Screen Unlocking Methods on Smartwatches Experimental Design

Weiwei Li wwli194@cs.toronto.edu

Zhongtian Qiu zqiu@cs.toronto.edu

Yiqing Yang yangyiq2@cs.toronto.edu

Yue Zhao yuezhao@cs.toronto.edu

PART 1: Research question and hypothesis

This project is motivated by the question that whether the extended use of authentication methods for smartphone is effective on smartwatch. Performance of different unlocking methods will be investigated on accuracy, efficiency as well as security.

The research question will be addressed by our experiment on the accuracy rate of the input of passwords/unlock patterns, the time it takes each participant to input the passwords/unlock patterns, and the rate of passwords/unlock patterns that are successfully replicated by shoulder surfing. A series of existing popular unlocking ideas for smartphones will be applied to smartwatches with different sizes of display, while adjustment will be made to improve the performance regarding experimental results.

PART 2: Apparatus

Camera: the camera will be used in shoulder test to record the process of unlocking the screen using different methods, from the back of the person. The camera will be placed 20 centimeters above the person's shoulder. The resolution of the camera should be at least 8 Megapixels; this could guarantee we can record the process clearly.

Monitor: a monitor will be used to play the video recorded in the process of the shoulder test to the participants, in order to breach the system. To make sure participants will not have any problem with the monitor resolution, we will use a 21-inch monitor (1280*1024) to play the replication videos.

Mobile phone: a mobile phone will be used to simulate different type of experiment setting. Because we need to test multiple smartwatch design, the use of a mobile phone brings the coherence and easiness. The phone should be Android OS based and have a screen that supports to simulate a 340*272 pixels setting in a 38mm watch face setting and 390*312 pixels setting in a 42mm setting.

PART 3: Participants

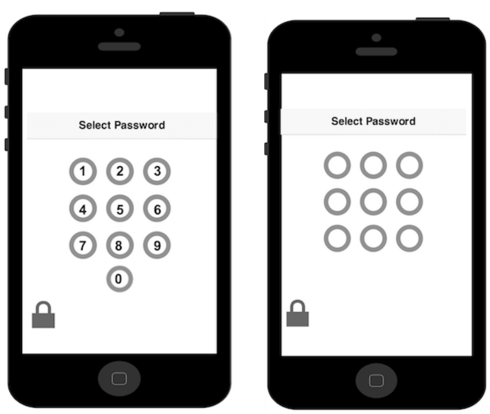
We will have 20 volunteers to take the experiments. Majority of the participants are CS graduate students in Master of Applied Computing program (MScAC) at University of Toronto, and there might be a few of participants from other backgrounds as well. Because we are all in the MScAC, we are going to talk to our fellow classmates in the same program directly and invite them to take the experiments. We also welcome our friends and families to the experiment to create the diversity and these people make up a small portion of the participant pool. Most of the participants are pursuing their master's degree and there is no bias in gender. All the participants have some basic knowledge or experience about using smartphone, and they are in the age range of 20 to 30.

PART 4: Experimental design

We design our experiment based on two independent variables, layout of the unlock screen and the size of the smart watch. By separately doing experiment on these two variables, we aim at finding the best layout and its corresponding screen size in accordance with our metric, dependent variables. Still, three dependent variables will be analyzed, i. e. accuracy, efficiency, and security.

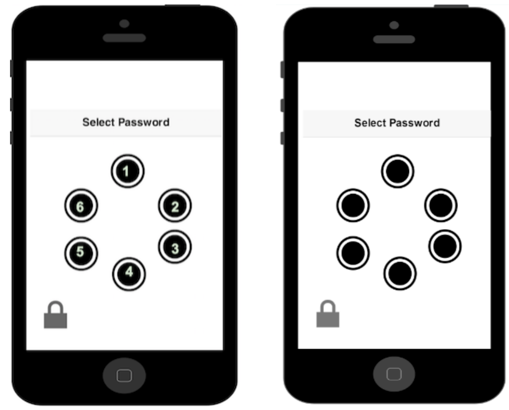
The first independent variable is the layout. We have four levels to compare: traditional digit keyboard versus gesture keyboard, circle grid versus square grid. The first level, we include two currently most popular ways of unlocking. Digit unlocking is usually seen on the iPhone while most Android phones use gesture to unlock. However, when it comes to the small screen on smart watch, the big finger effect might significantly affect user experience. The second level we present different layout of keys. Whatever digits keys or gesture keys are, people get used to the old-school way of dial number keypad. Yet, same problem as indicated above, small screen might make such layout uneasy to touch.

