Understanding Environmental Influences on Performing Password-based Mobile Authentication

Sergey V. Maydebura, Dong Hyun Jeong, and Byunggu Yu
Assurance Research Center for Trusted Information Computing (ARCTIC)
University of the District of Columbia
4200 Connecticut Avenue NW, Washington, DC 20008, USA
dc.20007@yahoo.com, djeong@udc.edu, and byu@udc.edu

Abstract

In a mobile environment, text-based passwords are still the most common mechanism for user authentication. Although various studies have been conducted to investigate what password composition policies are better oriented for mobile users, a limited study has been performed to understand the impact of password composition policies and environmental settings to mobile users' password typing abilities. In this paper, we present a study that investigates password strength, user behavior, and user sentiment across two password composition policies under two environmental conditions such as stationary (sedentary position) and on-the-go (while walking). From the study, we correlate our results with user behaviors under different environmental conditions to provide suggestions for password-composition policies for mobile-based authentication.

1. Introduction

Password-protected user accounts became popular with the advent of the time-sharing systems and through the development of the services based on inter-computer exchange of information. Early user accounts were open to password guessing attacks due to a lack of mandatory use of strong password composition policies. In 1988, the ARPANET community experienced a massive attack handled by the Internet Worm [2] the key intrusion mechanism of which, according to [8], involved attempts to discover user passwords with utilizing a dictionary as a potential source of user passwords. This incident guided us to consider adapting stronger password composition policies for strengthening and ensuring that new passwords are not exposed to guessing attacks.

To identify the effects of utilizing password composition policies, numerous studies have been performed on identifying users' password practices [5], the strength of resulting passwords [3], and the daily use of passwords [4, 5]. However, there is a lack of study on understanding the effect of password practices in a mobile environment. One of the primary differences between desktop and mobile environments is that mobile users are not bounded to a particular location and settings, therefore, the users are free to utilize their mobile devices to access and use password-protected services (e.g. online banking, email services, etc.) anytime and anywhere.

This paper investigates the effectiveness of using text-based passwords maintaining two entropy levels (18 bits and 24 bits) under two different environmental conditions as stationary (stable) and on-the-go (unstable). Our study involved six participants performing tasks of typing passwords on a designed mobile application. With the application, we measured the speed and accuracy of the users' password typing actions, and 3-axis accelerometer data required to track the users' behaviors (i.e. movements).

2. Previous Work

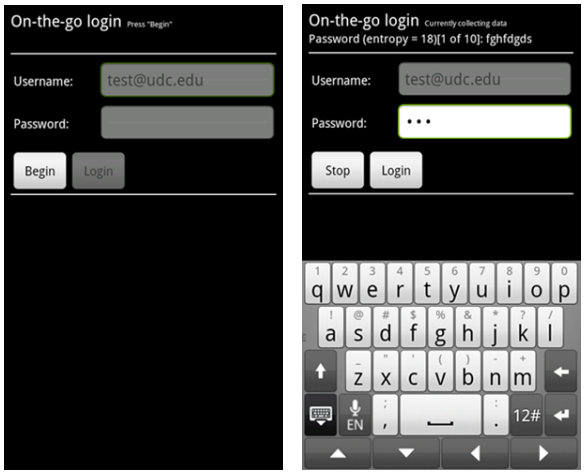
Numerous studies have been performed to understand the effectiveness of using passwords. Inglesant and Sasse [5] designed a structured diary study to understand the impact of utilizing unusable password policies within organizations. Although they cannot capture accurate measures of workload or time taken in password use, they identified that strict password composition and frequent password update policies might result in frustration and inability for users. Kelley et al. [6] identified that the most commonly used approaches of quantifying the effect of password-composition policies are (1) estimating the entropy of passwords using National Institute of Standards and Technology (NIST) guidelines [1] and (2) empirically analyzing passwords created under different password-composition policies with password-guessing tools. Komanduri et al. [7]

conducted a two-part online study using Amazon's Mechanical Turk service by recruiting 5,000 participants to understand the relationship between password-composition policies and the strength of the resulting passwords. From the study, they found that there is a negative correlation between entropy and usability. With a high-entropy composition policy, users face a difficulty of remembering passwords and tend to store them either on paper or electronically, therefore, increasing the vulnerability of their computing systems. They also found that a 16-character minimum with no additional requirements provides the most entropy and are more usable than other password composition policies.

