POSTER: TouchTrack: How Unique are your Touch Gestures?

Rahat Masood
University of New South Wales (UNSW), CSIRO Data61
Sydney, Australia
rahat.masood@student.unsw.edu.au

Benjamin Zi Hao Zhao
CSIRO Data61
Sydney, Australia
ben.zhao@data61.csiro.au

Hassan Jameel Asghar
CSIRO Data61
Sydney, Australia
hassan.asghar@data61.csiro.au

Mohamed Ali Kaafar
CSIRO Data61
Sydney, Australia
dali.kaafar@data61.csiro.au

ABSTRACT

This paper studies a privacy threat induced by the collection and monitoring of a user's touch gestures on touchscreen devices. The threat is a new form of persistent tracking which we refer to as "touch-based tracking". It goes beyond tracking of virtual identities and has the potential for cross-device tracking as well as identifying multiple users using the same device. To demonstrate the likelihood of touch-based tracking, we propose an information theoretic method that quantifies the amount of information revealed by individual features of gestures, samples of gestures as well as samples of gesture combinations, when modelled as feature vectors. We have also developed a purpose-built app, named "TouchTrack" that collects data from users and informs them on how unique they are when interacting with their touch devices. Our results from 89 different users indicate that writing samples and left swipes can reveal 73.7% and 68.6% of user information, respectively. Combining different combinations of gestures results in higher uniqueness, with the combination of keystrokes, swipes and writing revealing up to 98.5% of information about users. We correctly re-identify returning users with a success rate of more than 90%.

KEYWORDS

Touch-based Tracking, Mobile Privacy, Behavioural Biometrics, Touch Gestures

1 INTRODUCTION

In this paper, we postulate that the very distinguishability of touch-based gestures constitutes a major privacy threat as it enables a new form of tracking of individuals. We call this notion "touch-based tracking," which is the ability to continuously and surreptitiously observe, track and distinguish users via their touch gestures while they are interacting with touchscreen devices.

As opposed to "regular" tracking mechanisms (e.g., based on cookies, browser fingerprints) which track virtual identities [1–3], touch-based tracking is subtle and riskier as it allows the tracking and identification of the actual (physical) person operating the device. Touch-based tracking also leads to cross-device tracking where same user can potentially be traced on multiple mobile devices. Additionally, we also envision the risk of distinguishing and tracking multiple users accessing the same device. Not all use cases of touch-based tracking are negative. It can also be beneficial to users and service providers alike. For instance, the identification of multiple users using the same device may help in providing content more suitable for each of them. Nevertheless, touch-based tracking performed in any of the above cases can provide a more complete view of a user's behavior and can be used for a range of purposes including targeted ads, profiling, and spamming. Our main contributions are as follows:

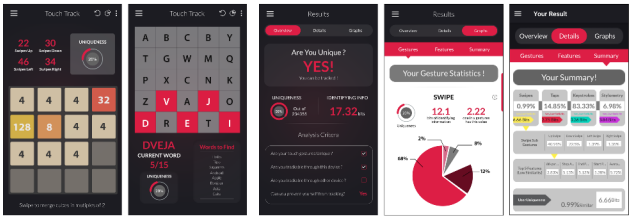
Contributions: We investigate the potential of using touch-based gestures for tracking, which we call touch-based tracking. We develop an analytical framework that measures the amount of identifying information (in bits) contained in these gestures, represented as feature vectors, at different levels of granularity. At the finest level, our framework quantifies the information carried by individual features, e.g., pressure on screen. At the second level, we quantify the information carried by a gesture sample, e.g., a single swipe. At the third level, our framework calculates the amount of information carried by multiple samples of the same gesture, e.g., a collection of swipes. Lastly, we measure the information carried by a collection of samples from multiple gestures, e.g., swipes and taps. We apply our framework on four widely used touch screen gestures: i) swipes, ii) taps, iii) keystrokes, and iv) handwriting.