The second independent variable is how big the watch round is appropriate to be, which we basically focus on the size of the Apple iWatch. Typically, we compare two sizes of traditional iWatch, 312 x 390 for 42mm Version, 272 x 340 for 38mm Version. We expect different size of the screen would be another issue largely influencing the performance of unlocking screen. Combined with the layout, we hopefully find the best combination from our experiment. As for the size of the button, we are going to use the existing size and button to eliminate the bias of not taking full advantage of the space.



(a) (b)

We decide to use within subject design. Since our experiment requires record as many as possible to better averaging the performance of each combination and we are able to reach around 15 - 20 people in our lab, within-subject could help us accumulate more samples. Also, this method makes it easier to detect the difference between conditions. Since this method



(c) (d)

Figure 1. Screen unlocking methods. (a) numeric PIN on square grid. (b) graphical pattern on square grid. (c) numeric PIN on circular grid. (d) repetitive selection on circular grid.

takes longer time for experiment and might make user fatigue, we will strictly control the time within 45 minutes.

We would invite around 16 candidates for this experience and give them a period of time to practice. Then we will let them try every combination for at least 10 times to collect adequate statistic data. To eliminate the order effects, we will divide the participants into 8 groups and use 8*8 Latin Square for the experiment. Orders of passwords/patterns differ between different groups.

PART 5: Tasks and Procedures

A pilot study will be conducted before the experiment to make sure the procedures are reasonable. For the experiment, we will first introduce what the project is about and what the participant will be asked to perform during this experiment. All the participants will be requested to sign a consent form for allowing us to use their experiment data. Each participant will start from a training phase on the unlocking methods that he or she is unfamiliar with. A video or a word document of instructions will be provided. A password list containing 10 different combinations of password will be used as testing dataset.

When logging starts, participant will be asked to enter the password set on each combined method at the speed they are comfortable with. The sequence of combined methods will be different between groups to eliminate the impact of order. Time used for each trial as well as whether the participant enters password correctly will be recorded by the built-in function to calculate the average entering time and the error rate for each design.

Another experiment will be conducted to testify the security level. Video will be recorded of volunteer entering password sets on all eight designs. Camera will be placed at 20 cm over volunteer's left shoulder from the same angle. Participants will be asked to write down the passwords or draw down the patterns they see. Percentage of correct guessing over all trials for shoulder surfing will be the measured as an estimation of security level.

At last, participants will be invited to fill out a questionnaire about their feeling of each unlocking method.

PART 6: Measures

User experience in a screen unlocking method is affected by the convenience, i.e. the accuracy and efficiency, and security of the method. Users hope to unlock screen easily and fast with minimum errors, whereas maintaining high resistance to unauthorized access. Therefore, a tradeoff between security and convenience should be made. In this project, three dependent variables will be analyzed, i.e. accuracy, efficiency, and security.

Accuracy of each unlocking method will be evaluated by the average accuracy rate of the input passwords/unlock patterns obtained by the participants. An app will be developed which records whether or not the participants have entered the right passwords/patterns displayed on the screen previously.

Efficiency refers to how fast users can input passwords/patterns. To evaluate the input efficiency, the app will be engaged to log timestamps so the cost of time for each input of passwords/patterns will be recorded. The efficiency of a specific unlocking method will be estimated by the average time cost of passwords/patterns input for all participants.

Security is determined by how easily the password/unlocking pattern is replicated by unauthorized users. It will be estimated by error rate of shoulder surfing. A camera will be placed 20 cm above the participant's shoulder in the back to record the process of unlocking the screen. A second participant will watch the video and write down the password/unlocking pattern he or she observes. The security level of each unlocking method will be estimated by the average error rate of replicated password/unlocking pattern obtained from the shoulder surfing experiment.

PART 7: Data collection

Because we have three dependent variables in part 6, the method of data collection varies. In collecting the data for accuracy (correctness) and efficiency (time), we will extract the data from the mobile simulation application directly. In the application we will develop, the correctness of an unlocking trial and how long a person uses on the try will be recorded in the log file. The data can be generated in a plain text format for the future investigation.

As for the data collection of security (the easiness of unauthorized replication), we will ask participants to draw the unlocking pattern or write down the pin on the paper we prepared--they do not need to actually replicate the password on a simulated smartwatch. We will check if the replication is successful and record this result. The reason we use this method is we want to decrease the impact of the familiarity of smartwatch of a participant in this experiment. For this specific dependent variable, the success of the replication is the most important factor. The use of paper could also help us to decrease the difficulty of data mining and bring the coherence.