, e.g., take out the smartphone. A pre-designed behavior to capture distinctive feature may impose additional burden to users and show distrust to another user.

Similarity based: Similarity based behavior recognition methods primarily maintain a group of well-defined behavior profiles. They distinguish behaviors based on similarity metrics (e.g., DTW and EMD) to evaluate the similarity between the sampled signals and the pre-constructed behavior. For example, GTGF [9] proposes a continuous mobile authentication that converts the touch traces to images and computes the score between two users' image, it achieves an EER of 2.62%. Sae-Bae et al. [13] provides a user authentication system based on 22 designed touch gestures that compute DTW

distance and Frechet distance between users' traces to authenticate legal and illegal user. Luca et al. [8] proposes a user authentication system that directly computes the distance between touch traces using DTW algorithm.

However, it may perform low accuracy in such an identification scenario due to the variability of both the user and smartphone scenario, even the same motion (e.g., take-out smartphone) performed by the same user rarely have a fixed pattern [12].

Different from all past work, MAS combines users' order of using smartphone motions with independent motion patterns to authenticate owner, friendly user or attacker implicitly and instantly.

3 Preliminary

It may perform low accuracy in such an authentication scenario due to the variability of both the user and smart phone scenario, even the same motion (e.g., take-out smart phone) performed by the same user rarely have a fixed pattern [12]. The user behavior is short in duration and has little information to refer to. Unlike previous work, MAS combines a user's smart phone action sequence with independent action patterns to implicitly and instantly verifies an owner, user-friendly or attacker.

3.1 User Classification This Style for Level Two

The users are divided into three categories:

i) Owner: The owner of the smart phone device with using his/her smart phone freely.

ii) Friendly users: Smart phone owners share their device to them, but hiding the sensitive information unobtrusively.

iii) Attackers: The users who use the smart phone without owner's permission, the smart phone system does not provide any access permission with them.

In the first experiment, we observe how owners use their smart phones. The experiment lasts for 10 days and collects 489 samples for the key behavior of owner. The experiments are conducted in two scenarios: static scenario and dynamic scenario. In the second experiment, we firstly ask the owner to share their smart phones to 31 volunteers and then ask 20 volunteers to borrow a smart phone from 4 owners. We observe the key behavior of owners and borrowers respectively, collecting 489 samples. In the third experiment, we ask 20 volunteers to pretend to be attackers to 'peep' privacy without being noticed. We collect 40 samples for the key behavior of attackers when 'stealing' sensitive information. Some samples, for example, user picks up the smart phone, but does not unlock or use the smart phone, are invalid samples, and we throw them away.

3.2 Data Analysis

In this subsection, we will further analyze the smart phone use behavior of the owner, friendly users and attackers. We abstract the user's key behavior into a motion sequence

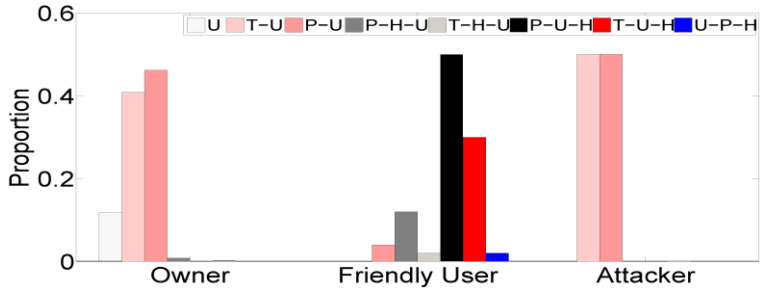
composed of several independent motions. We found that users' independent motion mainly include:

i) Pick-up: picking up the phone from the desktop or other places;

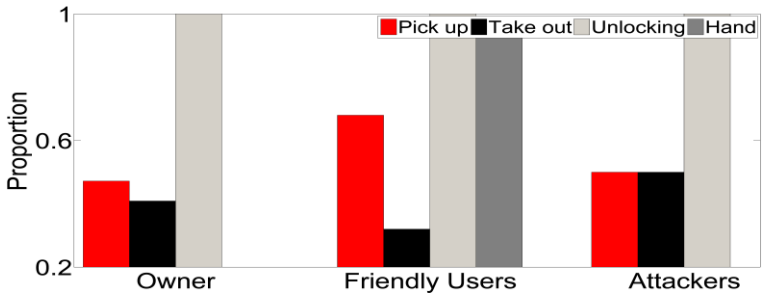
ii) Take-out: taking the phone out of the bag or pocket;

iii) Unlocking: unlocking the smart phone;

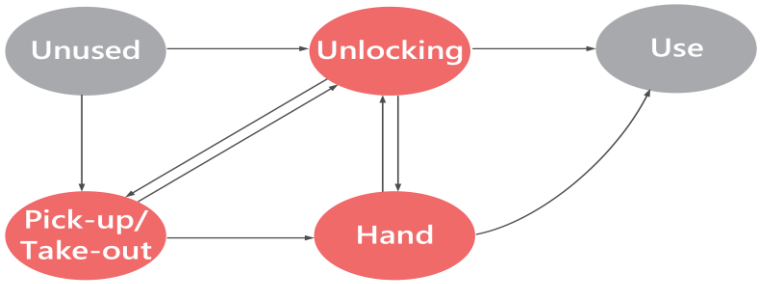
iv) Hand: handing the phone to others. They will form different motion sequences in different usage scenarios.



