A Behavioral Biometrics User Authentication Study Using Motion Data from Android Smartphones

Javid Maghsoudi and Charles C. Tappert
Seidenberg School of CSIS, Pace University
Pleasantville, NY 10570, U.S.A.
javidm1@gmail.com, ctappert@pace.edu

Abstract- This study examined the behavioral biometrics of smartphone motion to determine potential authentication accuracies on Android phones. The study used machine learning algorithms to analyze data from the accelerometer and gyroscope sensors. Android smartphone data were captured from sixty different individuals, resulting in a large collection of datasets for training and testing. The data were filtered by removing noise and segmented into motion intervals prior to feature extraction. The classification algorithms employed in the study were Multilayer Perception, k-Nearest Neighbor, Support Vector Machines, and Naïve Bayes. Authentication accuracies achieved ranged from 81% to over 97%.

Keywords - Behavioral biometrics, machine learning, user authentication, smartphone sensors, accelerometer, gyroscope, Weka.

I. INTRODUCTION

The purpose of this study was to determine whether behavioral biometrics would be suitable for purposes of authenticating users of smartphones. User authentication via behavioral biometrics is comparatively less established than the use of physiological biometrics. This umbrella term, behavioral biometrics, encompasses a range of gestures including those normally applied to mobile devices such as touch gestures, motion and orientation as well as others such as mouse dynamics, handwriting, grip, and gait/stride.

The challenges inherent in authentication make behavioral biometrics appealing for a number of reasons. For one, it will likely be harder for someone with malicious intent to successfully capture a natural motion compared to a password or even a fingerprint. Natural motions also provide the option of continuous authentication since motions such as holding the device, walking around with it, and holding it up to the user's ear are ongoing activities.

The outcome of the study supports the use of behavioral biometrics, and strongly supports the use of behavioral biometrics if used in tandem with other authentication technologies.

II. RELATED WORK AND LITERATURE REVIEW

There has been significant growth in the number and usage of mobile devices in recent years, and this will only continue to grow. Mobile devices are rapidly becoming a key computing platform, transforming how people access business and personal information [9].

A. Information Security

The need to secure private or sensitive information in mobile devices is one of the main problems in information security. However, usual methods such as passwords and tokens fail to keep up with the challenges presented due to many drawbacks [1, 3].

Access to information is not limited to personal information only. Users increasingly need to access information for business reasons. Access to business data from mobile devices requires secure authentication, but traditional password schemes based on a mix of alphanumeric and symbols are cumbersome and unpopular, leading some users to avoid accessing business data on their personal devices altogether [9].

Biometric recognition could be a good alternative to overcoming the difficulties of password and token approaches. There are many different types of biometrics that we can use to authenticate users like touch-based or non-touch type biometrics [1, 3].

User authentication is categorized into three types of techniques which are: passwords (what one knows), tokens (what one has), and biometrics (what one is) [1, 4]. Using passwords and tokens, while they have provided some benefits, are vulnerable to theft. These tokens and passwords fail to keep pace with the challenges presented because they can be lost or stolen [1, 2]. The industry and research have turned their focus on finding more secure ways to authenticate users. Biometrics authentication has thus become a primary focus of academic research and industry adoption/implementation to provide users enhanced security and authentications.

B. Biometric Research

We can group biometrics into two general categories: (1) physiological biometrics and (2) behavior biometrics. Some physiological biometrics are briefly discussed. A fingerprint scanner requires an image of a finger to determine if the pattern of ridges and valleys in the image matches the pattern of ridges

and valleys in a stored profile [7]. Voice verification is another biometric that uses the pitch, tone, and rhythm of speech. Iris recognition has become a common site at most airports which provides services that allow frequent travelers to bypass immigration queues. Hand geometry has been available for some time and has been deployed in various organizations such as airports. Vein pattern recognition has become popular more recently and its recognition is based on the science that blood vessel patterns are unique to each individual and cannot be counterfeited since it requires a "live specimen" to work [7]. Facial recognition is another biometric also widely used at airports [7, 10].

Some behavioral biometrics are briefly discussed. Signature authentication and recognition encodes the dynamic movements of the signature signer. Keystroke dynamics identifies users based on their typing patterns. Mouse biometrics is based on learning a unique pattern from the user's mouse movement activity [7].

Mobile devices capture biometric data via built-in sensors that measure motion, orientation, and various environmental conditions. These sensors are capable of providing raw data with high precision and accuracy and are useful to monitor three-dimensional device movement or positioning. For example, sensors are used to get gestures and motions, such as tilt, shake, rotation or swing. The Android platform supports broad categories of sensors such as motion sensors that measure acceleration forces and rotational forces along three axes, or those that measure the physical position of a device [8].

