COMPUTERS & SECURITY 53 (2015) 143–155

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cose

Computers & Security

Input extraction via motion-sensor behavior analysis on smartphones

Chao Shen [a,*], Shichao Pei [a], Zhenyu Yang [a], Xiaohong Guan [a,b]

[a] Xi'an Jiaotong University, No.28 Xianning West Road, Xi'an 710049, China
[b] Tsinghua University, Haidian District, Beijing 10084, China

ARTICLE INFO

Article history:
Received 30 March 2015
Received in revised form
18 June 2015
Accepted 29 June 2015
Available online 6 July 2015

Keywords:
Human-computer interaction
Smartphone security
Behavior analysis
Motion sensor
Input inference
Performance evaluation
Algorithm comparison

ABSTRACT

Smartphone onboard sensors, such as the accelerometer and gyroscope, have greatly facilitated people's life, but these sensors may bring potential security and privacy risk. This paper presents an empirical study of analyzing the characteristics of accelerometer and magnetometer data to infer users' input on Android smartphones. The rationale behind is that the touch input actions in different positions would cause different levels of posture and motion change of the smartphone. In this work, an Android application was run as a background process to monitor data of motion sensors. Accelerometer data were analyzed to detect the occurrence of input actions on touchscreen. Then the magnetometer data were fused with accelerometer data for inferring the positions of user inputs on touchscreen. Through the mapping relationship from input positions and common layouts of keyboard or number pad, one can easily obtain the inputs. Analyses were conducted using data from three types of smartphones and across various operational scenarios. The results indicated that users' inputs can be accurately inferred from the sensor data, with the accuracies of 100% for input-action detection and 80% for input inference in some cases. Additional experiments on the effect of smartphone screen size, sampling rate, and training data size were provided to further examine the reliability and practicability of our approach. These findings suggest that readings from accelerometer and magnetometer data could be a powerful side channel for inferring user inputs.

1. Introduction

Keyboard has long been the most common target for external or insider attackers to intercept users' sensitive or important input, such as password, email content, and transaction code (Asonov and Agrawal, 2004; Aviv et al., 2012; Cai and Chen, 2012; Foo Kune and Kim, 2010; Schlegel et al., 2011; Vuagnoux and Pasini, 2009). Currently, smartphones have become omnipresent platforms of personal computing for users to access the Internet and online services, and more and more personal information has been stored in smartphones. The obsolescence of physical keyboard on smartphones seems to provide a solution for the information leakage of sensitive user inputs. Extant smartphone operating systems (e.g., Android) also constrain that the input information and touch-input position cannot be directly read from the hardware (Mylonas, 2008). However, current smartphones are commonly equipped with various motion and location sensors, which can be used to measure the motion and posture

* Corresponding author. Tel.: +86 82663939.
E-mail address: cshen@sei.xjtu.edu.cn (C. Shen).
http://dx.doi.org/10.1016/j.cose.2015.06.013
0167-4048/

change of the smartphone. Also, Android smartphones have no limitation for third-party applications to access and read the sensor data (Mylonas et al., 2013). Along with the fact that input actions on touchscreen will cause small shaking of the smartphone, these findings offer us an opportunity of analyzing the behavior data from motion sensors to infer the position of touch actions, thus further to acquire users' inputs.

In this paper, we present an empirical study of inferring users' input on Android smartphone by analyzing the characteristics of motion sensors' data. The rationale behind our work is that different input actions would cause different levels of posture and motion change of the smartphone, which is based on different positions and strength of touch actions. We utilized an Android application (run as a background job) to monitor the data of accelerometer and magnetometer sensors. We then employed Kalman filter and denoising methods to obtain unbiased sensor data, and developed a procedure to detect the occurrence of input actions by analyzing the change of accelerometer data. We next extracted descriptive statistics features and wavelet transform features from the accelerometer data and magnetometer data to characterize the motion and posture changes of the smartphone, and applied classification technologies to establish an inference model of the positions of input actions. Finally, through the mapping relationship from input positions and common layouts of keyboard or number pad, the user inputs could be easily speculated. Experimental results across various types of smartphones and different operational scenarios showed that user inputs on number pad could be accurately inferred from the data of accelerometer and magnetometer sensors, with the accuracies of 100% and 80% for input-action detection and user-input inference in some cases. We also examined our performance against different smartphone screen sizes, different sampling rates, and different training data sizes, to further testify its reliability and practicability in practice. The main purpose and contributions of this work are fourfold.

