Improving user authentication on mobile devices: A Touchscreen Graphical Password

Hsin-Yi Chiang
Carleton University
hychiang@scs.carleton.ca

Sonia Chiasson
Carleton University
chiasson@scs.carleton.ca

ABSTRACT

Typing text passwords is challenging when using touchscreens on mobile devices and this is becoming more problematic as mobile usage increases. We designed a new graphical password scheme called ***T**ouchscreen **M**ulti-layered **D**rawing* (TMD) specifically for use with touchscreens. We conducted an exploratory user study of three existing graphical passwords on smart phones and tablets with 31 users. From this, we set our design goals for TMD to include addressing input accuracy issues without having to memorize images, while maintaining an appropriately secure password space. Design features include *warp cells* which allow TMD users to continuously draw their passwords across multiple layers in order to create more complex passwords than normally possible on a small screen. We compared the usability of TMD to Draw A Secret (DAS) on a tablet computer and a smart phone with 90 users. Results show that TMD improves memorability, addresses the input accuracy issues, and is preferred as a replacement for text passwords on mobile devices.

Author Keywords

authentication; touchscreen; mobile devices

ACM Classification Keywords

K.6.5 Computing Milieux: Security and Protection – Authentication

General Terms

Experimentation, Human Factors, Security

INTRODUCTION

Mobile devices have the ability to connect to the internet and access various personalized remote services. These services often ask users to authenticate using text passwords; this requires typing on mobile devices. However, modern mobile device are heavily graphic-oriented and touchscreens are often the primary input method even for typing. Physical constraints like screen sizes make typing less accurate and less efficient than on physical keyboards [2]. Similar results are available for soft keys (simulated buttons on touchscreens) in general [8]. The question arises: *Given the physical constraints of mobile devices, is there an alternative type of authentication that can (1) be easily deployed, (2) improve usability, and (3) maintain security?* A potential alternative is to explore the use of graphical passwords. A graphical password [3] is a secret that is entered or displayed in the form of drawings, icons or graphics.

MobileHCI'13, August 27 - 30 2013, Munich, Germany

http://dx.doi.org/10.1145/2493190.2493213

This paper includes two main contributions. First, it provides an evaluation of exemplar schemes from the three main categories of graphical passwords on both tablet computers and smart phones. Second, it offers a new authentication scheme specific for mobile devices with improved usability.

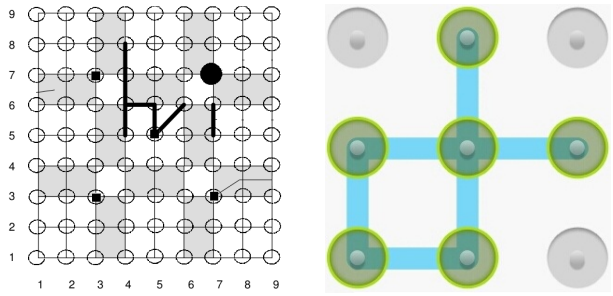
The initial exploratory study with 31 users was conducted to identify problems encountered when using graphical password schemes on mobile devices. We chose a representative scheme from three main categories [3]: *recall*, *cued-recall*, and *recognition*. We identified their strengths and weaknesses when applied to mobile devices. Based on our findings, we propose a new graphical password scheme designed for use with touchscreens; we call this scheme ***T**ouchscreen **M**ulti-layered **D**rawing* or *TMD*. TMD is a user-drawn graphical password scheme using a grid of large detached cells intended to address the identified issues with input accuracy. We used multiple layers of grids to encourage password complexity.

We evaluated the usability of TMD on tablets and smart phones by comparing it with DAS, and found favourable results for TMD. In our 90-user study, TMD showed superior memorability one week after password creation and TMD avoided the input accuracy issues common in DAS. User responses indicated that participants were more willing to use TMD than DAS to replace text passwords on mobile devices.

BACKGROUND

User authentication must be both secure and usable in order to be successful. Security issues lead to compromised accounts, leaked private information, and potential losses for individuals or organizations. Usability problems lead users to develop insecure coping strategies, circumvent security mechanisms altogether, or turn to less obtrusive alternatives. To address these issues, a number of potential alternatives have been proposed for user authentication. Of these alternatives, graphical passwords seem especially well-suited for the mobile environment because of their graphical nature. Our background discussion summarizes this line of research.

A detailed survey of graphical passwords is available from Biddle et al. [3]; we include only a brief overview here. Graphical passwords schemes are classified into three categories according to the type of memory task involved: recall,



(a) Pass-Go [15] (b) Android pattern unlock

Figure 1: Recall-based graphical password schemes

recognition, and cued-recall. Recall schemes require recreating drawings without hints (e.g., Draw A Secret [10]); recognition schemes require recognizing previously memorized objects from among decoys (e.g., Passfaces^{TM} [12]); cued-recall schemes require selecting a sequence of previously-chosen points on a provided visual cue such as an image (e.g., Persuasive Cued Click-Points [4]).

In the following discussion, we use *password space* as a measure of password strength. In computer security, this refers to the set of all possible password combinations which a password scheme can produce. The total is converted to base-2 and reported as the number of bits.