We can use accelerometers to identify and authenticate smartphone users based on a person's movements since a person's movements form a unique signature. We can collect data from individuals having a mobile phone equipped with the sensor to collect data as they walk, jog, and climb stairs. The acceleration data collected while walking, jogging, ascending and descending stairs all have the potential to function as biometric signatures [5]. The way a phone is held or kept at different positions through motions can be used to authenticate users. Primo, et al, [6] tackled this question on the impact of variations in the position of the phone to perform continuous authentication.

III. Data Capture

The study used the Sensor Kinetics Pro App to collect data from six Android-based phone models. All the phone models except for the Motorola Moto G had both accelerometer and gyroscope sensors. The Motorola Moto G did not have a gyroscope sensor so it provided only accelerometer data. Each of the six phones was used to gather data from 10 different subjects, who each performed 20 trials, for a total of 200 runs per phone or 1200 runs in total. The captured data, 40 CSV files per participant (20 trials using 2 sensors) were accumulated.

Figure 1 shows the Sensor Kinetics Pro screenshot of a data sample and Figure 2 shows a sample image of the numeric data captured.

Figure 1. Sensor Kinetics Pro screenshots on an Android device demonstrating the X, Y and Z components of an accelerometer measurement.

time	X_value	Y_value	Z_value
0.000	0.00000	0.00000	0.00000
0.060	0.20000	0.64000	10.28000
0.130	-0.05000	0.64000	9.94000
0.199	0.06000	0.62000	10.18000
0.259	0.05000	0.66000	10.18000
0.329	0.06000	0.63000	10.14000
0.400	0.03000	0.68000	10.14000
0.464	0.03000	0.64000	10.12000
0.529	0.03000	0.64000	10.11000
0.590	0.01000	0.67000	10.13000

Figure 2. Data captured from the Sensor Kinetics Pro application.

This study used a total of 60 subjects to perform 1200 distinct trials, which resulted in 2200 separate datasets (note that there was no gyroscope data from one of the phones). The machine learning algorithms were used against the 2200 datasets.

During the process of gathering the 20 trial runs from participants, the testers would find that poorly obtained trials would have to be discarded as outliers. The rate of discards per test subject ranged from 1 to 4, or a rate of 5% to 20%.

IV. Data Processing and Machine Learning

The Weka software was used for data analysis and predictive modeling. Data was processed using two different methods. First, a simple division of the recorded trial runs was executed. Then an advanced feature extraction process isolated the motions and pauses into separate segments.

A. Simple Division

The Java program read in the 20 trials of subject motion data, provided in .CSV. Then that data was divided into 16 sections, which was chosen to maintain parity with the second method, the algorithmic feature extraction detailed below (4 segments, each broken into 4 sections).

The data points in the 16 sections were averaged and then had their variances calculated in the x, y, and z axes separately. With 2 sensors, this totaled 192 data points.

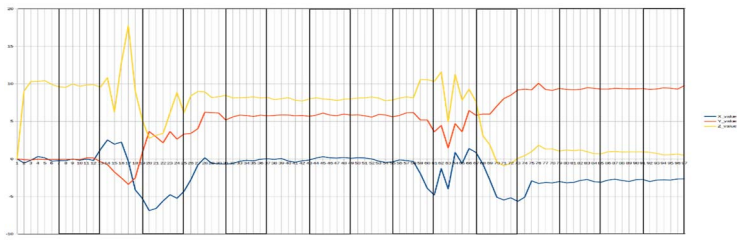

Figure 3. Simple division of a trial run into 16 sections with no attention paid to areas of motion or stillness.

B. *Complex Division (Algorithmic Extraction)*

For the algorithmic feature extraction, the following steps were performed: Using a graph of the accelerometer data, a moving average and variance were calculated for each time sample (separate x, y and z axes). Summing the variances for each sample gave a picture of the motion of the phone at each point in the trial run. A temporary threshold of motion was then adjusted via processing each run until it produced four segments: two in motion, and two at rest (held in front of the face, and at the ear). Depending on whether the accelerometer and gyroscope captured data at the same sampling rate (the Nexus 5 did not), the time indices for the gyroscope were adjusted to match the timestamps identified for the accelerometer.