Firstly, we lay empirical work for user input inference on smartphones by a side channel that relies on the analysis of the characteristics of accelerometer and magnetometer sensors. We investigate the reliability and practicability of inferring user inputs on smartphones based on the change of the sensors' data, without dedicated and explicit attention from users.

Secondly, we model characteristics of the accelerometer and magnetometer sensors by proposing newly-defined procedural features, such as fluctuation features and wavelet features, to characterize different positions of input actions in an accurate and robust manner. These features can lead to a performance boost both in input action detection and input position inference. We also apply four types of two-class classifier [*Random Forest*, *Support Vector Machine*, *Neural Network*, and *Nearest Neighbor*] to build the inference model, so that we can examine whether an observed effect is specific to one type of classifier or holds for a range of classifiers.

Thirdly, we examine the proposed approach in terms of the sensitivity to training data size, scalability to screen size, and flexibility to sampling rate, to further examine the applicability and generalization capability of the proposed approach in practice.

Finally, this study systematically evaluates the user-input inference by analyzing data of accelerometer and magnetometer sensors on smartphone, and extensive experiments across various types of smartphones and different operational scenarios show the proposed approach can perform input inference with a high accuracy. These results suggest that readings from accelerometer and magnetometer data could be a powerful side channel for inferring user inputs.

The structure of the paper is as follows. Section 2 discusses related work. Section 3 describes data acquisition process. Section 4 details the proposed approach for user-input inference. Section 5 presents our experimental design and results. Section 6 discusses and concludes.

2. Background and related work

2.1. *Input inference on smartphone*

Virtual keyboard is the most common input interface on smartphones, by which users may enter some security and privacy information, such as password, PINs, credit card numbers, or phone numbers. Therefore how to sniff or infer these inputs has recently aroused the interest from both attackers and researchers. In 2004, Asonov and Agrawal (2004) found that user inputs could be inferred by analyzing the sound emanated by the pressed keys. Later on, Foo Kune and Kim (2010) found that the timing information between two adjacent keystrokes could be used to accurately speculate users' input. Aviv et al. (2010) investigated the feasibility of inferring the touch pattern from oily residues on the touchscreen by users' fingers. Then some researchers (Raguram et al., 2011) exploited the visual feedback of magnified keys provided by smartphones' virtual keyboards to infer users' input. These ideas can be used for shoulder-surfing, whereby an attacker uses his own smartphone to video-record other users while they type PINs or messages.

Yet only recently have researchers come to the find of smartphone sensors as a sensitive source for users' input sniffer or speculation. Mylonas (2008) first showed the privacy risk of sensors utilized in smartphones. Xu et al. (2009) provided a way by a Trojan with access to the camera sensor for extracting private information. Later, Schlegel et al. (2011) analyzed the audio sensor for the same purpose. Cai and Chen (2011) also raised the concerns regarding the camera, the microphone, and the GPS signal for input inference in modern smartphones. Recently, Simon and Anderson (2013) demonstrated such an attack by exploiting the camera and microphone sensors and Spreitzer (2014) demonstrated this attack by exploiting the light sensor. But these attacks are less insidious because these sensors are protected with access permission by mobile operating systems. Besides, some researchers also investigated the reliability of the usage of accelerometer sensor in smartphones to capture inter-key timing measurements for extracting users' input on a nearby PC-keyboard (Marquardt et al., 2011), or to detect motion changes caused by users' taps on the screen and further to infer users' input on smartphones (Aviv et al., 2012; Cai and Chen, 2011; Owusu et al., 2012; Xu et al., 2012).