SilentSense: Silent User Identification Via Touch and Movement Behavioral Biometrics

Cheng Bo
Illinois Institute of Technology
Chicago, IL 60616, USA
cbo@hawk.iit.edu

Lan Zhang
Tsinghua University
Beijing, China
zhanglan03@gmail.com

Xiang-Yang Li
Illinois Institute of Technology
Chicago, IL 60616, USA
xli@cs.iit.edu

Qiuyuan Huang
University of Florida
Gainesville, FL 32605, USA
idfree@ufl.edu

Yu Wang
University of North Carolina at Charlotte
Charlotte, NC 28223, USA
yu.wang@uncc.edu

ABSTRACT

In this work, we present *SilentSense*, a framework to authenticate users silently and transparently by exploiting the user touch behavior biometrics and leveraging the integrated sensors to capture the micro-movement of the device caused by user's screen-touch actions. By tracking the fine-detailed touch actions of the user, we build a "touch-based biometrics" model of the owner by extracting some principle features, and then verify whether the current user is the owner or guest/attacker. When using the smartphone, the unique operating pattern of the user is detected and learnt by collecting the sensor data and touch events silently. When users are mobile, the micro-movement of mobile devices caused by touch is suppressed by that due to the large scale user-movement which will render the touch-based biometrics ineffective. To address this, we integrate a movement-based biometrics for each user with previous touch-based biometrics. We conduct extensive evaluations of our approaches on the Android smartphone, we show that the user identification accuracy is over 99%.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

Keywords

SilentSense, Mobile Device, Identification

1. INTRODUCTION

The blooming digital service for mobile devices has attracted more privacy concern, especially when people are sharing their personalized device to guest users. Since device owners are not willing to take distrust action to reduce permission deliberately before sharing [4], it would be good for devices to silently know exactly who is using it, so as to provide necessary privacy protection and access control.

The most popular mechanism for authentication is using enhanced password patterns [2] with an additional security layer, and establishing guest profiles for access control. Such methods are overelaborated, inconvenient and time consuming. Facial recognition [1] by front camera is another optional strategy to identify user. But it is still annoying to require users to accept frequent shooting. Besides, the accuracy is unstable with changing environment and frequent shooting is power-consuming.

The latest solution exploits the capacitive touch communication as a mechanism to distinguish different users [6], which has potential risk of being imitated. TapPrints [5] indicates that taps on the touch screen could be observed through sensitive motion sensors. Touchalytics [3] only exploits scrolling as biometric for continuous authentication while [8] only considers tap behaviors on certain digit patterns.

In this work, investigating the feasibility of utilizing the behavioral biometrics extracted from smartphone sensors for user identification, we propose *SilentSense*, a non-intrusive user identification mechanism to silently substantiate whether the current user is the device owner or a guest or even an attacker. Exploiting the combination of several interacting features from both touching behavior (pressure, area, duration, position) and reaction of devices (acceleration and rotation), *SilentSense* achieves highly accurate identification with low delay. A great challenge comes from the circumstance when the user is in motion, such as walking. The perturbation generated by the interacting will be suppressed by larger-scale user movement. While most of existing works neglect this circumstance, *SilentSense* is capable of identifying user in motion by extracting the motion behavior biometrics. As long as the current user is identified, necessary access control is triggered automatically.

Continuous monitoring sensors will provide minimum guest identification delay, but could cause unwanted energy consumption. Facing this challenge, we propose a novel model to estimate the current user leveraging the observation of owner's sociable habit. An online decision mechanism is designed for the timing to turn on or turn off sensors, which provides a balance between energy cost, delay and accuracy. Our online decision mechanism results in an adaptive observing frequency according to owner's social habit.

2. SYSTEM MODEL

SilentSense is designed as a pure software-based framework, running in the background of smartphone, which non-intrusively explores the behavior of users interacting with the device without any additional assistant hardware.

2.1 Main Framework

The framework model consists of two basic phases: *Training* and *Identification*, as shown in Figure 1. The training phase is conduct-

MobiCom'13, September 30–October 4, Miami, FL, USA.

ACM 978-1-4503-1999-7/13/09.
http://dx.doi.org/10.1145/2500423.2504572.

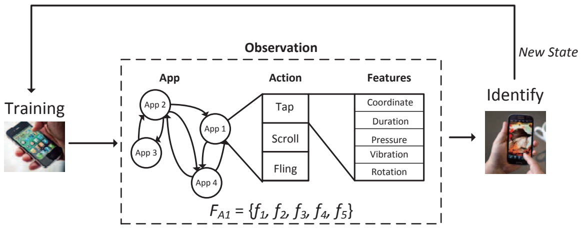


Figure 1: Framework Overview.

ed to build a behavior model when the user is interacting with the device, and the identification phase is implemented to distinguish the identity of the current user based on the observations of each individual's interacting behaviors.

Initially, the framework trains the owner's behavior model by retrieving two types of correlated information, the information of each touch-screen action and the corresponding reaction of the device. With the owner's behavior model, the current user will be identified through SVM classification. As personalized devices, the initial identification accuracy is usually not high enough because of lacking of non-owner's pattern. The model will upgrade the SVM model by adding newly observed features gradually through self-learning, which provides more accurate judgement.