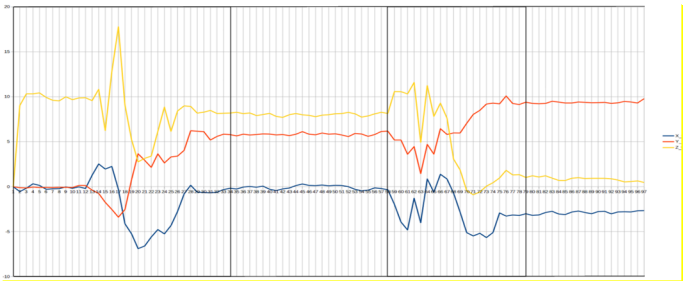

Figure 4. Complex processing of the motion to isolate movement from stillness: Not shown: the division of these 4 segments into 4 sections.

Each segment was divided into quarters, and then averaged ({x,y,z} separately) and had a variance taken ({x,y,z} separately), so that each segment had 24 attributes for the accelerometer, and 24 attributes for the gyroscope, if present. Each trial consisted of four segments, and thus consisted of 192 total data points.

The average and variance per dimension {x,y,z} are taken in each section, intended to capture each individual's idiosyncrasies. As the test subjects performed the motions in accordance with their own preferences, the relative speed/acceleration of those motions were distinguishable from others. The quarter-section averages were used in each dimension representing the speed/angle of the motion of that section on average. The variances were used to approximate individual variation within each section.

The use of quartiles was intended to recapture individual idiosyncrasies, while making the data amenable to the Weka library's numerous classification algorithms. Dividing up the data into smaller partitions would possibly create samples with too few data points, resulting in meaningless variances.

For both the simple division and the feature extraction algorithm, the information was aggregated into different .CSV files that were processed using Weka, and each 192-data-point row was post-pended with the test subject's number.

C. *Machine Learning Algorithms*

We used the following algorithms to learn the subjects and to identify them: Naïve-Bayes, k-Nearest Neighbors (K-NN), Support Vector Machines (SVM), and the Multilayer Perceptron (MLP).

Naïve-Bayes is a well-known and a very fast algorithm to run. The performance of Naïve-Bayes was not strong. It assumes features are independently formed from one another.

The k-Nearest Neighbors algorithm serves as a benchmark as being a standard in machine learning, and was also a fast and simple algorithm to run. It is known for sensitivity to local structures in the feature space.

Support Vector Machine (SVM) was also used because of its popularity and its wide use in many applications.

Multilayer Perceptron (MLP) was also used. During the training, the weights are adjusted after an initial arbitrary setting to close in upon the intended results. The executions took longer with MLP.

V. Test Results

Cross validation using 10-folds was used with each of the algorithms. The total test data was split into 10 equal partitions, and the algorithms were trained on 9/10ths. Then the remaining 10th was tested against the trained algorithms.

Table 1 shows the results for the 60 users, tested using six phones (five different models), using algorithms Naïve-Bayes, k-NN, the MLP, and Support Vector Machines.

First, there was clear improvement when the gyroscope was added to the analysis for three of the algorithms. SVM's performance stayed the same. Having the extra sensor aided those algorithms, which resulted in improvements of 1.1%. SVM did very well with one sensor.

Second, the improvement in specifically isolating the motions using the complex segmentation method is visibly evident in the results. An improvement of 1.1% is noted by doing this additional work using K-NN algorithm.

Third, the choice of algorithm mattered, for complex segmentation and using both sensors, going from N-B to k-NN improved about 6.2%, and moving to MLP resulted in an additional 2.9% gain. From the worst to best there was a difference of about 9.1%.

%	Simple (acc)	Simple (both)	Complex (acc)	Complex (both)
N-B	82.7	83.2	81.1	83.6
k-NN	86.3	87.0	88.7	89.8
MLP	91.4	92.5	92.9	92.7
SVM	96.3	92.8	97.7	92.2

Table 1. Results of Naïve Bayes, k-NN, MLP; using both simple and complex segmentations, and using the accelerometer only, or both sensors.

Improvement	Percent
From 1 sensor to 2 sensors (for MLP)	1.1%
From simple to complex (for K-NN)	2.1%
From NB to k-NN (For Complex)	6.2%
From k-NN to MLP (for Simple)	5.5%
From NB to MLP (for Complex)	9.1%

Table 2. Approximate values of improvements when increasing the number of sensors, using complex segmentation, or using a sophisticated algorithm.