(a)



（b）



(c)

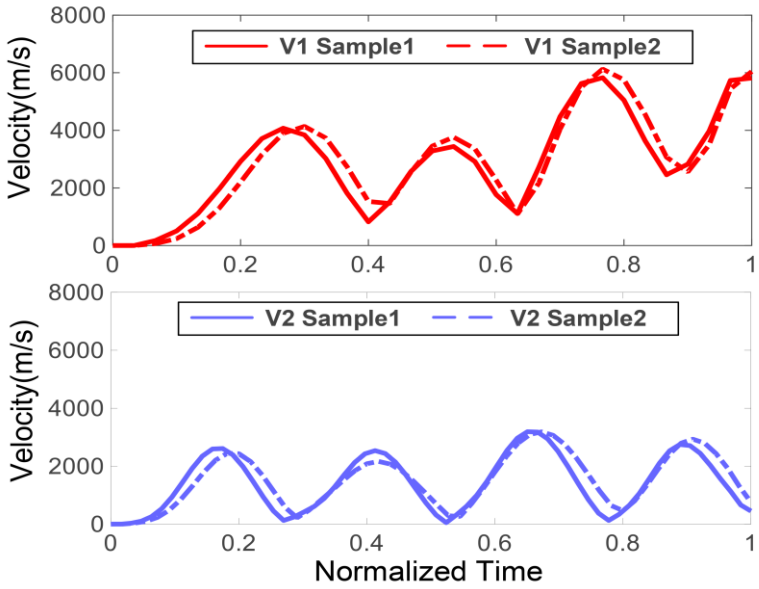
Fig. 1. Motion sequence of key behavior performed by different user: (a) probability of motion sequence in different users; U: Unlocking motion, T: Take-out motion; P: Pick-up motion, H: Hand motion (b) probability of each independent motion during a key behavior; (c) Motion sequence recognition process.

Motion Sequence of the Key Behavior Performed by Different User. In fig. 1(a), we plot the statistical results of different users' habits (or motion sequence) during the key behavior. We observed that, when the owner uses the smart phone, the sampled motion sequences are {unlocking} or {pick up/take out, unlocking}. However, when the friendly users use the smart phone, the sampled motion sequences are {pick up/take out, unlocking, hand}, {take out/pick up, hand, unlocking}. When the attackers use the smart phone, the sampled motion sequences are {pick up/take out, unlocking}, which a hand motion will not appear in the key behavior of attacker. Also, we show the transitions between each motion in fig. 1(c), including eight motion sequences which is performed by owner, friendly users or attackers. The main cause behind such a difference is:

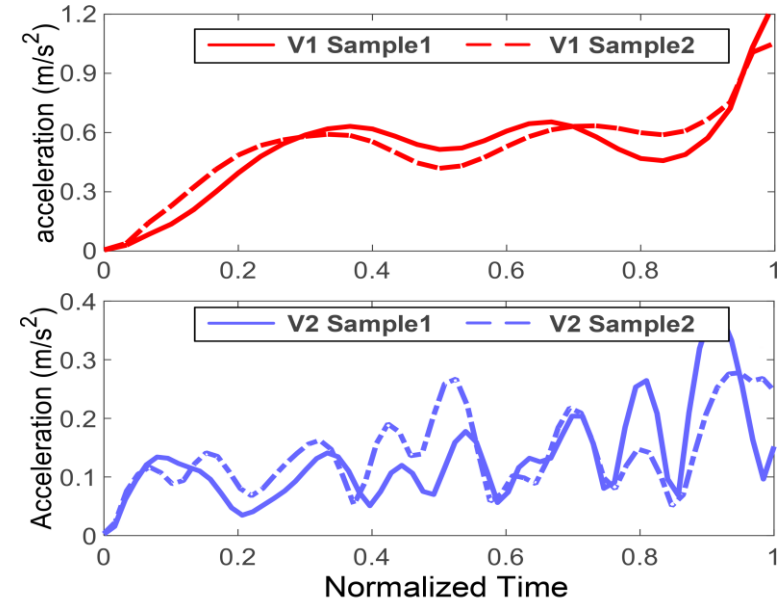
i) Smart phone owners share their device to friendly users, usually accompanied by a handing motion.

ii) Attackers peep private information of smart phone sneakily, they may steal the smart phone from owner's pocket, desk, etc., with a take-out or pick-up motion.

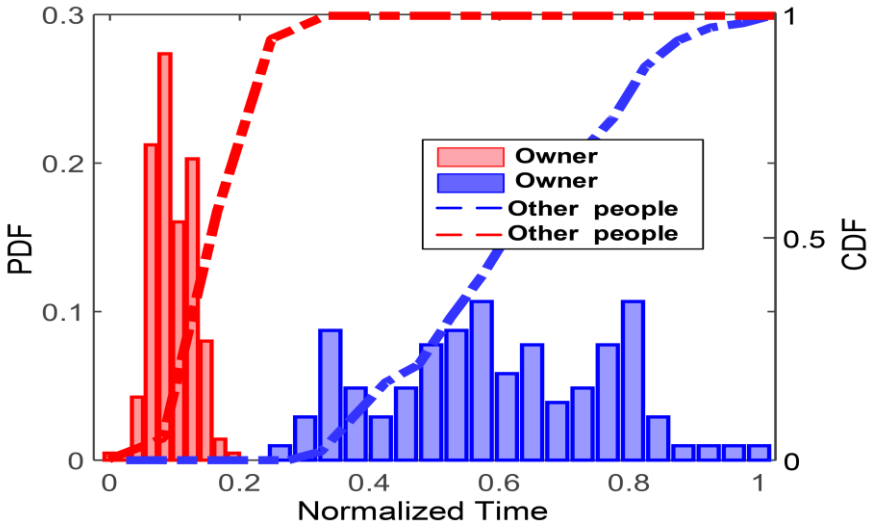
In fig 1(b), we then show the statistical results of independent motion in the habit of users. With the above observations, we can find that different users have distinctive habit when performing the key behavior.



