Android Pattern Unlock Authentication - effectiveness of local and global dynamic features

Nasiru Ibrahim[1] Harin Sellahewa [2]

Abstract: This study conducts a holistic analysis of the performances of biometric features incorporated into Pattern Unlock authentication. The objective is to enhance the strength of the authentication by adding an implicit layer. Earlier studies have incorporated either global or local dynamic features for verification; however, as found in this paper, different features have variable discriminating power, especially at different extraction levels. The discriminating potential of global, local and their combination are evaluated. Results showed that locally extracted features have higher discriminating power than global features and combining both features gives the best verification performance. Further, a novel feature was proposed and evaluated, which was found to have a varied impact (both positive and negative) on the system performance. From our findings, it is essential to evaluate features (independently and collectively), extracted at different levels (global and local) and different combination for some might impede on the verification performance of the system.

Keywords: Biometrics, Authentication, Behaviometrics, Mobile Security, Pattern Recognition, Multi-factor authentication, Android Pattern Unlock, Mobile Computing, Graphical Password, Touch Gesture-Based Authentication.

1 Introduction

All Android operated devices have Pattern Unlock implemented as an alternative to PIN / Password and biometrics authentication. The pattern authentication uses a grid of 3x3 nodes where users connect at least four nodes in a predefined order to gain access to the device. Pattern Unlock is the most popular of the many efforts in graphical authentication. Some others are: PassFaces [Co], Draw-a-Secret [DY07], PassShapes [WL08] and Microsoft Picture Password [St11]. Android Pattern Unlock has the advantages of being memorable and usable compared to PIN and Password [CL15]. However, it has significant demerits, chief of which is security. The method is susceptible to shoulder surfing [Ye17], smudge attack [Av10], guessing (statistical and probabilistic) and heuristic attacks [IS17, Av10, Ue13]. Further, it has been reported that users' choose patterns in a way that is easy to compromise. Biases were reported in the start node selection (top-left), high frequency of patterns, highly associative patterns with letters and digits, and under-utilisation of complex patterns characteristics - overlaps, crosses and intersections [ABk15, Ue13, IS17], which adversaries could use as the foundation of attacks. Due to these weaknesses, the authentication does not provide adequate protection to the users' data on the device.

We investigate how Pattern Unlock security can be enhanced by incorporating biometric features to provide multi-factor authentication. The hypothesis is that incorporating behaviour biometric features (behaviometrics) to Pattern Unlock would enhance its security and also preserve usability. To authenticate a user, the system would first check the pattern presented (if same with enrolled pattern), then verifies how the user performs it. Behavioural biometric features collected from the smartphone's touchscreen sensors are used to verify how the user performs the pattern. Both checks have to be true for access to be gained.

We focus on the performances (discriminating power) of different types of dynamic features extracted both locally and globally. Two evaluations were carried out: 1) investigate the performances of existing

[1] School of Computing, The University of Buckingham, United Kingdom, nasiru.ibrahim2@buckingham.ac.uk
[2] School of Computing, The University of Buckingham, United Kingdom, harin.sellahewa@buckingham.ac.uk

features, and 2) investigate the performance of a novel feature – Pattern Accuracy. In all evaluations, global and local feature extractions were studied. The results shows variable performances between features extracted locally and globally with best performance when they are combined - highlighting their significance. Further, results shows security of Pattern Unlock is enhanced by the proposed new feature and combined features give the highest discriminating power.

The contribution of this paper is in the comprehensive evaluation of features on different levels of extraction. Then, the introduction of a novel feature, Pattern Accuracy, which when incorporated with existing features, provides the best system performance.

The paper is structured as follows. Section 2 reviews existing studies on improving the security of Pattern Unlock with dynamic features. Section 3 describes the research methodology, including data, features and experimental protocol. Results are reported in Section 4 and finally conclusions in Section 5.

