

In Depth Review of “Side-Channel Inference Attacks on Mobile Keypads Using Smartwatches”

By Wesley Lam




1. Introduction

The following paper is an in depth review of *Side-Channel Inference Attacks on Mobile Keypads Using Smartwatches* written by Anindya Maiti, Murtuza Jadliwala, Jibo He, and Igor Bilogrevic. These individuals are referred to as “the authors” in this paper, and *Side-Channel Inference Attacks on Mobile Keypads Using Smartwatches* is referred to as the “original paper” in this review. The original paper covers the use of side channel attacks on smartwatches to gain information on smartphone handheld numeric touchpad key presses. This review quickly summarizes the original paper, followed by a section of highlight aspects, then weaknesses and suggestions for improvement, and finally concluding with any additional remarks. It is highly recommended that the original paper be read prior to this review.

2. Summary

The original paper discusses the use of smartwatch side channel attacks to infer smartphone numeric touchpad key presses. This is done by accessing smartwatch gyroscope and accelerometer sensor data, which is done through user installed malicious software. Classifiers are created and trained depending on the target’s hand configuration between the smartwatch and the smartphone they are wearing. The authors explain that the appeal of using gyroscope and accelerometer sensor data is due to two significant factors. One, users are typically unable to disable such sensors. Second, applications do not require permission to access these sensors unlike other functions such as GPS or microphone use. The paper then divulges into great detail about the training and testing of each classifier, and the variety of different experiments performed to consider real life scenarios. Such experiments include sampling at different frequencies, using different make and model smartwatches between adversary and target, natural (faster) typing scenarios, different hand configurations, transitional movement and tracing methods, smartphone, smartwatch, smartphone and smartwatch sensor data acquisition, and different classifier training datasets. A summarized list of all results can be observed in the table 1. The original paper lists some limitations such as posture, ambient movement and power consumption. Finally the original paper concludes with suggestions for improvement in defense, like a system monitoring mechanism and in tracing with *Random Walk Tracing* [1].

Table 1. Overall Results of Different Experiments

Different Holding Configuration and Setup Inference Accuracy	Smartwatch Only Classification Used (50hZ)	Smartphone Only Classification Used (50hZ)	Combined Smartwatch and Smartphone Classification Used (50hZ)	Lower Frequency 25hZ	Lower Frequency 10hZ	Faster/Natural Typing Speed (50hZ)	Different Adversary/Target Smartwatch
Same Hand and Non-Holding Hand Typing (SH-NHHT) SH-NHHT  (a) 28.44% SH-HHT 7.56% Vice Versa 36% Chance of Occurance	1 v 1 84.58%	1 v 1 78.7%	1 v 1 88.91%	1 v 1 ~72%	1 v 1 ~23%	52% Accuracy	1 v Rest 70.08% Accuracy for Samsung Gear Live 26.62% Variance 67.41% for Urbane W150 56.26% Variance
	1 v Rest 70.18%	1 v Rest 63.3%	1 v Rest 71.59%	1 v Rest ~76%	1 v Rest ~18%		
	All v All 88.16%	All v All 85.5%	All v All 88.65%	All v All ~70%	All v All ~27%		
Same Hand and Holding Hand Typing (SH-HHT) SH-HHT  (b) 32.83% SH-HHT 16.17% Vice Versa 49% Chance of Occurance	1 v 1 83.5%	1 v 1 84.0%	1 v 1 90.66%	1 v 1 ~65%	1 v 1 ~19%	61% Accuracy	1 v Rest 71.16% Accuracy for Samsung Gear Live 17.24% Variance 70.83% for Urbane W150 77.00% Variance
	1 v Rest 71.16%	1 v Rest 70.9%	1 v Rest 74.29%	1 v Rest ~60%	1 v Rest ~21%		
	All v All 85.83%	All v All 86.8%	All v All 89.78%	All v All ~75%	All v All ~20%		
Different Hand and Non-Holding Hand Typing (DH-NHHT) DH-NHHT  (c) 2.11% Chance of Occurance	Transition Accuracy 88.42% Inference Accuracy 43.75%	?	Transition Accuracy 88.42% Inference Accuracy 82.50%	?	?	Transition Accuracy 79.76% Inference Accuracy 38.33%	?