Next, we describe the three recall schemes most directly relevant to our proposal. Draw A Secret (DAS) [10] allows users to draw their passwords on a two-dimensional grid using one or more lines (Figure 2a). The password space is 48 bits for DAS passwords with a maximum length of 10 cells on a 5×5 grid [10]. Later security analysis has shown that DAS is susceptible to guessing attacks [16] because users tend to draw predictable passwords. In an improved proposal, Background DAS (BDAS) [6] includes background images to provide visual cues to users and help them create more complex passwords. Pass-Go [15] is a related scheme that require users to draw their passwords by connecting the intersections of grid lines (Figure 1a). The password space of Pass-Go is 64 bits for passwords of maximum length of 10 on a 9×9 grid. Neither DAS nor Pass-Go are designed for mobile devices and require relatively precise input to enter the password.

The most well-known deployed graphical password is the Android screen unlock mechanism (Figure 1b) for mobile devices. The scheme uses the intersections of a 3×3 grid where users select a series of intersections with one smooth gesture. The password space of the Android pattern unlock is 18 bits [1] which is approximately equivalent to a 5 digit PIN.

Recall-based schemes bear some resemblance to gestural biometrics (e.g., [13]) since both require users to perform a gesture with their finger. However, graphical passwords are only concerned with whether the discrete drawing exactly matches the stored password and do not attempt to distinguish between different users who may enter the same gesture.

Possible Attacks on Graphical Passwords: All knowledge-based authentication schemes, including graphical passwords, are susceptible to guessing attacks. In *Dictionary attacks*, attackers guess from a list of higher-probability passwords. They are a threat when users chose their own passwords and patterns emerge from user choice. *Brute force guessing attacks* are also possible when the password space is small because an attacker can iteratively guess all possible password combinations. Mitigation strategies are possible such as having a lock-out policy for incorrect login attempts, employing salts and iterative hashing when storing passwords, and enforcing minimum password strengths.

Shoulder surfing occurs when an attacker obtains users' passwords by direct observation or using video recording devices. On mobile devices, shoulder-surfing is a threat for text [14] and graphical passwords [5], although the severity of this threat varies considerably with context of use.

Aviv et al. [1] examined the feasibility of *smudge attacks* on touchscreen devices using the Android unlock screen. Smudge attacks analyze the path of oily residues left on touchscreens to determine passwords. While successful, this attack relies on gaining access to a user's device.

EXPLORATORY STUDY

Our overall goal is to design and implement a new graphical password scheme optimized for mobile devices with touchscreens. To have a better understanding of how touchscreens affect graphical passwords, we tested an existing scheme from each of the three different categories (recall, recognition, and cued-recall). Table 1 summarizes the three test schemes selected and Figure 2 shows their user interfaces.

We used two multi-touch devices to assess whether screen size impacts the usability of the schemes. We used a 4th generation *iPod touch* (3.5-inch, 960×640 pixels resolution display) and a 3rd generation *iPad* (9.7-inch, 2048×1536 pixel resolution display). Since no phone functionality was needed, an iPod Touch serves as a suitable phone proxy. The password schemes were implemented using JavaScript and Scalable Vector Graphics (SVG) technology and displayed using the built-in Safari browser on both test devices.

According to Florencio et al. [7], a 20 bit password space with additional login rules (e.g., limiting the number of retries) is considered sufficient defense against web attackers for most applications. As show in Table 1, we implemented minimum password length restrictions to ensure that the schemes complied with their recommendations.

Participants

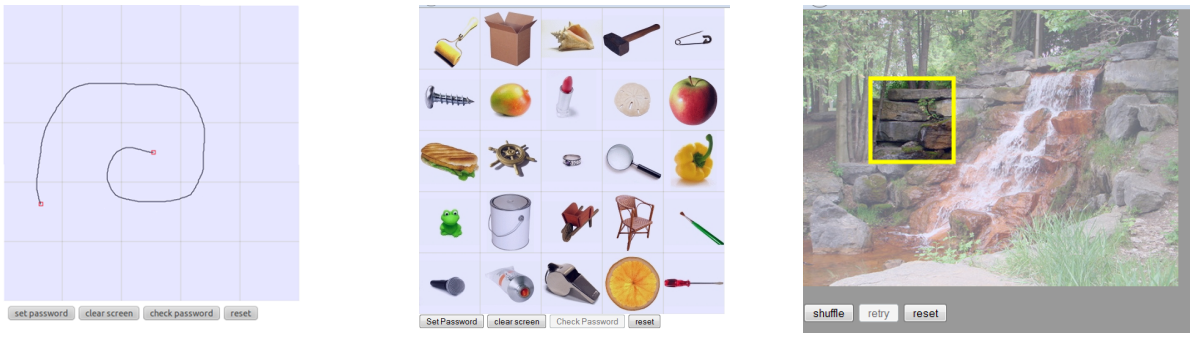
We recruited 31 participants (11 female, 20 male, mean age = 24.8 years) from various fields on the university campus. Twenty-four participants owned at least one mobile device. Fifty one percent of participants had seen or used a graphical password before, and almost all their experience was based on the unlock screen for Android.

Protocol

This experiment was approved by our institutional review board. Each participant completed an individual one-hour-session. The study used a mixed design: participants were

Category	Scheme Name	Basic Operation	System Configuration
Recall	DAS (Draw A Secret) [10]	Users draw their password on a grid	Uses a 5×5 grid (75×75 pixels per cell) with a minimum password length of 5 cells; each cell can be used multiple times. The password space is 23 bits.
Recognition	Object Recognition [9]	Users select their previously memorized icons from among decoys	Passwords consist of at least 8 icons from a panel of 25 icons (75×75 pixels per icon), the order of selection matters. The password space is 20 bits.
Cued-recall	PCCP (Persuasive Cued Click Points) [4]	Users select points on a series of images. To create a password, users choose each point from a randomly positioned viewport.	Passwords consist of 5 points. Images are 451×331 pixels each; the viewport is 100×100 pixels. The acceptable tolerance region for points is 25×25 pixels. The password space is 39 bits.