(a)



(b)



(c)

Fig. 2. The macro button chooses the correct format automatically. The unlocking motion performed by different people: (a) time series of velocity sampled during unlocking motion; (b) time series of acceleration sampled during unlocking motion; (c) distribution of normalized time distance between different users (e.g., owner and other people). PDF： Probability Density Function.

Analysis of Independent Motion Performed by Different User. In this subsection, we analyze the differentiation and consistency features of different people when performing an unlocking motion. As shown in fig. 2(a) and fig. 2(b), we plot the time series of velocity and acceleration sampled of unlocking motion performed by different users. Specifically, at the top of the figures show the velocity and acceleration data when two unlocking motions performed by the same owner, at the bottom of the figures show the velocity and acceleration data during two unlocking motions performed by another person. We can see that the pattern of unlocking motion is very similar when performed by the same owner, while different from that performed by another non-owner.

Fig. 2(c) shows the distribution of normalized duration time when unlocking motions performed by the same owner and other people, which shows that the duration time of owner is smaller than other people and overlap is very small.

The main reason behind this difference is:

i) The owner is more familiar with their password. As a consequence, showing a faster velocity, shorter time, more stable and greater acceleration;

ii) The duration time of other people is more decentralized.

The above observation implies that a person has consistent and distinguishing features when she/he performs unlocking motion in the key behavior.

4 Overview

Based on the observation in Section 3, we propose MAS, a multi-user identification detection system that is able to instantly and inconspicuously identifying who is accessing the smartphone (the owner, a friendly user or an attacker) using only the inertial sensors. Figure 4 shows the overview of MAS. To distinguish different users, we go through the following four steps:

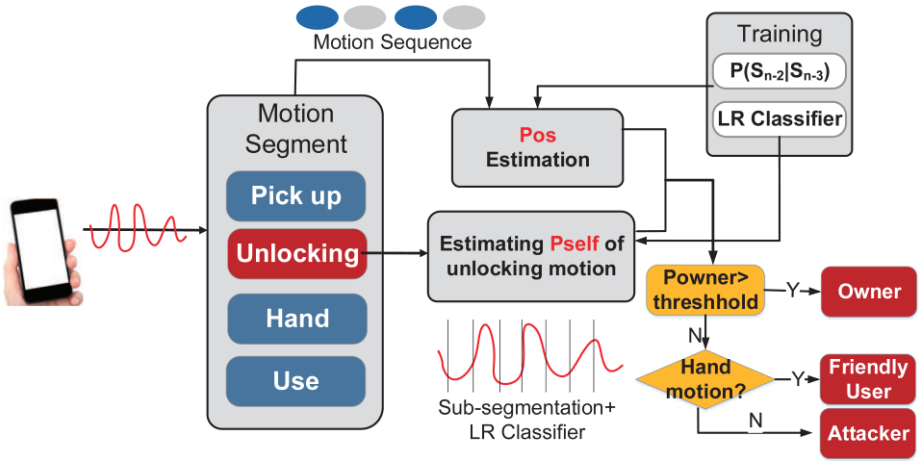


Fig. 3. Overview of MAS(Motion-sequence Authentication System)

- Pre-processing: the sampled sensor data are processed via a low-pass filter and a sliding mean filter to remove high frequency noises.
- Motion sequence detection and segmentation: in this component, MAS segments the smoothed sensor data into a sequence of sub-segments, where each segment contains an independent motion S_i. The output of this component is a motion sequence S_1, S_w.
- P_{self} estimation: after obtaining the motion sequence, we take the unlocking segment as the input of P_{self} estimation component, which evaluate how likely the unlocking motion is performed by owner.
- User authentication: We take a joint consideration of both the order of the motions and the P_{self} of the unlocking motion as input to distinguish the owner, the friendly user and the attacker.

5 System Design

5.1 A Model for User Authentication

In this paper, we present a Markov-based model to address this problem. When the user performs a key behavior (denoted as S_n). After segment, we get a motion sequence $S_n^{(p)} = \{S_{n-1}, \dots, S_{n-w}\}$,where w is the number of independent motions in the key behavior. Assume that the transition probability between S_i and S_{i-1} is $P_{(S_i|S_{i-1})}$. The probability that user perform the key behavior in the order of $\{S_{n-1}, \dots, S_{n-w}\}$ can be calculated as:

$$P_{seq} = \prod_{i=n-w}^{n} P(S_i|S_{i-1}, S_{i-2})$$

We use two metrics, i.e., P_a and P_b to evaluate the probabilities that the 8 motion sequences (which is shown in fig. 1(c)) performed by the owner.

Take-out motion detection: The difference in magnitude of take-out motion, step on a stair and hand motion. Himself/herself and by other people, respectively. Specifically, we have where n is the number of possible motion sequences. Specifically, the larger gap between P_a and P_b indicates a higher discrimination between the order of motion that performed by different users. We propose a metric (denoted as C_{of}) to describe the confidence of using P_a and P_b for user authentication. It can be defined as follows:

$$P_a = (P_{a_{seq_1}}, P_{a_{seq_2}}, \dots, P_{a_{seq_n}}), \quad P_b = (P_{b_{seq_1}}, P_{b_{seq_2}}, \dots, P_{b_{seq_n}}) \quad (1)$$

$$C_{of}(seq_i) = \max\{\frac{rank(a_{seq_i}) - rank(b_{seq_i})}{|S|}, \quad 1 - \frac{\min(P_{a_{seq_i}}, P_{b_{seq_i}})}{\max(P_{a_{seq_i}}, P_{b_{seq_i}})}\} \quad (2)$$

where $S = \{seq_1, \dots, seq_n\}$ is the set of all motion sequences, $|S|$ is the number of the motion sequences. Rank (a_{seq_i}) and rank (b_{seq_i}) are the ranking of P_{seq} among all the P_a and P_b, respectively (we have $a_{seq_i} \in a, b_{seq_i} \in b, a, b \in S$). At last, the probability that the seq_i is performed by the owner can be calculated as:

$$P_{os}(seq_i) = \alpha P_{ai} + (1 - \alpha) C_{of}(seq_i) \quad (3)$$

Where α is the weighing coefficient, which is set as 0.7 in our implementation.