3. Highlight Aspects

The authors should be praised for covering a lot of different aspects regarding this project of side-channel inference attacks on smartphone numeric key presses via smartwatch gyroscope and accelerometer sensors. One aspect that should be highlighted is the use of the ensemble classifier. The ensemble classification algorithm combines the use of five different classification algorithms, specifically:

1. Simple-Linear Regression (SLR)
2. Random Forest (RF)
3. K-nearest neighbors (k-NN)
4. Support vector machine (SVM)
5. Bagged Decision Trees (BDT)

The ensemble classification algorithm uses a majority vote system where it takes the result of the most popular answer from each individual classifier. This project could have been attempted by using only one of the original five classification algorithms. As a result of this ensemble classification, the overall inference accuracy is improved.

The comparison between smartwatch, smartphone, and smartwatch with smartphone data acquisition experiments provides a solid foundation of how relevant the authors' research is. Previous similar research performed side channel on smartphones but extracted gyroscope and linear accelerometer data from the target's smartphone instead of their smartwatch. Comparing the two different sets of research led to similar results, which proves that the authors are not performing an irrelevant or outdated study in this field of side channel attacks.

Another important aspect the authors covered is the variety of situations such as the varied frequency, different adversary and target smartwatch, different hand configurations and more. To realistically perform this kind of attack, a vast amount of different scenarios and factors must be considered. Also extending this approach of side channel attack into QWERTY keyboards is a good example of how their attack framework can be used in other attack applications. Despite the poor results from the QWERTY keyboard experiment, this experiment shows that potentially other information can be gathered using the same set-up and information from the original project.

The creation of the key press detection, energy threshold, and tracing algorithms are all important factors in designing the project. In addition, demonstration of the computation time shows the practicality of the time required to perform an attack. The maximum amount of time required for computation is for an *All versus All* scenario which took about 34 minutes in the experiments ran for this project.

4. Weaknesses and Suggestions

Over 87% of smartphone/smartwatch hand configurations are considered which means that 13% of the smartphone/smartwatch hand configurations are not studied. This implies that with the current status of the project, 13% out of any random target would be invulnerable to an attack. The authors acknowledge that they are missing the both hand typing scenario. One important fact to realize is that even if they know which hand configurations are missing, studying them may be a new challenge since it is not guaranteed that all hand configurations information can be pre-processed in the same way. In this very project, the Different Hand and Non-Holding Hand Typing (DH-NHHT) configuration and the other two (SH-NHHT and SH-HHT) configurations required different set-up due to the different limitations of the gathered information. Since DH-NHHT recorded information did not contain similar patterns as the SH-NHHT and SH-HHT recorded information, the authors had to develop the transition and tracing method for pre-processing DH-NHHT data which was extremely inferior in inference accuracy. To improve this, new hand configurations should be studied which has already been recommended by the authors.

The authors conclude that using different make and model smartwatches between target and adversary is feasible with reasonable accuracy. However, one major issue with this conclusion is that the authors only tested one particular set of different smartwatches with the Samsung Gear Live smartwatch and the LG Urbane W150 smartwatch. No other different combinations of smartwatches were tested which leaves an extremely low sample size of different pairs of combinations for different smartwatches. Wikipedia lists over 80 types of smartwatches which may yield different results than the two different smartwatches used for this experiment [2]. The amount of variance between different smartwatch combinations may cause unreliable attacks on certain targets as accuracy dropped to as low as nearly 60% without consideration of other hindering factors such as low sampling rate. To improve the authors claim that using different make and model smartwatches is feasible and reasonably accurate, they should carry out more experiments between multiple different make and model smartwatches since a sample size of 1 combination pair seems inefficient.

A weakness that is addressed by the authors is the smartwatch battery consumption rate at 50Hz sampling. The resulting sampling rate for an hour led to 31% smartwatch battery depletion which is potentially very alerting to a target. This depletion rate would raise suspicion and potentially ruin the side channel attack as the attack is supposed to be performed in a stealthy manner. The authors proposed that a mechanism to only record when the target is typing can resolve this issue.

The largest weakness the original paper contains is no consideration for a combination of realistic scenario factors. Low sampling rate, different make/model

smartwatches, and natural typing speed scenarios are all considered individually in their own controlled setting. No attempt at combining any of these two scenarios ever occur which is concerning when considering an actual real life scenario. Each one of these realistic scenarios lead to a drop in inference accuracy which when considered individually, still yields an acceptable accuracy result. However when a combination of these scenarios occur, it is expected that each one of these realistic scenario factors will cascade upon each other causing the inference accuracy to drop significantly. This does not even include other hindering factors not studied such as posture and movement which would also lower accuracy. To address this issue the authors should carry out mixed realistic scenario experiments to see if the results still are acceptable and reasonably accurate.