Hayashi and Hong [4] examined password usage in daily life. By quantitatively analyzing frequencies of using passwords to log into computers for over two weeks, they observed that most participants reuse their passwords for multiple online accounts. They also found that most participants believed that writing down important passwords is risky, but did not realize that reusing passwords is also risky.

3. Mobile Application

To perform our study, we developed a mobile application (see Figure 1).



(a) (b)

Figure 1. An experimental application is designed to for Android OS mobile phone. By pressing the "Begin" button in the initial layout (a), users are able to type-in the password using a virtual keyboard (b).

The application displays predefined, yet randomly generated, passwords to the user and records all actions performed on the mobile device. It has a capability of capturing the user interaction (i.e. keys pressed), timings, and accelerometer data. Figure 1(a) shows the initial layout of the developed application, which includes two input boxes and two buttons. If the "Begin" button is pressed, one of the generated passwords is posted on screen. At the same time, the "Begin" button is replaced with a "Stop" button and a virtual keyboard becomes visible (Figure 1(b)). The study can be terminated at any time when the "Stop" button is pressed by the user. After typing the password in the password input box, the user can press the "Login" button to commit the password. If the typed password is incorrect, the user is required to retype the password. If successful, the other password is prompted on screen. For each authentication event, the application creates three log files to capture key pressed information, accelerometer data, and time spent.

4. Evaluation Procedure and Results

PROCEDURE Our experiment comprises a simulation of mobile authentication events and collection of corresponding data such as timings, input accuracy, and accelerometer data for each event. The primary target of our interest is to identify how user authentication can be affected by different (1) password composition policies and (2) environmental settings. For the study, 2 sets (18-bit and 24-bit entropy) of 5 passwords(Table 1) were randomly generated by following the NIST (National Institute of Standards and Technology) password composition rules [1]. With the passwords, participants were requested to perform password typing activities within two environmental conditions (i.e. stable and unstable). For the stable condition, we asked our participants to input passwords while seated in a distraction-free, quiet, and familiar environment. For the unstable condition, the participants were asked to input passwords while walking. Each participant was required to repeat the procedure of entering the same 2 sets of passwords 6 times: 3 times in the stable condition and 3 times in the unstable condition (the conditions were interchanged every 2 password sets).

In the study, both accuracy and speed were measured quantitatively by analyzing the captured log files. Specifically, accuracy is measured in identifying the total number of mistakes while typing in passwords and the total number of incorrect login attempts. The total number of mistakes is determined by measured the frequency of the "Backspace" button pressed.

RESULTS From the data analysis, we found that the overall time spent ($M = 106.90sec., SD = 63.26$) for completing each task (with 18-bit and 24-bit passwords) was substantial, $F(1,5) = 20.73, p < .01, \eta_p^2 = .16$. Number of login attempts ($M = 0.70, SD = 1.16$) was significant ($F(1,5) = 4.28, p < .05, \eta_p^2 = .05$) depending on the the type of passwords. However, the number

Table 1. Two types (18bits vs. 24 bits) of randomly generated passwords were used in the study.

Entropy	Password 1	Password 2	Password 3	Password 4	Password 5
18 bits	fghfdgds	jkermcnd	ldfnsdlf	vvkslsfr	mcmdnjjf
24 bits	hD1r)f*j	d*j3a&mC	mV!cr&n4	hG7!gdc*	ft6y&&Nc

of mistakes ($M = 1.90, SD = 2.62$) was not significant, $p = .72$. As shown in Figure 2(a), we found that participants made slightly more mistakes when using 18-bit passwords compered to 24-bit passwords. Although there is no statistical correlation between the number of mistakes and time spent, they spent less time with 18-bit passwords ($M = 81.70sec., SD = 48.54$) than with 24-bit passwords ($M = 132.11sec., SD = 65.31$). Initially, we assumed that participants might spend more time with 24-bit passwords because they are more difficult to type than 18-bit passwords. However, we found an unexpected result, participants had slightly more login failures with 18-bit passwords ($M = 0.97, SD = 1.40$) than with 24-bit passwords ($M = 0.44, SD = 0.77$).