Table 1: System configuration for the three graphical password schemes in the exploratory study



(a) DAS (b) Object Recognition (c) PCCP

Figure 2: Interfaces for the graphical password schemes used in the exploratory study

randomly assigned to either the smart phone (15 participants) or tablet (16 participants) group; both groups tested all three password schemes. We used a Latin Square to assign the presentation order of the schemes. The experiment consisted of the following steps, repeated for each scheme:

1: Introduce scheme. Users received verbal instruction on how to use the scheme.

2: Learn scheme. Users practiced using the scheme on a desktop computer. A practice set of images were used for *Object Recognition* and *PCCP* to avoid memory interference. Users could explore until they were comfortable. The intention was to familiarize users with the scheme so that usability problems arising during testing on the mobile device were in fact due to the form factor.

3: Create password. Moving to the mobile device, users created a password. If using *PCCP*, users could shuffle the viewport as many times as they wished.

4: Confirm password. Users confirmed the password by re-entering it. If unable to confirm, they returned to step 3.

5: Answer questionnaire. On the desktop computer, users answered an online questionnaire providing feedback about the scheme.

6: Login. On the mobile device, users re-entered their password. They could retry as many times as needed. If unable to login, users could stop and move on to the next step.

7: Answer questionnaire. On the desktop computer, the users completed a second questionnaire about their perceptions of the scheme.

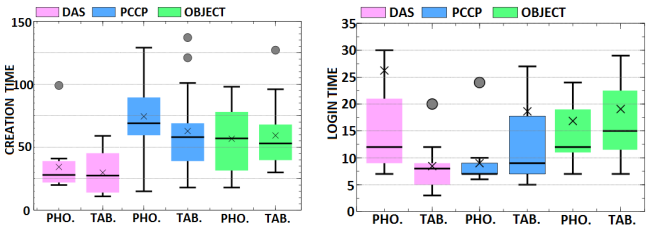
At the end of the session, users completed a questionnaire on demographics and past experience with mobile devices.

Results

We descriptively and statistically compared different data to identify usability issues and user preferences when using graphical passwords on mobile devices. We mainly focus on determining how screen size affects performance by comparing differences between devices on creation time, login time, login success rate, and password length for each scheme. For the creation time, login time, and password length, we used mixed-design ANOVAs to look for overall differences and t-tests to determine where the difference occurred; Fisher's Exact tests were used for the login success rate. The Likert scale questions were analyzed using Mann-Whitney U tests (independent ordinal data). The alpha value was set to 0.05.

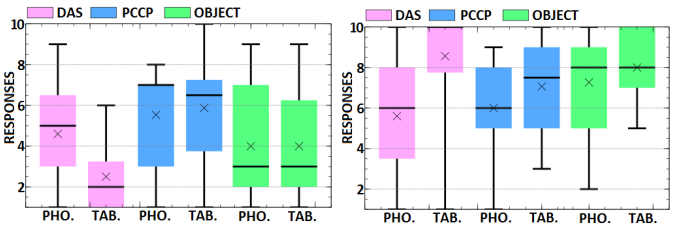
Creation Time

The creation time (Figure 3a) measured the time between the first touch event and pressing the submit button. A mixed-design ANOVA showed no main effect of form factor, but showed a significant effect for the types of scheme ($p < 0.000$, $F = 17.928$, and $df = 1$). Paired t-tests show a significant difference between the DAS-Object Recognition pair ($p < 0.001$, $t = -5.468$, and $df = 30$) and DAS-PCCP pair ($p < 0.001$, $t = -6.978$, and $df = 30$); no significant



(a) Password creation time (b) Login time

Figure 3: Times in seconds for the exploratory study.



(a) Difficult to create (1 = +ve) (b) Easy for screen size (10 = +ve)

Figure 4: Likert scale responses for the exploratory study.

	DAS		PCCP		Object	
	Tab	Ph	Tab	Ph	Tab	Ph
First try	100	64	87	85	87	85
≤ 3 tries	100	93	100	100	100	100

Table 2: Exploratory study login success rates (percentage).

difference were found in PCCP-Object Recognition pair. In summary, we found no evidence that screen size affects creation time, but did find differences between schemes. DAS passwords can be set up most quickly (median of 27 seconds) while the other schemes take significantly longer.

Password Length

Users could choose the length of their password in DAS and Object Recognition. A mixed-design ANOVA showed a main effect of form factor ($p = 0.031$, $F = 5.153$, and $df = 1$). Further analysis showed that only DAS password lengths were affected by form factor ($p = 0.016$, $t = -2.553$, and $df = 29$); tablet users created significantly longer DAS passwords (tablet mean = 16 cells, phone mean = 10 cells). We suspect that the small screen made it difficult to draw DAS passwords accurately, so users created shorter passwords. Object Recognition can be operated by pressing virtual buttons and was not affected by screen size (mean = 8 icons per password on both devices). We further noticed that DAS users generally created passwords that were much longer than the minimum required length while Object Recognition users just met the minimum requirement.

Login Time

Figure 3b shows the login time for successful attempts. Other than DAS on the phone, users took less than 20 seconds to login successfully on all schemes. For DAS and Object Recognition, the time of a login attempt is measured from the first touch event until the *check password* button is pressed; this includes any time spent clearing and starting over. In PCCP, it is possible for users to notice an erroneous click before completing the login process. Therefore, when measuring the login time for PCCP, we treated user resets as failed attempts. A mixed-design ANOVA test indicated no main effect for either the type of scheme nor form factor for login time.