2 Literature Review - Pattern Unlock with Dynamic Features

Existing literature on Pattern Unlock improvement can be categorised into two. First, static approach, considers modification of the grid layout, selection rule alteration and persuasion [Ue13, KN14, GMM16, Co16]. Second, dynamic approach, involves incorporation of an implicit layer into Pattern Unlock. We review the main literature in the second category (and refer the reader to the literature on the static approach for a detailed discussion).

De Luca et al. [Lu12] first enhanced Pattern Unlock with an additional implicit layer of security. The study was carried out with data collected from 31 participants on an Android device. Participants were assigned unique patterns to perform 21 times in 21 days (one per day). After the 21 inputs, the participants were asked to draw all other users' patterns three times as forgery samples. For each sample, *(x,y) coordinates, finger pressure, finger size, time* and *speed* were recorded. Dynamic Time Warping (DTW) was used for evaluation with a total of 645 genuine samples and 2790 forgery samples. To enrol a user, the first five samples were used to create a template which was selected based on the lowest warp distances (median, mean, minimum and maximum) to the other four samples. After the template has been selected, the warp distance (median, mean, minimum, maximum and standard deviation) to the other four samples are used for comparisons with the remaining samples. The study achieved optimal performance with the template chosen based on median warp distance with a threshold of maximum warp distance. The system obtained 398 true positives, 231 false positives, 858 true negatives, 92 false negatives with 19% false rejection rate (FRR), 21% false acceptance rate (FAR) and accuracy of 77%.

Angulo and Wastland [AW12] conducted a similar experiment with an implicit layer on Pattern Unlock. In the study, 32 participants completed the experiment, including 12 females, ages 19 to 56 years. Four smartphones were used for the experiment: Samsung Galaxy SII (18), Nexus S (8), HTC Legend (4) and HTC Vision (2). Participants were required to perform three six-node patterns 50 times (each) consecutively. For every sample, two features were extracted: *finger-in-dot time* - the time (in ms) from a finger touching a node to finger dragging outside the node and *finger-in-between-dot time* - the finger movement time from one node to the next (speed). The feature vector had a total of 11 features - six finger-in-dot times and five finger-in-between-dot times. For analysis, the first ten samples were discarded; the next 25 were used for training and 15 for testing. Performances were evaluated with six different classifiers: Euclidean Distance, Manhattan Distance, Mahalanobis Distance, Recursive Partitioning (RPart), Support Vector Machine (SVM) and Random Forest. The EER reported for all classifiers were: 27.35% - Euclidean, 25.60% - Manhattan, 23.03% - Mahalanobis , 29.70% - RPart, 14.06% - SVM and 10.39% - Random Forest (best).

de Wilde [dWSV15] investigated the performance of Pattern Unlock with biometric features, although in identification mode. The study involved 144 participants (25 female and 12 left handers) which were required to perform a five-node pattern ten times in eight days. For analysis, a likelihood-ratio based classifier with *x,y coordinates* and *time* features were used. 96 classes were used as training sets and 48 for testing. The true match rate (TMR) was reported based on static false match rate (FMR) for different feature combinations. At 10% FMR, the TMR reported was 53.1% (*x,y coordinates* and *time*), 58.0% (*x,y coordinates*) and 54.8% (*x,y coordinates* and *normalised time - time/max time*). The study reported an average EER of 19.0% with *x,y coordinate* and *time*, 18.2% with *x,y Coordinate* and *normalised time (time/max time)* and best EER of 16.9% with only the *x,y coordinate*.

The above literature shows that biometric features can be used to enhance the security of Pattern Unlock authentication. However, the studies have shortcomings that we considered in our investigation i.e. low accuracy and use of only features extracted on one level.

3 Data and Methodology

3.1 Data

Due to the unavailability of public data sets on Pattern Unlock with dynamic features, the authors collected a new data set, which can be obtained by other researchers free of charge by contacting the authors. The database consists of 140 participants including 82 males, 20 left-handers, and ages between 18 and 70. Data were collected on Xperia Z3 with 5.2" touchscreen display running Android 5.1.1 OS using an app developed for this research. Participants were tasked with drawing a nine-node pattern (in a static position - sitting - while holding the device) seven times. Before drawing the patterns, the participants were allowed to practice drawing patterns on the smartphone until they felt comfortable.