We denote the probability that the unlocking motion performed by the owner as P_{self} (the method to calculate P_{self} is given in Section 3.3). Based on $P_{os}(seq_i)$ and P_{self}, we

propose a normalized metric to evaluate how likely the detected motion sequence $S_{current}$ is performed by the owner as follows:

$$P_{owner} = P_{os}(S_{current}) \times P_{self}. \quad (4)$$

Clearly, a high P_{owner}, namely a high probability that the smart phone is using by the user himself/herself, is achievable under the following two conditions:

i) The motion sequence $S_{current}$ comparisons with the habit of the owner, leading to a high P_{os};

ii) The unlocking motion exhibits high similarity with that performed by the owner himself/herself, leading to a high P_{self}. We will discuss the feasibility of the above proposed model in section 2.3.

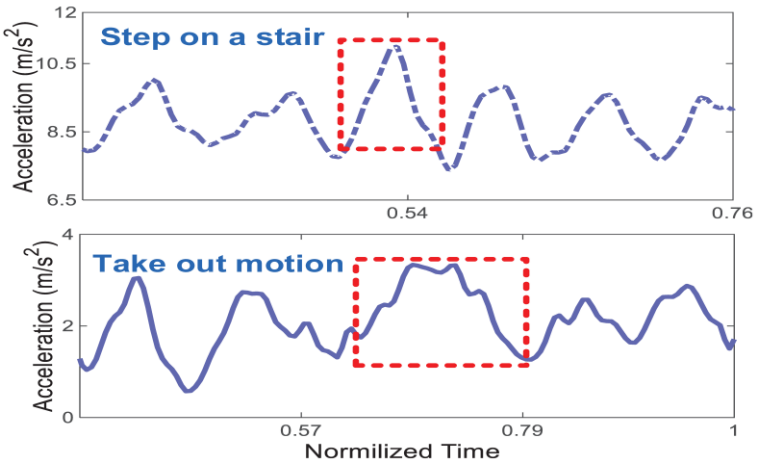
5.2 Motion Detection and Segmentation

The goal of the motion detection and segmentation component is to extract the pick up, take out, hand or unlocking motion, from the smoothed sensor data, and output a motion sequence $\{S_{n-1}, \ldots, S_{n-w}\}$.

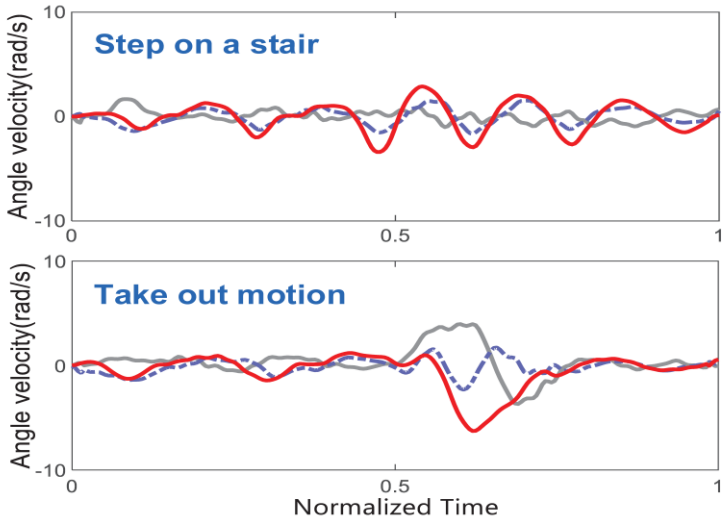
Pick-up/Take-out Detection. The pick-up motion refers to the motion that the user picks up a smart phone from a desk. The take-out motion refers to the motion that the user takes out a smart phone from his/her pocket or handbag.

We consider the pick-up and take-out motions as the same motion, since the only difference between them is the initial attitude of the smart phone. One naive solution to detect the pick-up/take-out motions is to detect the sudden change of the acceleration data with a predefined threshold. However, some other motions, such as step on a stair, might have a similar impact on the acceleration data, as shown in fig. 4(a).

Our idea to solve this problem is to use the gyroscope. Specifically, comparing with pick-up/take-out motion, fig. 4(b) shows that the motion of step on a stair has a marginal impact on the gyroscope data. However, we meet another problem that the handing motion has a similar impact on gyroscope data. Fortunately, we find that different from the handing motion, the pick-up/take-out motion will lead to the changes in the altitude of a smart phone.



(a)



(b)

Fig. 4. The difference between take-out motion and step on a stair: (a) accelerometer data; (b) gyroscope data.

Unlocking Detection. To unlock a smart phone, the user usually first lights the touch screen, then enters a password/pattern to unlock the smart phone. We use the built-in API of the smart phone to detect this motion.

Handing Detection. To detect the handing motion, one potential solution is to capture the sudden change in gyroscope data. However, pick-up/take-out motion would incur

similar change in the gyroscope data, as shown in fig. 4. Fortunately, when we observe the components of the gyroscope data on different axis, we find that different from the pick-up/take-out motion which incurs significant changes in all the three axes, the handing motion only incurs changes in the Z-axis.

In MAS, we take both the changes in Y-axis and Z-axis of gyroscope sensor data into account, which detect the hand motion with two predefined thresholds（threshold 3 and threshold 4）. When a handing motion is performed at the time of t with coming the new data (gyroscope sensor data on y-axis w_y^t and gyroscope sensor data on z-axis w_z^t), the algorithm collects the samples w_y^m and w_z^m within last T second to compare with w_y^t and w_z^t find the maximal difference. If $|w_z^t - w_z^m| < threshold3$ and $|w_y^t - w_y^m| < threshold4$, we identify the t as the start of handing motion.