Practicality in performing the proposed attack contains a lot of additional aspects not covered in the original paper. For example implementing or making the target somehow install malicious software may be challenging on its own. Also identifying a target's smartwatch/smartphone hand configuration requires a method of surveillance or prior knowledge. Knowing when a target is using their numeric keypad is critical since otherwise the adversary would obtain useless information that they may miscategorize as legitimate data. Obtaining this knowledge would require either physical surveillance, or a built in mechanism into the smartwatch. Or some other means of observation is required to make this attack work as intended. Consideration for a target that constantly changes their hand configuration may be necessary since there is no guarantee that the target always holds their phone or wears their smartwatch the same way every time. The original paper can be improved by addressing these practicality issues. Another factor that damages this research that is not addressed is fingerprint identification, which entirely skips over the need for passcode identification which is common among newer smartphones. It is expected that over 70% of smartphones shipped in 2018 will contain the fingerprint identification feature [3].

One thing that was not considered in the original paper is the scale could this side channel attack could be performed at. The following preconditions exist for a successful attack:

1. The adversary must know the users hand configuration beforehand, and the user must use one of the hand configurations studied (87% chance for attack)
2. The adversary must somehow access the user's smartwatch accelerometer and gyroscope data, which is done with malicious software
3. To do this, the adversary must first find a way to present the software to the user
4. The user must then decide to install the software or the adversary can forcibly install the software onto the users smartwatch covertly

These preconditions make it difficult for the adversary to perform a large-scale attack with the data of many users being compromised at once. There are a few potential scenarios where this can occur however, such as sending out a large-scale scam email that advertises the malicious app or advertising the app itself for users to download. However whether or not the adversary has the financial resources to perform the latter is another different matter. To obtain the large scale hand configuration data, a survey could be a prerequisite to fill out on the users mobile or smartwatch device to download the app, or the app could require the user to select or somehow indicate how they hold their phone (if applicable since it would be a smartwatch app) and how they wear their smartwatch. Of course even in this scenario, the user could simply input false information into the survey which would provide the adversary with invalid data. A large scale attack is probably not realistic but it is still possible regardless. Therefore this type of side channel attack would be expected to narrow down to target specific launched attacks.

5. Use for Non-Malicious Intent

Traditionally performing a side-channel attack would assume the adversary is attempting to perform criminal activity like accessing a target's credit card number or identifying a target's passcode to access their smartphone. However, the use of this type of side-channel attack could potentially be used for good intentions like by the police or military. Gaining information against criminals is a valid way of using these methods of attacks lawfully since it could prevent damage from occurring. Smartphones have the capability of erasing all data stored on themselves if the user inputs a certain number of incorrect passcodes. This can provide a backup method for criminals to erase their tracks of illegal activity stored on their smartphone in case their smartphone gets compromised. An example of this situation is the 2015 San Bernardino Attack, in which 14 people were killed and 22 were injured due to a terrorist attack [4]. The FBI was able to obtain the attacker's smartphone, an Apple iPhone 5C, which was suspected to have important information related to the terrorist attack. However, the FBI could not initially unlock the smartphone since they did not know the 4 digit passcode. Additionally the phone was programmed to automatically erase all of its data after ten failed password attempts. This led to complications with a FBI vs Apple court case since Apple did not want to create an updated operating system that could disable security features. This is an example of an unpleasant situation that could be avoided if the passcode information was made available earlier which is possible by using the smartwatch key press inference side channel attack. The example provides at least one practical use that raises the appeal of the research performed by the authors.

References

- [1] Maiti, A., Jadliwala, M., He, J. and Bilogrevic, I. (2018). *Side-Channel Inference Attacks on Mobile Keypads Using Smartwatches*. IEEE TRANSACTIONS ON MOBILE COMPUTING,. IEEE, pp.1-15.

- [2] En.wikipedia.org.(2018). *Smartwatch*. [online] Available at: <https://en.wikipedia.org/wiki/Smartwatch>

- [3] Counterpoint Research. (2018). *More Than One Billion Smartphones with Fingerprint Sensors Will Be Shipped In 2018 - Counterpoint Research*. [online] Available at: <https://www.counterpointresearch.com/more-than-one-billion-smartphones-with-fingerprint-sensors-will-be-shipped-in-2018>

- [4] En.wikipedia.org. (2018). *2015 San Bernardino attack*. [online] Available at: https://en.wikipedia.org/wiki/2015_San_Bernardino_attack