A static approach was adopted in which all users performed the same pattern to simulate skilled forgeries. The researchers ensured all conditions, procedures, instructions and devices were the same for all participants, ensuring consistency and avoiding variations caused by variable experimental conditions. Further, no information was disclosed that could influence the natural behaviour of participants until after the data collection.

The raw data recorded were: finger press time, finger release time, finger press pressure, finger release pressure, (selected) nodes coordinates, finger coordinates (finger path from press to release), pressure on nodes, times at nodes, finger movement pressure and time, from which features were created.

3.2 Features

Features were extracted on two levels: globally and locally. In global extraction, the features were extracted from data obtained on the entire pattern, i.e. from pattern start-to-finish, while in local extraction, the features were extracted per node of the grid, i.e. between two nodes. The features used include both existing (have been used in earlier studies) and a novel feature are described in sections 3.2.1 and 3.2.2 respectively. In all, 112 features were extracted both locally and globally.

3.2.1 Existing Features

The existing features extracted are *duration, distance, pressure and speed*. These features were used in [Lu12, dWSV15, AW12] and are described below:

1. Duration: Time-stamps were recorded at each point on the trajectory as the users perform the patterns from which global and local durations are calculated.

$$Global\ Duration = \mid press\ time\ -\ release\ time\ \mid \tag{1}$$

$$Local\ Duration_i = \mid Node\ time_{i+1}\ -\ Node\ time_i\ \mid \tag{2}$$

2. Distance: The distance is the numerical measure of how much a user's pattern has covered between two points. Two points (with x,y) and one-dimensional (with x or y) distances are calculated using equations 3 and 4 respectively.

$$Two\ Points\ Distance = \sqrt{\sum_{i=1}^{i<m}((x_i - x_{i+1})^2 + (y_i - y_{i+1})^2)} \tag{3}$$

$$One\ Dimensional\ Distance_i = \sum_{i=1}^{i<m} \mid x_i - x_{i+1} \mid \qquad where\ m = number\ of\ coordinates \tag{4}$$

3. Pressure: The pressure sensor measured the amount of force exerted on the touchscreen along the pattern trajectory. *Minimum*, *maximum*, *median*, *average* and *standard deviation* of the pressure were calculated both globally and locally.

4. Speed: The speed is the rate at which the users draw the pattern. It is obtained by dividing the pattern distance (in pixels) and pattern duration (in time, ms), as shown in equation 5.

$$Speed = \frac{Distance}{Duration} \tag{5}$$

Global and local speeds are used as features, depending on extraction, the respective distances and durations are used.

3.2.2 Proposed Feature

The proposed feature in this study is Dynamic Pattern Accuracy. Dynamic Pattern Accuracy measures the consistency of the user in performing a pattern. The measurement was based on the distance between a reference pattern path and the path from a user's pattern (green line in Figure 1). The reference path is represented by a line (invisible to the user - red line in Figure 1) that goes through the centre of nodes. Dynamic Time Warping (DTW) was used to measure the *accuracy* and the global and local *mean* and *standard deviation* of accuracy were calculated and used as features.

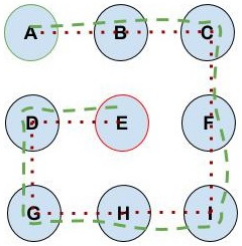

Fig. 1: Example of accuracy variation (from which global and local accuracy are extracted)

3.3 Experimental Protocol

Every participant contributed seven samples (totalling 980 samples). The data set was divided into two sets of 70 users. A genuine class was selected from one set, and the next five classes were selected as imposters (i.e. five imposters per genuine user). We used a leave-one-out data partition for train/test (and repeated seven times; hence, results are average of the runs). Genuine (train and test) samples were selected from the same set while imposter (train and test) samples were selected from the other set - thus, the imposter train and test samples do not overlap - simulating practical scenario where the imposters used in training a system would not be the same imposters that would try to access the system.