Login Success Rate

Table 2 shows percentages of successful logins. Fisher's exact test showed that DAS on the tablet has a higher login success rate than the phone ($p = 0.04$) on the first try and no significant differences were detected for the other schemes. No schemes showed significant differences for the login success rate within three tries. DAS users found it difficult to accurately draw on the small screen. It is unlikely that users forgot their password because the time between the password creation and login is very short and most users could correctly enter the passwords within three tries.

Questionnaire Responses

We compared smart phone versus tablet computer users' questionnaire responses for each scheme. Likert scale response ranged from 1 (strongly disagree) to 10 (strongly agree). Consistent with the previous results, we observed that screen size influences users' experiences only for DAS.

Figure 4a shows a box plot for the question *I find it hard to create a graphical password using this scheme*. Participants thought that DAS was more difficult to use on the small screen ($p = 0.009$, $U = 54.5$, and $z = 2.606$), but no statistically significant differences were found for PCCP or Object Recognition. Other than DAS on tablet which scored an average of 2.5, other schemes on both devices scored between 4.0 and 5.8, indicating that participants had some overall difficulty creating passwords.

Figure 4b illustrates responses for *This scheme was easy to use given the size of the device screen*. Participants felt that DAS is more difficult to use on the smart phone than tablet computer ($p = 0.004$, $U = 49.0$, and $z = 2.842$) while no statistically significant differences were found for PCCP and Object Recognition. Except for DAS on smart phone, other scheme-device combinations had averages between 6.0 and 8.5 indicating that participants felt that the schemes are somewhat usable.

Observations and User Feedback

We observed users and noted relevant behaviours and feedback. Users also had the opportunity to provide open-ended feedback on the questionnaire. In general, they appreciated not having to type their password, but identified opportunities for improvement for each of the schemes.

DAS: Several users complained that the phone canvas was too small, making it difficult to draw their secrets accurately. They wanted an easier way to draw diagonal lines and had difficulty remembering the drawing order of their sketch. Several users also chose passwords that had "fuzzy boundaries": tracing grid lines or crossing cell via the corner (Figure 5).

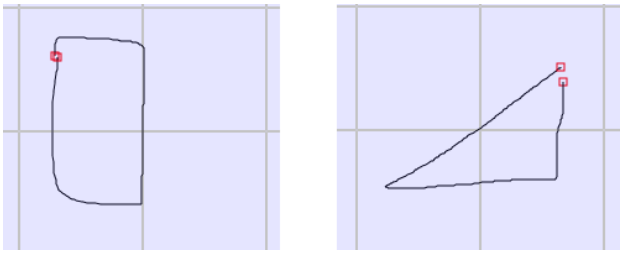


Figure 5: Passwords containing fuzzy boundaries

Since passwords are recorded as a series of grid cells, users may be unaware which side of the grid line was used [6] and may not be able to exactly replicate their passwords.

PCCP: The main observation was about shuffling of the viewport. The concept of viewport is to encourage users to select more random points as their passwords. However, many users selected a point outside of the viewport then continuously clicked the *shuffle* button until the viewport moved to the desired location or tried to drag the viewport with their fingers because of their past experiences with touchscreens. Some mentioned that it was difficult to memorize 5 images and the points in a short time, while others mentioned that it was difficult to click on the points precisely.

Object Recognition: Users tended to pick their icons to form a semantic pattern. Most users tried to group icons based on properties such as size or color. They then selected groups of icons to help with memorability. The icons appeared sufficiently large to click on both devices, but some users said that it was difficult to differentiate similar icons on the small screen. They also mentioned that it was sometimes difficult to find objects that share common properties and difficult to memorize that many images.

DESIGN OF THE NEW TMD SCHEME

Based on our exploratory study, we concluded that a major problem with user-drawn graphical passwords on touchscreens is the accuracy problem. Furthermore, we also discovered that users disliked having to memorize unfamiliar images or icons. With that in mind, we designed a graphical password scheme that has larger target areas and does not force the users to remember additional images or icons.

Our ***T**ouchscreen **M**ulti-layered **D**rawing* (TMD) scheme (Figure 6a) consists of three types of "cells": *Unselected*, *Selected* and *Warp* cells. The edges of the cells do not touch each other to eliminate the fuzzy boundary problem but visual cues are included to tell the user which directions of motion are possible. To create a password, users choose any unselected cell as their starting point; they then can make the password longer by choosing any unselected adjacent cell as the next cell in the password. Once chosen, the color of the cell changes to indicate that the cell has been selected. Selection of the cells must be done within a single dragging gesture; lifting the finger from the screen indicates the selection process is complete. Users may cross already selected cells to reach further unselected cells, but the path through the already selected cells is not recorded. TMD allows users to draw their secrets across multiple "layers" by going through the warp cells at the four corners. Figures 6b and 6c illustrate how a user reaches a warp cell and gets transferred to the next layer. A grid size of 5×7 was chosen because this configuration maximized the number of cells while maintaining a 10×10 mm size. Lee and Zhai [11] determined that buttons smaller than this resulted in degraded user performance.

Similar to DAS [10] or Pass-Go [15], TMD passwords are encoded using an alpha-numerical representation of each cell. As shown in Figure 6d, each TMD cell is assigned a value. To encode a password, simply concatenate the alphanumerical text of each cell separated by a comma in the order which they were selected by the user.