After identifying individual motion during performing using behavior, MAS obtains a motion sequence which might be {take-out, unlocking, hand}. In addition, we take independent motion as the input of P_{self} estimation component for further analysis.

5.3 A Subsection Sample

The target of P_{self} estimation component is to evaluate how likely the independent motion is performed by the owner. With the analysis in section 3, we find that a person has distinctive features (e.g., velocity, acceleration and time duration) when performing unlocking motion.

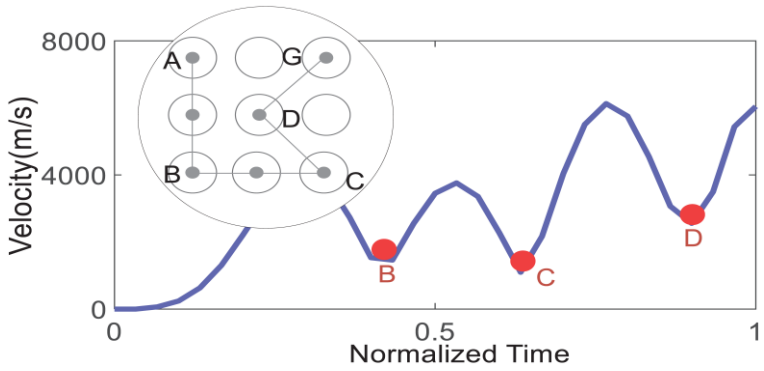


Fig.5. P_{self} Estimation

To obtain above fine-grained information, we find that, the distinguishable features may show at different sub-segment of an unlocking gesture. Thus, compare with extracting feature from the entire motion, we extract features from sub-segment of unlocking motion have higher differentiation degree. However, how many numbers of sub-segments should we segment an unlocking motion is a challenging task. First, too short time duration of a sub-segment may hide the consistent behavior. Second, too large time duration of a sub-segment may average out the distinctive information from the feature, failing to identify different users.

As shown in fig. 5, we show an example of lock pattern and its corresponding time series of velocity sampled when performing such a gesture. At the key point, such as B, C and D point, fig. 5 shows that they cause obvious changes in speed and direction. We propose an algorithm that we first leverage the key points to segment unlocking motion into sub-segment, such as {AB, BC, CD, CE}, then calculate P_{self} of entire unlocking motion based on the feature from each sub-stroke and use a logistic regression classifier.

To detect the key point, we leverage the minimum speed which typically indicate a key point. At the time of t, we assume that the speed of P point is v_t, the algorithm traces back to get the mean speed v_m within last T second, and trace forward to colorredget the mean speed vn within T seconds. With a given threshold, if $v_t - v_m <$ Threshold and $v_t - v_n <$ Threshold, we identify the P point as the key point.

We segment the unlocking motion based on the detected key point. We then extract features from sub- segment. Then, we exploit the logic regression algorithm to calculate P_{self}. The main reason for choosing such an algorithm is that:

i) its output ranges from 0 to 1 which is meet with the normalized result of P_{self};

ii) Logic regression algorithm with computational simplicity (O(n)) helps us detect the motion instantly

5.4 A Model for User Authentication

When the new data comes, we first remove high frequency noise in data using a low-pass filter. Then,

i) we can get the current motion sequence $S_{current} = \{S_{n-w}, ..., S_{n-1}\}$ based on the motion detection and segmentation (in section 3.2);

ii) Estimating the P_{self} of its unlocking motion by extracting feature from sub-segment and using logic regression algorithm. After getting above information, MAS estimates the P_{owner} using Equation. 4.

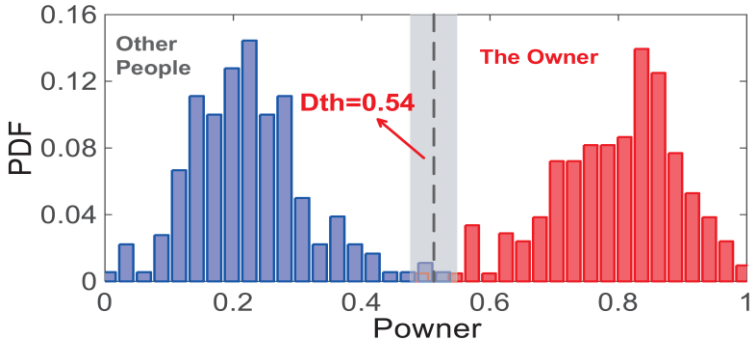
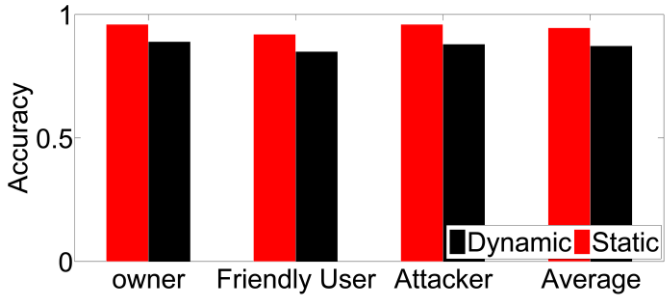
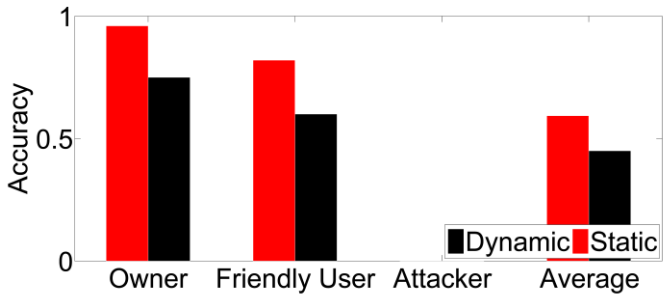


Fig. 6. Distributions of P_{owner} for the motion sequences of key behavior performed by the user and other people.