Two experiments were conducted: the first experiment investigated the performances of existing features extracted at different levels. The second experiment investigated the performance of proposed Dynamic Pattern Accuracy feature on the system performance, including the impact when incorporated with existing features. All experiments were carried out on MATLAB 9.0 (R2016a) using the LIBSVM library developed by Chang and Lin [CL11].

4 Results

4.1 Experiment One: Performance of individual features at local and global levels.

We present the performances of the existing category of features used in dynamic pattern authentication. The features used are: *distance*, *duration*, *pressure* and *speed*. In the studies above (Section 2), these features were extracted either globally or locally. Here, we provide a comprehensive presentation of the features extracted highlighting the significance of the holistic investigation.

The performances are presented for all extraction levels (global, local and combined), individual, and all features combined. The performances (EER and AUC) are highlighted in Table 1.

Features	EER %			AUC		
	Global	Local	Combined	Global	Local	Combined
Distance	33.55	18.13	15.20	0.7561	0.9016	0.9224
Duration	37.71	24.18	23.80	0.6689	0.8374	0.8404
Pressure	26.78	17.80	17.89	0.8255	0.9034	0.9029
Speed	29.78	29.73	24.51	0.8029	0.7579	0.8347
All (Benchmark)	12.70	10.78	08.39	0.9445	0.9565	0.9700

Tab. 1: Performances (EER and AUC) of individual and combined features measured at global, local and combined extraction levels

The table shows the discriminating power of features, individually and collectively, at varied extraction levels. Analysing the global and local features, the best performing individual feature was achieved using *pressure* feature with 17.80% EER and 0.9034 AUC on local extraction. On the other hand, the worst performing feature was achieved using *duration* feature with 33.55% EER and 0.6689 AUC on global extraction - performing slightly better than random guessing (0.5 AUC).

However, the performances when all features are combined in their respective extraction levels were interesting. When all global extraction of features were combined the performance was 12.70% EER (and 0.9445 AUC) and 10.78% EER (and 0.9565 AUC) when all local features were combined. Overall, the combined local extraction of the features performed better than the combined global extraction which suggests that there is higher discriminating potential (high variability between-class) in features at sampled points rather than the whole pattern, generally.

On combining feature extraction levels, the best performing individual feature was *distance* with 15.2% EER and 0.9224 AUC. The other features saw a substantial increase in performance except for *pressure*, where the performance was best in local extraction rather than combined - the decrease was marginal (0.09% EER and 0.0005 AUC). Although the decrease was marginal, it suggests that not all combination of extraction levels on features would increase the performance, some might even hinder the performance. The best performance of 8.39% EER (and 0.9700 AUC) was achieved when all features were combined, including global and local extractions.

Angulo and Wastland [AW12] who used local extraction of *speed (finger-in-between dot time)* and *time (finger-in-dot time)* features achieved the best performance of 10.39% EER with Random Forest classifier. Our best local extraction result was 10.78% EER with all features combined. [AW12] result was obtained with 25-10 genuine train-test samples while ours was with 6-1 genuine train-test sample. Further, on classifier comparison, they achieved 14.06% EER with SVM.

De Luca et al. [Lu12] used *(x,y) coordinates, pressure, finger size, time* and *speed* with DTW. They achieved best performance (77% accuracy, 21% FAR and 19% FRR) with *pressure*, *size* and *speed* combined. Using the same metrics, our performances were 9.84% FAR, 9.59% FRR and 90.20% accuracy (9.71% EER) with *pressure* and *speed*.

4.2 Experiment Two: Performance of Proposed Features

Dynamic Pattern Accuracy was proposed for incorporation into dynamic Pattern Unlock authentication. The feature was derived from touch movement – described in 3.2.1. The performances of the proposed feature are presented in two stages: 1) pattern with dynamic accuracy feature (alone), and 2) pattern with all features. The performances are shown on Table 2.