USER STUDY OF TMD

We conducted a user study to assess TMD's usability on mobile devices.We selected DAS as the comparison scheme because it is the most well-known recall-based graphical password scheme. While in common usage, the Android unlock scheme was not selected because its password space is too small to provide meaningful comparisons. We tested our schemes on the same tablet and smart phone as the exploratory study. This experiment was approved by the university ethics review board.

Implementation

We implemented TMD and DAS as web pages using JavaScript and Scalable Vector Graphics (SVG) technology. The web pages recorded user activities during the experiment such as time spent to setup a password, time spent entering a password, number of retries/clears before sending the password for verification, the length of password created, and the date and time of each login attempt.

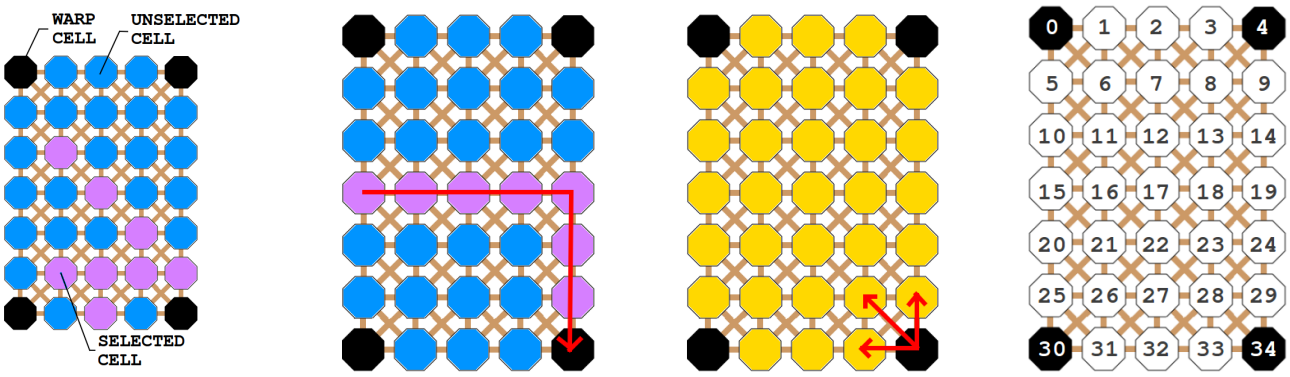
We set the minimum password length for TMD to be 10 cells and DAS to be 5 squares. The password spaces were 23 bits for TMD and 26 bits for DAS. Longer passwords are required for TMD because the interface does not record repeat passes on an already selected cell. This design decision was made so that users always had a visual cue (colour change) when a cell was selected. Both configurations meet Florencio et al.'s [7] suggested minimum password space requirements.

Participants

We recruited 90 participants (average age was 23 years). There were 13 PhD, 13 masters, and 59 undergraduate students; five did not answer the question. Seventy-two participants owned at least one smart phone for at least six months and 18 participants owned at least one tablet computer for at least six months. Fifty-seven participants had seen a graphical password before and 29 participants had previously used a graphical password scheme. Of these, almost all the participants reported experience with the Android screen unlock.

Protocol

The study used a 2×2 between-subjects design, with participants randomly assigned to one of four groups (TMD on



(a) TMD interface

(b) The User has reached a warp cell when entering the password

(c) The scheme displays the next layer and the user may continue the password on any of the 3 indicated cells

(d) TMD Encoding

Figure 6: Design of TMD

phone, DAS on phone, TMD on tablet, or DAS on tablet); there were 22, 23, 23 and 21 participants per group respectively. This lab study had two 30-minute sessions scheduled 5-10 days apart.

Session 1

Participants were asked to create a password which they thought was safe and memorable. The session was carried out with the following steps:

1: Introduce the scheme. Users received verbal instruction on how to use the scheme. They were also reminded to remember the created password for the next session.

2: Create password. On their assigned device, users created a password. They could try as many times as necessary.

3: Confirm password. Users confirmed the password by re-entering it. If unable to confirm, they returned to step 2.

4: Answer questionnaire. Users answered a questionnaire consisting of Likert scale and open-ended questions to provide feedback regarding their opinion and perceptions of their assigned scheme.

Session 2

For Session 2, participants logged in using the password from the first session; they had three chances to enter the password correctly. The session included the following steps:

1: Login. On the mobile device, users logged in with the password from Session 1. If unable to login within three tries, users stopped and move on to the next step.

2: Answer questionnaires. Users completed questionnaires consisting of Likert scale and open-ended questions. Topics included the user's perception of the scheme, demographics, and past experience with mobile devices.

Results

We compare TMD and DAS's performance and user perceptions on several measures.

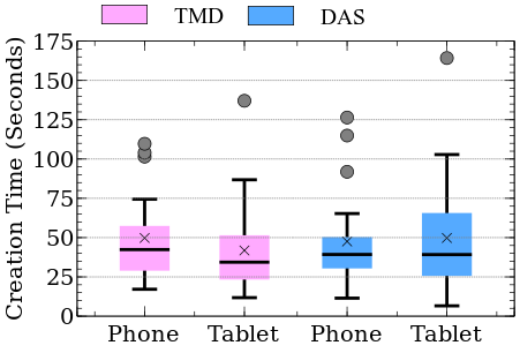


Figure 7: Password creation time (TMD and DAS)

Creation Time

Figure 7 shows the password creation times. The password creation time is measured from the time when the *Create* button is pressed until the password has been confirmed and sent to the database for matching. Participants in all groups could create a password within a minute. A two-way ANOVA showed no main effect of form factor or scheme.