By analyzing the data depending on the conditions (i.e. stable or unstable), time spent ($p = .075$), number of mistakes ($p = .40$), and number of failed login attempts ($p = .25$) were insignificant. However, there are distinctive results depending on the conditions. Although there was no correlation between the number of typing mistakes and the number of failed login attempts, we found a trend that our participants tend to spend more time in the unstable condition ($M = 118.35sec., SD = 55.45$) than in the stable ($M = 95.46sec., SD = 69.10$). Even though they spent more time in the unstable condition, they still made more typing mistakes and login failures (Figure 2(b)).

In addition, captured accelerometer data was analyzed to identify the primary factor that acts as a distraction component, which comprised the main difficulty for the participants to perform mobile authentication precisely. From the accelerometer data analysis, we were focused on extracting RMS power delivered to the mobile device and correlating it with three outcomes: time, number of typing mistakes, and number of unsuccessful logins. The RMS power from raw accelerometer data was extracted according to the following procedure:

1. Normalize accelerometer data by excluding the direct current (DC) component.

2. Calculate the power for each of three axes: $P = (\int_{-\infty}^{+\infty} |a(t)|^2 dt) \times m/t$, where $a(t)$ is time domain accelerometer data on x, y, or z axes, t is the time of data acquisition, and m is the mass of the mobile device (final units are joules/second).

3. Calculate the total RMS power:

$$P(RMS) = \sqrt{\tfrac{1}{3} \times (Px^2 + Py^2 + Pz^2)}.$$

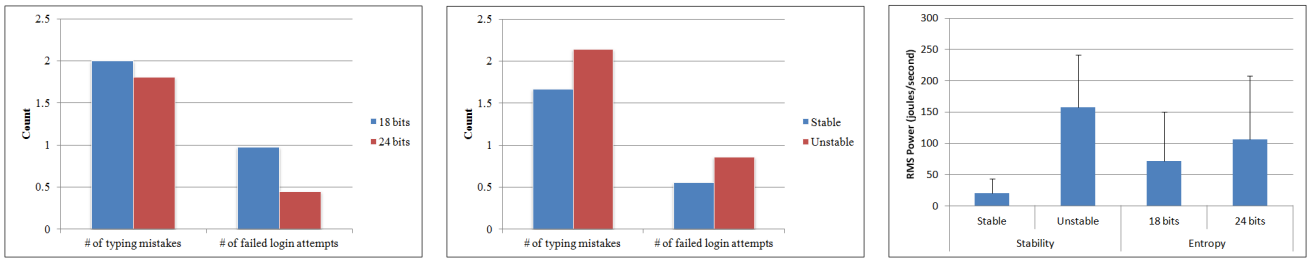
When analyzing the accelerometer data depending on the password entropy (16 bits vs. 24 bits), the RMS power was insignificant, $p = 0.10$. However, it was quite substantial when analyzing the RMS power depending on the conditions (stable vs. unstable), $F(1, 5) = 198.14, p < .01, \eta_p^2 = .56$. After calculating Pearson's correlation coefficients, we found that the total RMS power for all participants was positively correlated to the completion time ($r(72) = 0.48, p < .01$) and the number of typing mistakes ($r(72) = 0.24, p < .05$). Only in performing login attempts, the correlation was weaker ($r(72) = 0.22, p = .06$).

We expected that the RMS power will be higher in the unstable condition because participants performed the given tasks while in locomotion, which causes an increased level of hand unsteadiness. However, as shown in Figure 2(c), the RMS power did not significantly depend on the level of password strength, ($r(72) = 0.19, p = .10$), however, the power was slightly higher when using 24-bit passwords. This is because participants were required to perform additional key strokes to switch virtual keyboard layouts - 24-bit passwords include two special characters (not included in the main keyboard layout) and at least one upper character.