We develop and deploy a "game-like" Android app called "TouchTrack" which consists of three well known open source games and one purpose-built handwriting module. We test our framework on a total of 40,600 gesture samples collected from 89 participants and identified features that contain high amount of identifying information using the *maximum-relevancy minimum-redundancy* (mRMR) algorithm [4]. With the same dataset, we measured the amount of information contained in samples from the same gesture and from multiple gestures. We found that 50 features in a single handwriting sample contribute 68.71% of information about users, which increases to 73.7% with multiple samples. We further identified that different gestures combined together reveal more information about users. For instance swipes, handwriting, and keystrokes carry 98.5% of information. Among users who performed all the four gestures, our framework showed 98.89% of information about users.

We also validated our framework in terms of correctly identifying a returning user. We found that with multiple samples, swipes and handwriting show a TPR of 90.0% and 91.0%, respectively. For a

CCS '17, October 30-November 3, 2017, Dallas, TX, USA

ACM ISBN 978-1-4503-4946-8/17/10.
https://doi.org/10.1145/3133956.3138850



(a) 2048 (b) Lexica (c) Results (d) Graphs (e) Summary

Figure 1: TouchTrack Result Screens

combination of gestures we found that swipes and handwriting combined together had a TPR of 93.75%.

2 DATA COLLECTION

To illustrate the potential of touch-based tracking, our *TouchTrack* app collects gesture samples as raw readings from the touch sensors, sends them to our server, and informs users about the uniqueness of their gestures by displaying the uniqueness results. The app is made up of four games, three of them are based on popular open-source games and a fourth game was purposely developed by us. These games are i) 2048[1] to collect up, down, left and right swipes, ii) Lexica[2] to collect taps, and iii) Logo Maniac[3] to collect keystrokes, and iv) "Write Something" (developed by us) to collect writing samples. The four games were selected so as to capture user touch interactions in a natural way. Screenshots of the TouchTrack App are displayed in Figure 1. When a new user uses TouchTrack, he/she is required to sign up using a unique username, which together with the device ID is hashed and stored in our database. Prior to data collection, we underwent and obtained ethics approval. The users were informed about the purpose of TouchTrack and what data is being collected.

The Raw Dataset and Statistics: We gathered raw touch and motion features from the `MotionEvent` and `SensorEvent` Android APIs of the device. From these raw features we derived more features to capture information such as averages, standard deviations, mins, maxs etc. These derived features are called extracted features. Examples of the extracted features are *median of first 5 acceleration points, 80-percentile of pairwise x-tilt,* and *standard deviation of change of area position.* We extracted 229 features for swipes, 7 for taps, 8 for keystrokes, and 241 for handwriting. Table 1 shows the summary statistics of our data collection. There were a total of 89 users who downloaded and used our app, however, only 30 users used all four games and hence provided samples for all gestures.

Table 1: Touch Gesture Data Statistics

Gesture	No. of Users	No. of Samples	Gesture	No. of Users	No. of Samples
Swipes	81	16611	Up Swipes	78	3568
Down Swipes	71	4781	Left Swipes	63	4252
Right Swipes	65	4010	Handwriting	36	1291
Taps	89	16225	Keystrokes	49	6473
All Gestures:	30	25186			
Total:	89	40600			

[1] https://github.com/gabrielecirulli/2048 [2] https://github.com/lexica/lexica
[3] https://github.com/Luze26/LogoManiac

3 THE METHODOLOGY

Our quantitative methodology is based on relative mutual information. To illustrate this, we consider quantifying uniqueness of a single feature by fixing a gesture, say right swipe. Let $\mathbb{U}$ denote the set of users and let $\mathbb{F}$ denote the range of values of the feature. Let U and F denote the random variables associated with these sets. Then, the relative mutual information is defined as

$$I_R(U;F) = 1 - \frac{H(U \mid F)}{H(U)},$$

where $H(U) = \log_2 |\mathbb{U}|$, and $H(U \mid F)$ is defined as $H(U \mid F) = -\sum_{f \in \mathbb{F}} \Pr(F = f) H(U \mid F = f)$. Finally,

$$H(U \mid F = f) = -\sum_{u \in \mathbb{U}} \Pr(U = u \mid F = f) \log_2 \Pr(U = u \mid F = f).$$