Length and Depth

The average password length of TMD is 18 cells for the tablet and 17 cells for the smart phone. DAS passwords had an average length of 16 blocks for both devices. A two-way ANOVA indicated that there was no significant difference in length between the devices or schemes.

In the design of TMD, we incorporated the concept of *password depth*, i.e., the number of layers used in a password. The password depth is 0 at the initial state and incremented when users passed though a warp cell. On both devices, the average password depths are over 1 (span over two panels) which indicated that the users were actually using this feature when creating the passwords even though they were not explicitly encouraged to do so in our instructions. T-test results showed no significant difference in depth between devices.

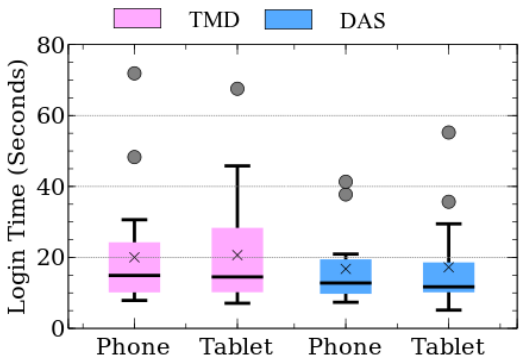


Figure 8: Session 2 login time (TMD and DAS)

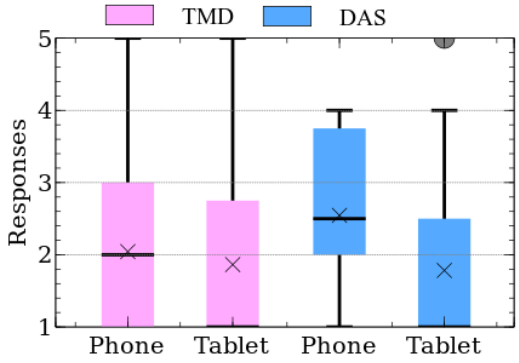


Figure 9: Screen size makes it hard to use (1 = +ve)

	TMD		DAS	
	Phone	Tablet	Phone	Tablet
First try	86	86	67	57
≤ 3 tries	100	95	71	71

Table 3: Session 2 login success rate (percentage)

Session 2 Login Time and Success Rate

Figure 8 shows the average login time of a successful attempt approximately a week after password setup. The time for a successful login attempt starts when the *login* button on the initial page is pressed until the password is sent for verification. A two-way ANOVA of login times showed no main effects for form factor or scheme. The average login time varied between 15 and 18 seconds. No participant complained that logging in took too long. We expect that times would decrease once participants became more familiar with their password.

Table 3 shows the percentage of successful logins in Session 2. A week after password setup, TMD shows a login success rate of at least 95% within the first 3 tries whereas DAS only has a 71% success rate. Fisher's Exact test showed that for logging in successfully without any mistakes (first attempt), TMD has a higher login success rate than DAS on the tablet computer ($p = 0.017$) but no significant difference when used on the smart phone. For logging in successfully within 3 tries, test results showed that TMD has a higher login success rate than DAS on both tablet computer ($p = 0.04$) and smart phone ($p < 0.01$). These results suggest that TMD passwords are more memorable than DAS passwords, especially if users are allowed a few attempts.

Questionnaire Responses

In both sessions, we asked users to complete questionnaires to provide us with feedback. We grouped the Likert scale questions into 3 categories for discussion: form factor, graphical password scheme, and password memorability. For all Likert scale questions, responses ranged from 1 (strongly disagree) to 5 (strongly agree).

Form Factor. For form factor, we compare responses between the smart phone and tablet computer groups for each scheme. Consistent with the exploratory study, participants felt that DAS is harder to use with a small screen. However, the size of the screen did not affect the participants' perception of TMD.

Figure 9 shows responses to the question *The size of the screen on this device makes the scheme hard to use*. Favourable responses are represented by lower scores. A Mann-Whitney U test indicated that DAS users generally found it more difficult to use the scheme on a small screen ($p = 0.021$, $U = 154.0$, and $z = 2.304$). No significant differences were found for TMD.

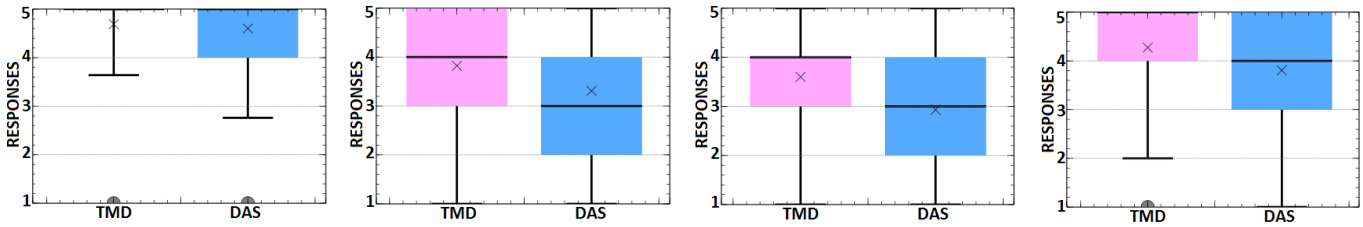
Graphical Password Scheme. The design of the graphical password scheme has a direct impact on users' perception and opinion. A good graphical password scheme should be easy to understand and use. In this section we compare responses between TMD and DAS regardless of form factor.

Figure 10a represents responses to *It was easy to understand how the scheme works*. Mann-Whitney U test showed no significant difference between TMD and DAS ($p = 0.517$, $U = 941.5$, and $z = 0.649$). The average scores for the two schemes are both 4.6, showing that participants found the schemes are easy to understand.