With the dynamic accuracy as the implicit feature, the system performed similar to the existing features in that local extraction performed better than the global. The global dynamic accuracy had an EER of 35.27% while having 24.38% EER on local dynamic accuracy. The *dynamic pattern accuracy* feature with the benchmark improves the performance by 3.4% EER (global) while marginal (0.13% EER) improvement on local extraction. In this case, the global extraction performed better than the local extraction, an opposite outcome to the benchmark features. When all extraction from the benchmark features and *dynamic pattern accuracy* were combined, the best system performance was obtained of 8.08% EER. The result compared to the benchmark was a increase by 0.31% EER - indicating a variable impact of the local and global dynamic accuracy feature.

The two instances where global dynamic accuracy was used, the system performance has increased while a decrease was observed with the use of local accuracy. Hence, we hypothesise that the accuracy features extracted locally provide a negative influence on the system performance. Further, the influence is highlighted in the combined case. To investigate, we look at the performances of benchmark with global dynamic accuracy, benchmark with local dynamic accuracy, and benchmark with combined dynamic accuracy levels. Between benchmark with combined levels of dynamic accuracy (8.08% EER) and benchmark with local dynamic accuracy (9.49% EER), there is a drop of 1.41% EER while between benchmark with combined levels of dynamic accuracy (8.08% EER) and benchmark with global dynamic accuracy (7.71% EER), there is an increase of 1.78% EER. Hence, we conclude that local dynamic accuracy feature has a negative contribution to the performance of the system, and the best performance is achieved with benchmark and global dynamic accuracy features.

The findings in this section confirm that Pattern Unlock possess dynamic features that could be used enhance its security and without significant degradation to the usability. Furthermore, the dynamic features

Features	EER %			AUC		
	Global	Local	Combined	Global	Local	Combined
Benchmark	12.70	10.78	08.39	0.9445	0.9565	0.9700
Accuracy	35.27	24.38	21.90	0.7214	0.8345	0.8577
Benchmark with Accuracy	09.30	10.65	08.08	0.9661	0.9555	0.9696
Benchmark + Global Accuracy	07.71			0.9738		
Benchmark + Local Accuracy	09.49			0.9662		

Tab. 2: Performances (EER and AUC) of Proposed Features

were extracted at different levels (global and local), and the findings found the extractions to have variable discriminating power. Interestingly, the local extraction performed better than the global extraction of same features, except for *dynamic pattern accuracy* where the global outperformed the local extraction. When all extractions were combined, the best performances were achieved. Therefore, we recommend investigating different extractions, and the combination of features as not in all cases would the best performance be the use or combination of all features, in fact, it might cause a reduction in the performance.

The paper evaluated the impact of incorporating dynamic pattern accuracy features to Pattern Unlock. First, the benchmark performance was found to have increased with dynamic pattern accuracy on all feature extraction levels compared to without dynamic pattern accuracy. The improvement was highest on the global level as the EER was reduced by 3.5 percentage points. On the local (and combined extraction) level, only a marginal reduction was observed which prompted further investigation into the feature. We found that the local extraction of the feature possesses lower discriminating power mainly due to high variability and instability in the small sampled points caused by the temporal effect of user pattern. But, system best performance was obtained with global dynamic pattern accuracy

5 Conclusion

The paper investigates the improvement approaches to Pattern Unlock authentication with dynamic features. It presented a comprehensives evaluation of features and their extraction levels. The results indicate, first, feature level extraction is essential and has a significant impact on system performance, second, adding more features does not necessarily translate to better system performance, Further, it is crucial when designing and implementing a biometric system especially touch-based authentication systems, to evaluate the extraction of features and their combinations as their fusion might not always produce the best system performance mainly due to correlation with the class label.

With these findings, feature selection techniques could be implemented to measure the correlation between a class and other features. The best features should have a high inter-class variation and low intra-class variation; otherwise, the feature would be redundant - making the classifier less efficient with more features, present over-fitting and achieve lower precision.

In conclusion, the paper has shown viable avenues in which the Pattern Unlock security could be significantly improved without affecting the usability of the scheme. Further, the findings show that the current layout and pattern creation restriction does not have to be modified or changed to improve the security of the scheme. However, even when the grid size is increased, pattern creation restriction lifted, users can choose longer and more complex patterns which would give more data points to the features and potentially higher discriminating power thereby increasing the classification accuracy and the effort needed to compromise the layers.