2.2 Interacting Model

The operation of touchscreen based mobile device mainly consists of four gestures: Tap, Scroll, Fling, and Multi-touch. Different gestures usually have different touch features and lead to different device reactions. Interacting with certain app often involves a certain set of gestures. For a touch action T_i, we combine the app with its touch gesture and the features captured by this framework as one *observation*, denoted as $O_i = \{A_i, G_i, f_{i1}, \cdots, f_{in}\}$. here A_i is the app being used, G_i represents the gesture (*e.g.* tapp), and $f_{i,j}$ ($j \geq 1$) are features of the observed action.

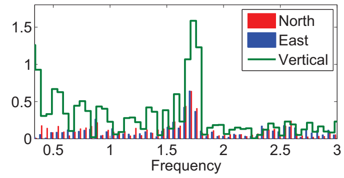
Two types of features are used in this system: the touch features and reaction features. The touch features include touch coordinate on the screen, touch pressure and duration, which can be obtained from system API. To capture the reaction features, we notice that diverse gestures and positions for holding the device by individual users infer different amplitudes of vibration caused by each touch, which has already been proved by previous works ([5, 7]). Such tiny reaction of the devices produces an identifiable patterns which could be observed via accelerometer and gyroscope.

2.3 Identification Process

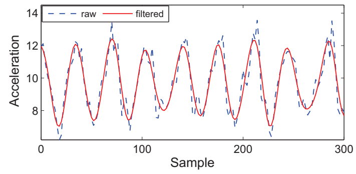
The purpose of the framework is to identify the current user of the device, and prevent sensitive information leakage if the user is not the legal owner. Generally, the three important issues that users concern about are *delay*, *accuracy*, and *energy consumption*.

We employ SVM to judge the identity of the current user according to each interacting behavior observation. Since it is difficult to validate the correctness of the results because of lacking of ground truth, we denote ε_i as the probability of the result, which is available for the SVM. Obviously, the accuracy of the identification process depends on the amount of observations, thus we denote $\theta_i(X_1, X_2, \cdots, X_i)$ as the probability of the identity based on the sequence of accumulated observation until X_i. With the number of observation increases, the framework will be more credible to provide the correct result, and the overall probability is generated according to historical creditability:

$$\theta_i(X_1, X_2, \cdots, X_i) = 1 - (\prod_{j=1}^{n}(1 - \varepsilon_j(X_j)))$$



(a) The FFT result of 4 step acceleration in the earth coordinate system.



(b) The raw vertical acceleration and filtered acceleration.

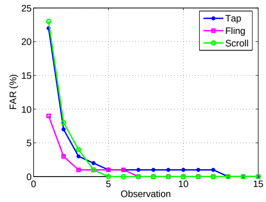
Figure 2: The frequency feature of acceleration in the earth coordinate system while walking.

On the other hand, framework in mobile device cannot neglect the energy consumption, coming from feature extraction and running the identification. We assume the function of $U(E_t, \theta_t)$ as the utility that could be achieved under the energy budget E_t and maintaining the reliable overall probability for the identification. The framework will make dynamic decision to balance the two factors so that the expected utility could be maximized. We can also set the threshold on the overall probability while minimizing the energy cost.

2.4 Motion Analysis

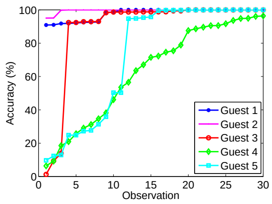
The amplitude of the sensory data extracted from the motion will be much larger than that of small perturbation caused by touch action, and the latter may be swamped by the former so that it fails to be extracted as a feature. In our work, we analyze the motion features, and use the walking features as part of the behavioral biometrics for identification.

To accurately capture the walking features of different users, three steps are conducted in our method. Firstly, considering a user could hold the phone in any attitude, we convert the raw acceleration vector in phone coordinate system into the earth coordinate system in real time. There are a lot of walking independent noise in the acceleration, which will greatly confuse the walking feature detection. We analyze the acceleration while walking in the frequency domain, Figure 2(a) shows that, the energy mainly locates around 2Hz, which is the user's walking frequency. The energy in other frequency comes from noise. To extract the pure walking acceleration, secondly, we filter linear acceleration with a band pass filter. Then we get the vertical acceleration in the gravity orientation and horizontal acceleration. Figure 2(b) shows the filtered vertical acceleration. A simple step detection algorithm can be performed on the filtered vertical acceleration in real time. Thirdly, we extract the walking feature from the processed acceleration data. The vertical displacement of a walker is directly correlated to his/her stride length and height, hence it is an important feature. Besides, the step frequency and horizontal acceleration pattern also vary with different users. To sum up, we extract four features of walking from EA_v and EA_h: (1) Vertical displacement of each step by double integration of EA_v; (2) Current step frequency, calculated by the duration of each step; (3) Mean horizontal acceleration for each step; (4) Standard deviation of EA_v for each step.