5. Discussion

Initially, we assume that password typing capability is directly proportional to users' physical environments. If users are walking or running, their ability to type passwords will be degraded due to situational impairments (e.g. walking vibration and divided attention, etc.). However, despite the fact that the experiment was conducted according to task-oriented evaluation protocol, our study includes limitations, and therefore, we cannot conclude that environmental condition takes an important role of causing numerous password typing failures. In the mobile environment, users use their memorized passwords to perform authentications. Since this memorization process requires a long-term study with having numerous experimental conditions, we limit our study on understanding the users' ability in performing password typing in different conditions (i.e. stable and unstable) without requiring them to memorize passwords, instead the participants were prompted with passwords during each authentication event.

Although the order of the passwords were counterbalanced, we found a learning effect. In the first 2 trials,



(a) (b) (c)

Figure 2. Evaluation Results. (a) Participants made slightly more typing mistakes with easy passwords (i.e. 18-bits), (b) participants made less login failures in the stable condition, and (c) measured RMS powers were high in the unstable condition and when using difficult passwords (i.e. 24-bits).

participants significantly were getting used to the application as well as the mobile device. In addition, they started to remember particular sections of certain passwords unintentionally. This often happened when they were using easy passwords (i.e. 18-bit passwords). Finally, they got a high accuracy in the third trial. Therefore to type a whole password, the participants experienced less attention distraction caused by transferring the sight from the on-screen keyboard to the section of the screen where passwords were being displayed.

6. Conclusion and Future Work

In this paper, we performed an experimental study to find answers for the environmental influence to the performance of typing passwords and the effectiveness of using password composition policies. There was no statistically significant relationship to the overall performance of typing passwords. However, participants spent more time in the unstable environmental condition. Although they spent more time, they still made more mistakes and more login failures in the unstable condition. With strong passwords (24-bits), participants spent more time, as we expected. But, we found a statistically significant result in making fewer mistakes and having less login failures with using 24-bit passwords.

As a future work, we plan to investigate more on identifying factors causing high accuracy in performing password-based mobile authentication with 24-bit entropy passwords. To understand better about the environmental influence to mobile authentication, we plan to extend our study having more environmental conditions such as walking, running, and riding a bus, bicycle, or escalator. Since people unintentionally slow down the walking speed when typing special characters, it is also important to isolate factors by capturing more data.

7. Acknowledgement

This work was supported by a grant from the National Science Foundation (Grant NSF Award No 0911969).

References

[1] W. E. Burr, D. F. Dodson, E. M. Newton, R. A. Perlner, Polk, S. Gupta, and E. A. Nabbus. Electronic authentication guideline. *NIST Special Publication 800-63*, January 2006.

[2] P. J. Denning. The internet worm. Technical Report TR89-3, Research Institute for Advanced Computer Science, NASA Ames Research Center, February 1991.

[3] P. Dourish, E. Grinter, J. Delgado de la Flor, and M. Joseph. Security in the wild: user strategies for managing security as an everyday, practical problem. *Personal Ubiquitous Comput.*, 8(6):391–401, Nov. 2004.

[4] E. Hayashi and J. Hong. A diary study of password usage in daily life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2627–2630, New York, NY, USA, 2011. ACM.

[5] P. Inglesant and M. A. Sasse. Studying password use in the wild: practical problems and possible solutions. In *SOUPS 2010: Workshop on Usable Security Experiment Reports*, 2010.

[6] P. G. Kelley, S. Komanduri, M. L. Mazurek, R. Shay, T. Vidas, L. Bauer, N. Christin, L. F. Cranor, and J. Lopez. Guess again (and again and again): Measuring password strength by simulating password-cracking algorithms. In *Proceedings of the 2012 IEEE Symposium on Security and Privacy*, SP '12, pages 523–537, Washington, DC, USA, 2012. IEEE Computer Society.

[7] S. Komanduri, R. Shay, P. G. Kelley, M. L. Mazurek, L. Bauer, N. Christin, L. F. Cranor, and S. Egelman. Of passwords and people: measuring the effect of password-composition policies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2595–2604, New York, NY, USA, 2011. ACM.

[8] E. H. Spafford. The internet worm incident. Technical Report CSD-TR-933, Purdue University, September 1991.