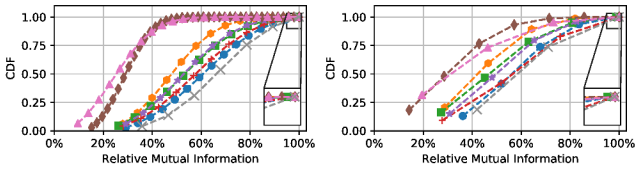
It boils down to quantifying the probabilities in the above two equations, which we do empirically through our dataset. Given a feature value f, $\Pr(U = u \mid F = f)$ is calculated by counting the number of times f has been exhibited by user u divided by the number times f appears in the dataset. $\Pr(F = f)$ is calculated from the empirical distribution of F. The methodology for quantifying uniqueness of gesture samples, modelled as featue vectors, is different. This is because due to high dimensionality of the feature vector, it is unlikely that any two feature vectors from the same user will be exactly the same. Thus, the probabilities are calculated differently using "fuzzy" matching (as opposed to exact). A similarity metric (in our case cosine similarity) is used to decide whether a received feature vector is similar to a given feature vector in the dataset. The rest of the calculation of the relative mutual information is similar to above.

4 RESULTS

In this section, we present the results of applying our framework on the touch gestures. We identify a set of features for each gesture type, and apply our methodology on the selected features to show uniqueness results.

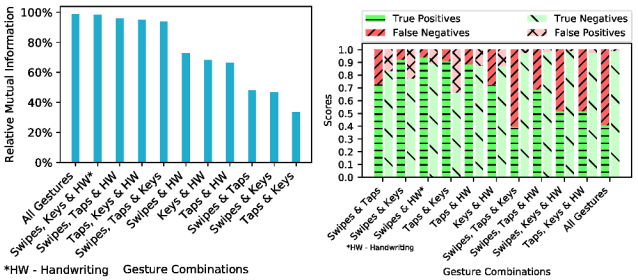
Feature Subset Selection (FSS): We intend to find the uniqueness of gestures as a function of increasing number of features. To do this, we needed a ranking of features in terms of their distinguishing capacity. We use the maximum-relevance-minimal-redundancy (mRMR) algorithm that attempts to constrain features to a subset which are mutually as dissimilar to each other as possible, but as similar to the classification variable as possible [4]. We used sets of top i features from each gesture according to their mRMR rank, where i was incremented in discrete steps until m (the total number of features). We then evaluated their relative mutual information using our framework for the uniqueness of a single gesture sample and multiple samples from the same gesture.

We note that for all gestures, the relative mutual information increases with increasing number of features. Also, the uniqueness of a set of gesture samples is generally higher than single samples, and in all cases surpasses the uniqueness of single samples as we increase the number of features. The uniqueness of multiple swipe samples is the highest, with 92.01%, followed by handwriting (85.93%) and downward swipes (77.52%). On the other hand, samples of taps and keystrokes exhibit least uniqueness carrying 34.73% and 41.02% of information. We observe that given a single



(a) CDF of a Gesture Sample (b) CDF of Set of Gesture Samples

Figure 2: CDF of a Gesture Sample and a Set of Gesture Samples on a Single Device. Relative Information of Respective Categories (sample, set of samples) are: -●- Swipes: (57.7%, 63.3%), -○- Up Swipes: (48.5%, 50.23%), -■- Down Swipes: (52.2%, 54.5%), -+- Left Swipes: (53.9%, 68.6%), -⋆- Right Swipes: (53.3%, 57.4%), -◆- Taps: (29.5%, 34.7%), -▲- Keystrokes: (26.2%, 41.0%), -×- Handwriting: (68.7%, 73.7%)



(a) Uniqueness Results (b) Classification Results

Figure 3: Uniqueness and Classification Results for Gesture Combinations.

gesture sample, handwriting provides 79.49% of information about the user and a keystroke gives the least amount of information i.e. 28.76%.

Uniqueness of Features: Before assessing the uniqueness of features we binned any continuous features or features with a large domain. We found that *80-percentile of area* in left swipe reveals 56.10% of information about a user, followed by *20-percentile of area* in down swipe 55.50%. Among features which are shared among all gestures,*start area* contains 52.5% of information, followed by *start pressure* yielding 45.4% of information. On the other extreme, *inter-stroke time* for a keystroke reveals minimum amount of user information, i.e,. 7%. We observe no trend (in terms of dependency) among features,except that relative information decreases in decending order of the features.