(a) FAR (b) FRR

Figure 3: FAR, and FRR by different actions and different number of actions observed.



(a) Guest Accuracy (b) Owner Accuracy

Figure 4: Identification based on combined operation.

3. PRELIMINARY RESULTS

We implemented *SilentSense* on Android phone as a service running background. This service obtains the current app and touch events from system API, and captures sensory data from accelerometer and gyroscope.

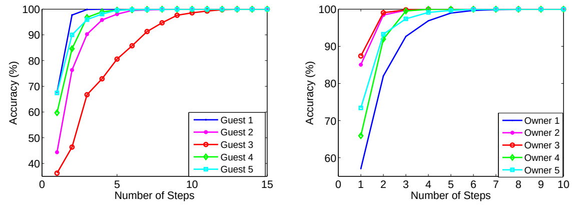
3.1 Identification in Static Scenario

First, we explore the uniqueness of the behavior biometric in the static scenario. With more than 100 actions for each user, we analyze the key features of users extracted from both the touching behavior and reaction of smartphone. The analysis shows that there exists big diversity of each interacting feature among different users.

We evaluate the performance of operating in three different gestures through three different apps, including *Message*, *Album*, and *Twitter*. In Figure 3, we plot the false acceptance ratio (FAR), and false rejection ratio (FRR) of identifying user by different actions with different number of total observed actions. From Figure 3(a), the mean FAR of identification by one observation of tap is 22%, by one fling action is 9%, by one scroll action is 23%. The FAR is reduced to below 1% after observing about 3 fling actions and with about 13 observations the FAR achieves 0 for all three gestures. Surprisingly, Figure 3(b) shows that FRR almost achieves 0 with only 2 observations for each gesture. The experiments result show great discrimination of three gestures based on multiple features extracted by *SilentSense*.

Now, we evaluate the performance of *SilentSense* in a more general scenario, with 100 users interacting with smartphones freely as their daily usages, *i.e.*, the action sequences are random combinations of three types of gestures. Since the amount of training behavior data for the guest user is much smaller than that of the owner's data set (we assume user usually not lend phone to others very often), our experiments show that it is much faster to achieve a high accuracy for identifying the owner than identifying the guest. Even so, the framework could reach over a 80% accuracy within 10 observations for identifying a guest. Figure 4(a) takes the results from five random selected guests and plots how soon the framework could identify the guest. Similarly, as shown in Figure 4(b), the owner will be identified with in 6 observations. Overall, in a general scenario, with only one observation, the FAR and FRR are about 20%. But, with about 12 observations of various actions, the FAR and FRR are both reduced to nearly 0, meaning that there is no incorrect identification.

3.2 Identification in Dynamic Scenario



(a) Guest Accuracy (b) Owner Accuracy

Figure 5: Identification based on walking feature.

In the dynamic scenario, we extract 4 walking features, including vertical displacement, step duration, mean and standard deviation of horizontal acceleration and establish a SVM model for dynamic walking features. The same users are required to use the smartphone while they are walking freely. We collect their processed vertical and horizontal accelerations in the earth coordinate system. After collecting necessary information, we combine the walking features with touch event features to establish the SVM model. And such touch event features only contains the duration, pressure, and the operation mode. Figure 5 presents the achieved identification accuracy increases with observed steps. As shown in Figure 5(a), after 12 steps, the accuracy to identify a guest can achieve 100%. Similarly, Figure 5(b) shows that after 7 steps, the accuracy to identify the owner can achieve 100%.

4. ACKNOWLEDGMENTS

The research of authors is partially supported by NSF CNS-0832120, NSF CNS-1035894, NSF ECCS-1247944, NSF CNS-1050398, National Natural Science Foundation of China under Grant No. 61170216, No. 61228202, No. 61272426, China Postdoctoral Science Foundation funded project under grant 2012M510029 and 2013T60119.

5. REFERENCES

[1] Visidon applock. http://www.visidon.fi/en/Home.

[2] 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 *CHI*, pages 987–996. ACM, 2012.

[3] 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 TIFS*, 2013.

[4] A. Karlson, A. Brush, and S. Schechter. Can I borrow your phone?: understanding concerns when sharing mobile phones. In *ACM CHI*, pages 1647–1650, 2009.

[5] E. Miluzzo, A. Varshavsky, S. Balakrishnan, and R. R. Choudhury. Tapprints: your finger taps have fingerprints. In *ACM MobiSys*, pages 323–336, 2012.

[6] T. Vu, A. Baid, S. Gao, M. Gruteser, R. Howard, J. Lindqvist, P. Spasojevic, and J. Walling. Distinguishing users with capacitive touch communication. In *ACM MobiCom*, pages 197–208, 2012.

[7] Z. Xu, K. Bai, and S. Zhu. Taplogger: Inferring user inputs on smartphone touchscreens using on-board motion sensors. In *the 5th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pages 113–124, 2012.

[8] N. Zheng, K. Bai, H. Huang, and H. Wang. You are how you touch: User verification on smartphones via tapping behaviors. WM-CS-2012-06, http://www.wm.edu/as/computerscience/documents/cstechreports/WM-CS-2012-06.pdf