References

[ABk15] Aviv, A. J.; Budzitowski, D.; kuber, R.: Is bigger better? Comparing User-Generated Passwords on 3x3 vs. 4x4 Grid sizes for Android's Pattern Unlock. Annual Computer Security Applications Conference (ACSAC), Los Angeles, CA. USA, 2015.

[Av10] Aviv, A. J.; Gibson, K.; Mossop, E.; Blaze, M.; Smith, J. M.: Smudge Attacks on Smartphone Touch Screens. In: Proceedings of the 4th USENIX Conference on Offensive Technologies. WOOT'10. USENIX Association, pp. 1–7, 2010.

[AW12] Angulo, J.; Wästlund, E.: Exploring Touch-Screen Biometrics for User Identification on Smart Phones. In (Camenisch, Jan; Crispo, Bruno; Fischer-Hübner, Simone; Leenes, Ronald; Russello, Giovanni, eds): Privacy and Identity Management for Life. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 130–143, 2012.

[CL11] Chang, CC.; Lin, CJ.: LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:1–27, 2011. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[CL15] Clark, G. D.; Lindqvist, J.: Engineering Gesture-Based Authentication Systems. IEEE Pervasive Computing, 14(1):18–25, 2015.

[Co] Passfaces: Two Factor Authentication for the Enterprise, URL: http://www.realuser.com/index.htm (accessed: 10-06-2019).

[Co16] Colley, A.; Seitz, T.; Lappalainen, T.; Kranz, M.; Hkkil, J.: Extending the Touchscreen Pattern Lock Mechanism with Duplicated and Temporal Codes. Advances in Human-Computer Interaction, 2016.

[dWSV15] de Wilde, L.; Spreeuwers, L.; Veldhuis, R.: Exploring How User Routine Affects the Recognition Performance of a Lock Pattern. In: International Conference of the Biometrics Special Interest Group (BIOSIG). pp. 1–8, 2015.

[DY07] Dunphy, P.; Yan, J.: Do Background Images Improve "Draw a Secret" Graphical Passwords? In:CCS '07 Proceedings of the 14th ACM conference on Computer and communications security, 2007.

[GMM16] Guerar, M.; Merlo, A.; Migliardi, M.: ClickPattern: A Pattern Lock System Resilient to Smudge and Side-channel Attacks. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 8:64–78, 2016.

[IS17] Ibrahim, N.; Sellahewa, H.: Touch gesture-based authentication: A security analysis of Pattern Unlock. In: IEEE International Conference on Identity, Security and Behavior Analysis (ISBA). pp. 1–8, 2017.

[KN14] Kwon, T.; Na, S.: TinyLock: Affordable defense against smudge attacks on smartphone pattern lock systems. Computers & Security, 42:137 – 150, 2014.

[Lu12] Luca, A. De; Hang, A.; Brudy, F.; Lindner, C.; Hussmann, H.: Touch Me Once and I Know It's You!: Implicit Authentication Based on Touch Screen Patterns. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '12, pp. 987–996, 2012.

[St11] Signing in with a picture password, URL: https://blogs.msdn.microsoft.com/b8/2011/12/16/signing-in-with-a-picture-password/ (accessed: 10-06-2019).

[Ue13] Uellenbeck, S.; Durmuth, M.; Wolf, C.; Holz, T.: Quantifying the security of graphical passwords: the case of Android Unlock Patterns. Conference on Computer & Communications Security (CCS), 2013.

[WL08] Weiss, R.; Luca, A. De: PassShapes: Utilizing Stroke Based Authentication to Increase Password Memorability. In: Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges. NordiCHI '08. ACM, pp. 383–392, 2008.

[Ye17] Ye, G.; Tang, Z.; Fang, D.; Chen, X.; Kim, K. I.; Taylor, B.; Wang, Z.: Cracking Android Pattern Lock in Five Attempts. In: NDSS. 2017.