Uniqueness of a Gesture Sample: Once we have fixed the set of features, we need to find the threshold τ of the cosine similarity metric that balances uniqueness of gesture samples and correctly (and uniquely) identifying a returning user. To do this, we split the data into an 80-20 partition, and then evaluated the equal error rate (EER). We observe that our methodology correctly re-identifies a returning user up to 81% (19% EER) of the time if given a handwriting sample. The worst performance is a TPR of 61% (39% EER) when a sample of keystroke is provided. After fixing the threshold, we computed the uniqueness through our relative mutual information metric. The results showed that a handwriting sample reveals the highest amount of information (68.71%), followed by swipes (57.77%). The four types of swipes, i.e., left, up, down, and right swipes, yield 53.9%, 52.2%, 52.2%, and 48.5% of information, respectively. However, taps and keystroke reveal only 29.5% and 26.2% of information. Figure 2a shows the CDF of a gesture sample. We observe a considerable difference in the range of information revealed by different gestures, with handwriting exposing more than 60% of information for half of the users in the database. Following this, the swipes also show high uniqueness, revealing 30% to 65% of information about 75% of users.

Uniqueness of a Set of Gesture Samples: We computed a different threshold of the cosine similarity metric for this category, and chose the one which resulted in the best EER. We found that the rate of re-identifying a returning user is higher reaching up to 91% (9% EER) for handwriting. This means that combining a few samples of the same gesture may allow for more accurate tracking. We then calculated relative mutual information and found that handwriting reveals 73.7% of information, followed by left swipe which yields 68.6% of information of user gestures. In accordance with previous results, taps and keystrokes reveal minimum amount of information about users, i.e., 34.71% and 41.0%, respectively. The CDF is shown in Figure 2b.

Uniqueness of Combination of Gestures: Figure 3a shows the quantification results of multiple gestures in different combinations. We found that a combination of all gestures reveal a maximum 98.89% of information about users, followed by the combination of swipes, handwriting & keystrokes that yield 98.5% of information. We also tested these gesture combinations in terms of re-identifying returning users. Figure 3b shows the TPR and FPR of the different combinations of gestures. We see that as we increase the number of gestures in our combination, the FPR drastically decreases, but so does the TPR. For instance, all gestures together yields the 0.99% FPR but also a low TPR (just above 40%). The lowest FPR was recorded by the combination of swipes, handwriting and keystrokes (0.85%). The main reason for a big drop in TPR as compared to the rate of single gestures, is mainly due to the rather strict metric of only labelling a given combination as being from a user if the predicate for each gesture evaluates to 1.

5 FUTURE WORK

In the future we intend to extend our methodology to scenarios such as *single-device multi-user tracking* and *multi-device single-user tracking*. The first scenario distinguishes between multiple users accessing the same device. The second scenario is the tracking of the same user across multiple devices.

REFERENCES

[1] Peter Eckersley. 2010. How Unique Is Your Browser? *Proc. of the Privacy Enhancing Technologies Symposium (PETS)* (2010), 1–18. https://doi.org/10.1007/978-3-642-14527-8_1

[2] Pierre Laperdrix, Walter Rudametkin, and Benoit Baudry. 2016. Beauty and the Beast: Diverting Modern Web Browsers to Build Unique Browser Fingerprints. *Proceedings - 2016 IEEE Symposium on Security and Privacy, SP 2016* (2016), 878–894. https://doi.org/10.1109/SP.2016.57

[3] Łukasz Olejnik, Claude Castelluccia, and Artur Janc. 2012. Why Johnny Can't Browse in Peace: On the Uniqueness of Web Browsing History Patterns. *5th Workshop on Hot Topics in Privacy Enhancing Technologies (HotPETs 2012)* (2012), 1–16. http://hal.archives-ouvertes.fr/hal-00747841/

[4] Hanchuan Peng, Fuhui Long, and Chris Ding. 2005. Feature selection based on mutual information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27, 8 (2005), 1226–1238. https://doi.org/10.1109/TPAMI.2005.159 arXiv:f