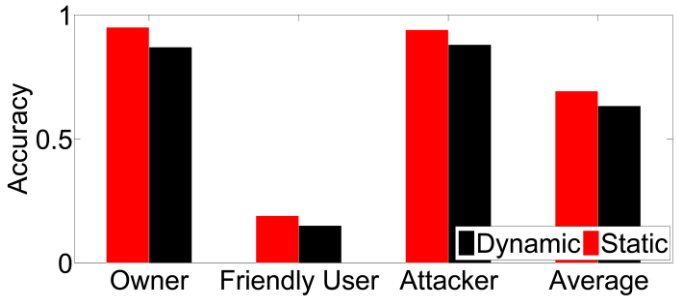
Fig. 6 shows the distribution of P_{owner} for the motion sequences of using smart phone performed by the owner himself/herself and other people. The result shows that the overlap is very small. As a consequence, we set the threshold $D_{th} = 0.54$, which $P_{owner} > D_{th}$ indicates the owner. Further, if $P_{owner} < D_{th}$ and a "and motion" is detected, MAS identifies it as friendly user.



(a)



(b)



(c)

Fig. 7. Performance under different condition: (a) Accuracy of MAS; (b) Accuracy of P_{os}; (c) Accuracy of P_{self}.

6 Performance Evaluation

In this section, we first explain how MAS trains its logic regression classifier and the transition probability for Markov-based model. Then, we conduct a set of experiments in several scenarios to demonstrate the accuracy and robustness of MAS.

6.1 Training

Training the Logic Regression Classifier. With the observation in section 2.3, we find that the unlocking motion performed by the same role is very similar and different from that performed by another role. Thus, we can only train a classifier for the same role, rather than each person. During the training progress, we ask the owner to unlock his smart phone for 100 runs, and 30 non-owner volunteers (including friendly users and attackers) to unlock the owner's smart phone for 120 runs. Then, we use these collected sensor data to train the logic regression classifier.

Training the Transition Probability. We obtain the transition probability based on the habit of user in real world. We found 50 volunteers with occupation ranging from students, faculty, to company staffs and age ranging from 18 to 55 to do data collection, which records the key behavior and lasts for three weeks. The key behaviors including the behavior of owner, friendly users and attackers.

During this period, our goal is to get the transition probability, there is no privacy concern. Finally, based on the collected motion sequence, MAS can infer the transition probabilities between each two independent motions.

6.2 Experiment Setup

We conduct extensive experiment on four different types of smart phones (including Samsung S4, m1 mental, MI 2s, and OPPO r9s, which are equipped with gyroscopes, accelerometers and gravity sensors) to validate the effectiveness of MAS. The experiments are conducted under two scenarios, including static scenario (e.g., sitting and standing) and dynamic scenario (e.g., walking). The sampling rate is set as 50Hz. There are 7 volunteers in our performance evaluation.

The performance metrics in this paper are as follows:

i) True Positive Rate (TPR): the fraction of cases where MAS correctly recognizes the other people;

ii) False Positive Rate (FPR): the fraction of cases where MAS mistakenly recognizes the owner of other people;

iii) Accuracy: the fraction of cases where MAS correctly recognizes owner, friendly user or attacker respectively.

In order to evaluate the performance of MAS, we implement the following approaches for comparison:

- MAS: the system in this paper.
- P_{os}: identify the user by exploiting only the P_{os} of motion sequence.

- P_{self}: identify the user by exploiting only the P_{self} of independent unlocking motion performed by the user.

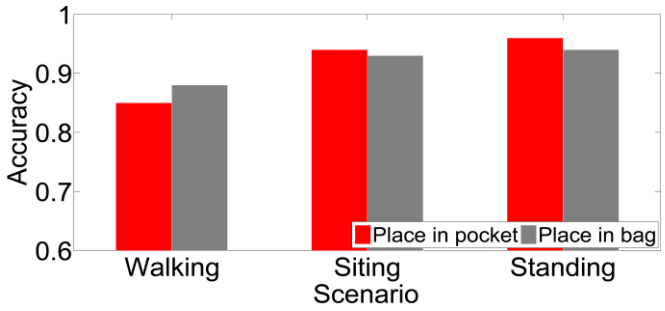
6.3 Accuracy of MAS

To evaluate the accuracy of MAS, we conduct three experiments in an office, in which users are sitting or standing (a static scenario). In the first experiment, we ask an owner to use his/her smart phone himself/herself for 36 runs whenever he/she wanted; In the second experiment, three of volunteers is required to borrow the smart phone from the owner for 36 runs. To imitate motions of the real attacker, three volunteers pretend to be attackers and try to use the owner's smart phone when the owner leaves or does not pay attention to. The results are shown in fig. 9.

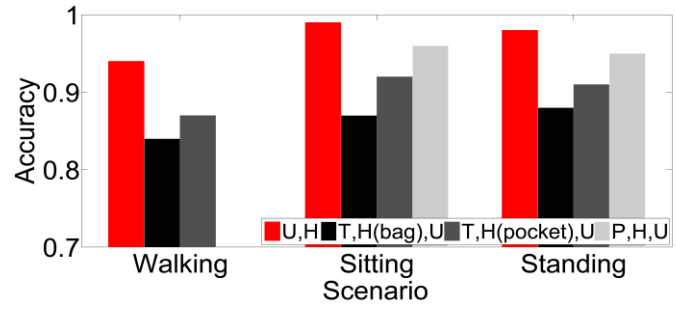
The results have shown that in the relatively ideal condition (a static scenario), when MAS distinguishes between owner and other people, the TPR of MAS can be as high as 95%, and the FPR is lower than 4%. Compared to LXG（ identify the user based on the features of users' finger movements when people interact with touch screen.） only identifies owner and attacker, MAS identifies multi-users (e.g., owner, friendly users and attackers) with high accuracy, as shown in fig. 7. Meanwhile, the FPR of P_{self} or P_{os} can be as high as 57% and 40% respectively, which means that MAS has a good ability to identify user (friendly user and attacker).