Figure 10b shows responses to *I am more willing to use this password scheme than traditional text-based passwords on this device*. A Mann-Whitney U test showed that users are more willing to use TMD than DAS ($p = 0.048$, $U = 771.0$, and $z = 1.979$) to replace traditional text-based passwords.

Figure 10c shows the box plot for *I would use this graphical password for my important accounts (e.g., online banking)*. The test results showed that participants gave TMD significantly higher scores than DAS ($p = 0.013$, $U = 673.0$, and $z = 2.496$). Users generally felt that TMD passwords could provide sufficient protection for their important accounts whereas DAS might not.

Memorability. Passwords that are easy to remember can reduce security risks such as writing down or reuse of passwords. Although actual password memorability has already been reported, the questionnaire questions can still provide us with useful information about users' perception. In this section we compare TMD and DAS regardless of device.



(a) Easy to understand (b) More willing to use than text (c) Use for important accounts (d) Easy to remember

Figure 10: User Responses to Likert-scale questions for TMD and DAS (5 = +ve)

Figure 10d represents responses to *It was easy to remember the password* which we asked the participants in session 2. Participants in both schemes felt that passwords were memorable. Mann-Whitney U test results showed that there is no significant difference between the two schemes. Although DAS participants thought that memorizing the passwords was easy, this is inconsistent with the login success rate discussed previously, which showed that they had significantly more difficulty than TMD users.

Observations and User Feedback

We recorded our observations and feedback from participants during the two sessions. This information gave us insight on how to improve the design of our scheme. Users also provided written feedback on the questionnaires. With both schemes, users found using the graphical scheme faster than entering text passwords and appreciated not having to type. Some participants said that while they could remember the shape of their password, they were not sure which cell was their starting point in the second session.

TMD. The first thing we noticed with TMD was the absence of the fuzzy boundaries problem. Participants tended to use more vertical and horizontal strokes in their passwords, but those who used diagonal lines had no problem confirming the passwords. Unlike DAS passwords which relied solely on shapes, some TMD users used the unique color from each panel to help memorize their passwords, providing us with positive feedback on this feature. Several appreciated being able to enter the password in one fluid motion.

A few items were also noted for improvement. Some found it difficult to think of a pattern using only one line. A couple of issues were related to using the mobile device itself: a few were worried about releasing their stroke too early and complained about friction from pressing so hard on the screen, while a few others accidentally activated the top panel in iOS when getting too close to the top of the screen.

DAS. Some participants mentioned that they liked being able to draw shapes as their passwords. Participants also noted problem areas. Some participants who used DAS on the phone complained that it was hard to draw the lines accurately because the line was covered by the participant's own fingers. As in the exploratory study, the fuzzy boundaries problem affected many participants and they could not correctly confirm their passwords. Participants also mentioned difficulty with drawing diagonal lines.

Interestingly, the things which the participant did not like about the two schemes were very different. In TMD, the complaints where more focused on the system-level problems. For DAS, the complaints centered on the design of the scheme and its use on a small touchscreen.

TMD password patterns and distribution.

When users select their own passwords, patterns tend to appear across users as some passwords are more popular than others. This is a security concern in both text and graphical password schemes because attackers could leverage patterns to prioritize dictionary guessing attacks.

We collected 45 TMD passwords and examined them for distinctive patterns, as described in Table 4. These patterns represent very broad categories and would also be present to some degree in random passwords. They do not uniquely identify specific passwords. Patterns may occur on a single layer or span multiple layers. Previous studies have already shown that DAS passwords are susceptible to patterns [16], so we focus on a preliminary analysis on TMD passwords.

We visually inspected each password and classified it into the most fitting category. Passwords that did not fall into one of the six identified patterns were classified as "abstract", meaning that they had no discernible pattern. For passwords which use multiple layers, we superimposed all the layers before inspecting the passwords. If a password qualified for more than one category, the higher priority category as specified in Table 4 claimed the password.

Figure 11 shows the distribution of TMD patterns. About two thirds of passwords fell into one of the broad pattern categories. While not immediately identifiable by attackers since a large number of passwords fit each category, this issue will need further exploration in TMD. We believed that by adjusting the password policy of TMD and providing guidance on password selection, we may be able to reduce this risk.

We also looked at the distribution of starting points for the passwords. Similar to the password patterns, the distribution of the starting points could be leveraged to create a more efficient list for dictionary attacks. Figure 12 shows the distribution of the 45 starting points collected. Given the number of cells, we would expect approximately half of starting points to be on an edge if the distribution was random. From our analysis, 93% of participants picked starting points along the edges of the grid. We further see a concentration of starting points next to warp cells. The lack of visual reference on

Category	#	Definition
Recognizable symbols	1	Symbols which are well known to others such as the English alphabet, Arabic numbers, or mathematical operators
Back-trace	2	The path of the password on the current layer is identical to the previous layer but the direction of the path is reversed
Recognizable patterns	3	Recognizable paths such as spiral or zigzag patterns
Symmetric	4	Shapes that are symmetric about an axis, the axis can be vertical or horizontal
Along the edges	5	Uses only the cells on the edges of the grid
Simple shapes	6	Closed simple polygon shapes
Abstract	-	Passwords that do not follow any obvious patterns

Table 4: Different categories of TMD password patterns

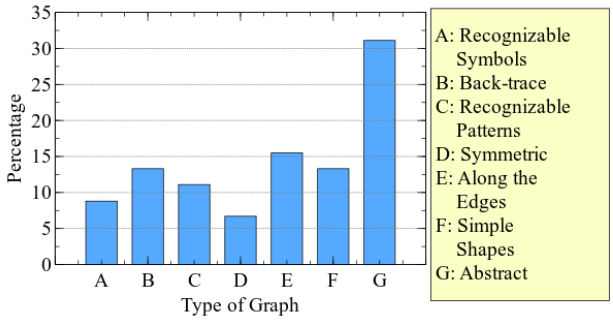


Figure 11: Distribution of TMD password patterns

the interface might have caused this concentration of starting points. Adding additional visual references and studying their effect on starting points should be included in future work.