One minor detail worth noting was that the multilayer perceptron algorithm seemed to perform slightly worse moving from one to two sensors in concert with the complex extraction. Given the general propensity of the algorithms to improve when (1) more sensors are used, and (2) the complex method was applied, the authors of this study believe that overall, the best combination of factors were: 2 sensors, complex extraction, and the multilayer perceptron or Support Vector Machine (SVM) algorithms.

VI. Conclusion

The results support the importance of using behavioral biometrics in user authentications in conjunction with a standard authentication like password to reach the highest level of accuracy or close to 100%. We also found that as the number of participants increased there was a slight decrease in performance.

The MLP algorithm, when fed 1200 test trials and 196+1 attributes (+1 being the identification label), struggled under the weight of the learning data. With the 236,400 data points, the MLP 10-fold learning took over 40 minutes to complete. The results dictate that MLP was accurate enough that combining it with any of the other algorithms would weaken the results. There were no instances were MLP misclassified a sample and both k-NN and N-B classified it correctly. Therefore, fusing the results of all three algorithms would serve no purpose other than to lower the results of MLP alone.

The learning algorithms performed better with additional data: having the gyroscope data added to the analysis improved the results across the board.

For phones with sensors that have a lower sampling rate but still wish to offer biometric authentication, cubic interpolation could possibly be an option. Cubic interpolation makes use of neighboring data points to approximate missing data. This idea however would require significant additional testing to see whether or not such virtual data would be able to compete with phone with high sample rates.

One particularly noteworthy feature of the results was the fact that the machine learning algorithms frequently confused the data samples provided by a set of twins, subjects 43 and 44, as their data samples were often not ascribed to the correct individual (Figure 5). This finding had some influence on the study's outcome but is a large factor in the final recommendation that the sort of behavioral biometric described in this study would not be effective as the only method of authentication.

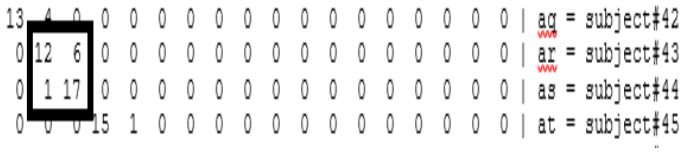

Figure 5. Confusion matrix: the effect of twins in the study. Results from the k-NN run against the complex feature extraction with both sensors.

VII. Future Development

After completing the analysis, there are some suggestions for future iterations of the experiment. One change would be to break up the data-gathering into sections. In this experiment, the test subjects had to perform all 20 trials in a single sitting. It would be interesting to see if collecting trials over multiple sessions would have any effect on the results.

VIII. Acknowledgment

The authors thank the various masters-level capstone project teams that gathered the data samples, ran the Weka algorithms, and coded the Java program.

IX. References

[1] A. A. Alariki, and A. A. Manaf, "Biometrics Authentication Using Touch-Based Gesture Features for Intelligent Mobile Devices," in *1st Int'l Conf. of Recent Trends in Information and Communications Technologies*, Johor, Malaysia, 2014, pp. 528-538.

[2] A. A. Alariki, and A. A. Manaf, "Behavioral Biometrics Authentication Using Gesture-Based." *Journal of Computational and Theoretical Nanoscience*, vol. 20, no. 2, pp. 492-495, Jan 2014.

[3] H. Crawford, "Keystroke Dynamics: Characteristics and Opportunities," in *Int'l Conf on Privacy Security and Trust (PST)*, Ottowa, Canada, 2010 pp. 205-212.

[4] R. Giot, M. El-Abed, B. Hemery, and C. Rosenberger. "Unconstrained keystroke dynamics authentication with shared secret." *Computers & Security*, vol. 30, no. 6, pp. 427-445, Sep 2011.

[5] J. R. Kwapisz, G. M. Weiss, S. A. Moore, "Cell Phone-based Biometric Identification." 2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)(2010): n. pag. Web.

[6] A. Primo, V. V. Phoha, R. Kumar, A. Serwadda, "Context-Aware Active Authentication Using Smartphone Accelerometer Measurements," *in IEEE CVPR*, Columbus, OH, 2014, pp. 98-105.

[7] "The Role of Biometrics in IT Security and Continuous Authentication." [Online]. Available: https://www.behaviosec.com

[8] "Sensors Overview," [Online]. Available: http://developer.android.com/guide/topics/sensors/sensors_overview.html

[9] 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" from *Annual Computer Security Applications Conf.*, Orlando, FL, 2012, pp. 159-168.

[10] "What Is Facial Recognition? - Definition from WhatIs.com," [Online]. Available: http://whatis.techtarget.com/definition/facial-recognition