6.4 Robustness of MAS

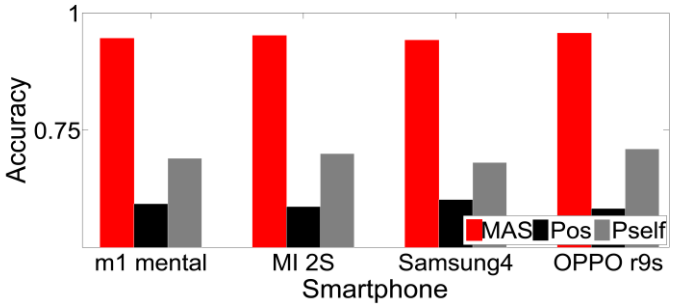
Performance in Identifying Each Independent Motion. Our results show a high accuracy in identifying pick-up motion, take-out motion and hand motion. As we know, if MAS mistakenly detects an independent motion in the key behavior (e.g. ,treating a hand motion as pick-up motion), which is likely to identify users incorrectly. Thus , a low motion detection accuracy leads to a low accuracy of MAS.



(a)



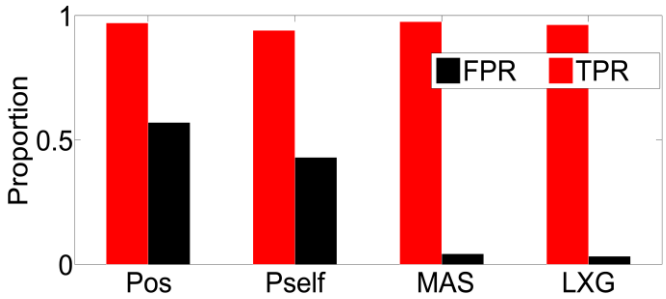
(b)



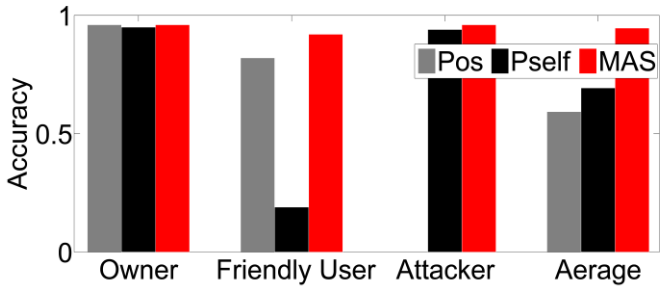
(c)

Fig. 8. (a) Accuracy of MAS in identifying take-out motion; (b) Accuracy of MAS in identifying hand motion; (c) Performance in different smart phone. "U, H": hand motion is performed after unlocking; "T, H(bag), U"/ "T, H(pocket), U": take out the smart phone from bag/pocket, hand, unlock; "P, H, U": pick up, hand, unlocking.

To study the effectiveness of MAS in identifying each independent motion, we ask 7 volunteers to perform the key behavior (including pick-up motion, take-out motion and hand motion) with random order in a static scenario and dynamic scenario respectively. Also, there is no apparent performance degrade of MAS is observed in different scenarios. The accuracy of distinguishing pick-up motion is as high as 98% in statics scenario. When people are in a dynamic scenario, the accuracy is more than 85%.



(a)



(b)

Fig. 9. (a) TPR and FPR comparison of different methods in distinguishing owner and non-owner; (b) Overall accuracy comparison of different methods in identifying multi-user.

Performance in Different Working Conditions. In this section, we evaluate MAS under different conditions (a static scenario and dynamic scenario) to show its robustness. In this section, we have done another experiment. The result is shown in fig. 7. According to the results in fig. 7, no matter the user is in a static scenario or dynamic scenario, MAS outperforms all other methods. Also, the results show that all the three methods perform worse in the dynamic scenario, with 7%, 9%, and 20% performance degradation for MAS, P_{os} and P_{self} method respectively, since the inertial sensor data exhibits larger variance when the user is walking.

Performance in Different Smart phone. We conduct experiments on different types of smart phones, including Samsung S4, m1 mental, MI 2s, and OPPO r9s. The results in fig. 8(c) have shown that MAS is applicable on different scenarios, and is better than P_{os} and P_{self} methods.

7 Conclusion

In this paper, we present a novel smart phone authentication approach, MAS, for non-contract multi-user authentication. In short, different users have different behavioral habits (or sequences of motions) when performing critical motions, and MAS uses a sequence of motions as a unique behavioral biometric. The system is easy to use, unobtrusive, and hard to counterfeit. It can be applied to multiple scenarios. We tested MAS in static scenarios (such as sitting and standing) and dynamic scenarios (walking), and evaluated the performance of seven volunteers. Experimental results showed that the system has certain accuracy and robustness. In particular, MAS was the most stable (only 7% less) in dynamic scenario, compared to P_{os}(9% less) and P_{self}(20% less). Therefore , it is reasonable to believe that MAS provides a novel, high-precision and high-stability viable approach to existing user identification technologies. At present, our work is focused on common scenarios. In the future, we plan to improve MAS's user recognition capability in special scenarios, such as high-risk scenarios (fire scene, underwater, Covid-19 care

unit with facial protection etc.). In addition, we will use other methods to further improve the accuracy of the system, such as using wavelet transform to extract feature parameter. This will be a new exploration in the field of human-computer interaction, but also put forward a new idea for user security authentication method.

References

1. Mobile device security threats, http://searchmobilecomputing.techtarget.com/guides/Mobile-device-protection-and-security-threat-measures/.
2. Ahmad Zairi Zaidi, Chun Yong Chong, Zhe Jin, Rajendran Parthiban, Ali Safaa Sadiq.: Touch-based continuous mobile device authentication: State-of-the-art, challenges and opportunities. Journal of Network and Computer Applications 191(1), 103162 (2021).
3. Torbjørnsen, A., Ribu, L., Rønnevig, M.：Users' acceptability of a mobile application for persons with type 2 diabetes: a qualitative study. BMC Health Serv Res 19（1）, 641 (2019).
4. M. Jin, Y. He, D. Fang, X. Chen, X. Meng and T. Xing.: iGuard: A Real-Time Anti-Theft System for Smartphones. IEEE Transactions on Mobile Computing 17(10), 2307–2320(2018).
5. Yongtuo Zhang, Wen Hu, Weitao Xu, Chun Tung Chou, Jiankun Hu.：Continuous Authentication Using Eye Movement Response of Implicit Visual Stimuli. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies1(1), 1-15(2018).
6. Bevan C., Fraser D. S..: Different strokes for different folks? Revealing the physical characteristics of smartphone users from their swipe gestures. International Journal of Human-Computer Studies88(12), 51-61(2016).
7. Xiao Wang, Tong Yu, Ole Mengshoel, and Patrick Tague.: Towards continuous and passive authentication across mobile devices: an empirical study. In: 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec' 17). Association for Computing Machinery, pp. 35–45, New York, NY, USA(2017).
8. Alghamdi S.J, Elrefaei L.A..: Dynamic Authentication of Smartphone Users Based on Touchscreen Gestures. Arabian Journal for Science & Engineering43(2),789-810(2018).
9. Ferrag M.A., Maglaras L., Derhab A., et al.: Authentication schemes for smart mobile devices: threat models, countermeasures, and open research issues. Telecommunication Systems: Modelling, Analysis, Design and Management8(6), 73(2020).
10. Meng, W., Li, W. & Wong, D.S.: Enhancing touch behavioral authentication via cost-based intelligent mechanism on smartphones77(2), 30167–30185 (2018).
11. Cao H , Chang K . Nonintrusive Smartphone User Verification Using Anonymized Multimodal Data. IEEE Transactions on Knowledge & Data Engineering1(1), 1(2018).
12. Nguyen T., Roy A. , N. Memon.: Kid on the phone! Toward automatic detection of children on mobile devices. Computers & Security8(4), 334-348(2019).
13. Alzubaidi A., Kalita J..: Authentication of Smartphone Users Using Behavioral Biometrics. Human-Computer Interaction46(12), 256(2019).
14. C. Song, A. Wang, K. Ren and W. Xu.: EyeVeri: A secure and usable approach for smartphone user authentication. In: 35th Annual IEEE International Conference on Computer Communications, pp. 1–9. IEEE INFOCOM 2016, San Francisco, CA, USA(2016).
15. Y. Guo, L. Yang, X. Ding, J. Han and Y. Liu.: OpenSesame: Unlocking smart phone through handshaking biometrics. In: IEEE INFOCOM, pp. 365–369. IEEE, Turin, Italy (2013).
16. Zou Q , Wang Y , Wang Q , et al.: Deep Learning-Based Gait Recognition Using Smartphones in the Wild. IEEE Transactions on Information Forensics and Security1(1),99(2020).

17. Raghavendra, R, C. Busch, and B. Yang.: Scaling-robust fingerprint verification with smartphone camera in real-life scenarios. In: IEEE 6th International Conference on Biometrics: Theory, Applications and Systems, pp. 1–8. IEEE, (2013).
18. Viebke A., Memeti S., Pllana S., et al.: CHAOS: a parallelization scheme for training convolutional neural networks on Intel Xeon Phi[J]. Journal of Supercomputing75(1), 197-227(2019).
19. Xu X , Yu J , Chen Y , et al.: TouchPass: towards behavior-irrelevant ontouch user authentication on smartphones leveraging vibrations. In: (MobiCom '20)The 26th Annual International Conference on Mobile Computing and Networking. (2020).
20. Izumoto D , Yamazaki Y .: Security enhancement for touch panel based user authentication on smartphones. In: 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE, (2020).
21. Saini B S , Singh P , Nayyar A , et al.: A Three-Step Authentication Model.
22. B. S. Saini et al.: A Three-Step Authentication Model for Mobile Phone User Using Keystroke Dynamics. IEEE Access 8(1), 125909–125922 (2020).
23. Kang, Taeho & JI, Sangwoo & JEONG, Hayoung & ZHU, Bin & Kim, Jong.: WearAuth: Wristwear-Assisted User Authentication for Smartphones Using Wavelet-Based Multi-Resolution Analysis. IEICE Transactions on Information and Systems12(1), 1976–1992 (2019).
24. Kim S J , Kim J M , Jo I J .: Multimedia image data processing on smartphone for authentication. Multimedia Tools and Applications 78(5), 5287–5303 (2019).
25. Liu X , Shen C , Chen Y .: Multi-source Interactive Behavior Analysis for Continuous User Authentication on Smartphones. In: 13th Chinese Conference, pp. 669–677. Urumqi, China (2018).
26. A. Acien, A. Morales, R. Vera-Rodriguez and J. Fierrez.: Smartphone Sensors for Modeling Human-Computer Interaction: General Outlook and Research Datasets for User Authentication. In: 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 1273–1278. IEEE , Madrid, Spain (2020).