We note that while these patterns are of concern, they should be considered within the context of the study. We made no attempt to influence user choice of passwords nor did we instruct users on how to select "secure" passwords because we wanted to see users' natural inclinations. Just as text passwords require password rules to enforce a minimum standard, we expect that some rules and instructions could greatly improve the security of user chosen TMD passwords.

Interpretation

TMD has several advantages over DAS on touchscreens. First, TMD passwords were more memorable than DAS passwords after one week. Second, users preferred TMD passwords over text passwords and were more willing to use TMD passwords to protect important accounts. Finally, TMD eliminated the fuzzy boundary problem that users experienced with DAS. In all other measures, TMD performed at least as well as DAS.

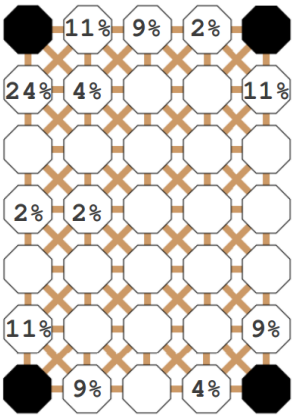


Figure 12: Distribution of the starting points of the passwords

DISCUSSION

Our motivation for this work was the mounting frustration with using text-based passwords on mobile devices. This problem is becoming more prevalent as the popularity of such devices increases. We took an iterative approach, first uncovering what worked and what was problematic with existing alternatives, then used these results to inspire the design of our new scheme.

Limitations of the studies include testing with a primarily young university population (although these do represent a significant portion of actual mobile users) and the lab environment. As a first exploration, we wanted to control for equipment differences and environmental influences so that we could focus on the actual schemes, but further testing is required in more ecologically valid scenarios.

We have identified potential weaknesses in TMD and have strategies for addressing them. Possible password policies include encouraging or requiring users to use multiple layers in their passwords, limiting the number of cells in the password that are on the edges of the grid, or asking users to go through randomly pre-selected cells at least once. Any such rules would need to be user tested to ensure that password memorability remains acceptable. Randomly-assigned passwords are another option that could be explored. Even in its present state, however, TMD has a larger password space than the Android screen unlock, its most similar deployed scheme.

Possible modifications to TMD may encourage users to select more random passwords. For example, the addition of visual cues could allow users to move away from the edges and still have confidence that they can remember their starting point. Learning from other similar schemes [6], we might add background images to serve as visual references. However, since users did not like to memorize unfamiliar images as part of their password in our exploratory study; an alternative solution is to allow users upload their own images. We would need to mitigate the risk that they might use an image containing obvious cues. Other possible visual cues include changing the colour or adding a small glyph or symbol on some cells to act as reference points.

It is also worth highlighting the features of TMD that were successful and that make TMD a usable scheme. One of these qualities is the concept of warp cells and multiple layers. During our test, users had no problem understanding the concept of layers and how to reach the next layer by using warp cells. The visual design of the scheme appears intuitive, making the scheme easy to understand and use. This simplicity encouraged users to create passwords that where nearly double the minimum required length, indicating that we could easily enforce a larger password space than we did in this study. These passwords were also memorable. Over 95% of the participants were able to remember the password after a week even though they have never used the scheme before. On the contrary, only 71% of the participants were able to remember their DAS passwords. To maximize the screen usage, the main interface screen does not have any buttons; the scheme senses the end of a password entry when users lift their finger. Also, it is possible to deploy TMD to an existing service without having to modify back-end databases because it converts the graphical password into an alphanumeric string locally before sending the data. Furthermore, although the scheme is designed specifically for touchscreens, it also functions with a mouse-based interface so could be utilized on user accounts that cross platforms or web-based services. In fact, the current implementation is entirely web based.

CONCLUSION

User authentication is a key issue that must be addressed for successful integration of mobile devices into end-users' daily lives. This paper presents two distinct contributions to the mobile HCI literature. First, it presents an evaluation of three existing graphical password schemes (one from each major category) on tablet computers and smart phones. This exploratory study revealed significant usability problems with all three schemes. DAS suffered from accuracy problems while users disliked PCCP and Object Recognition because they had to memorize unfamiliar images.

Our second contribution is a new password scheme which combines the advantages of the three existing schemes while addressing the usability problems that were uncovered. The interface of TMD is composed of large detached cells. The large cells were intended to reduce errors caused by accuracy problems and the space between the cells were intended to eliminate the fuzzy boundaries problem. TMD also avoids having users memorize images. TMD uses *layers* to increase the password space so the length of the password is bounded by the device memory or system configuration but not the screen size. User testing of TMD shows that it has superior performance to DAS, the closest existing scheme, on a number of measures. On the remaining measures, it performs equally well. While improvements are possible, we are encouraged by these positive results and believe that TMD is worth further exploration.

REFERENCES

1. Aviv, A. J., Gibson, K., Mossop, E., Blaze, M., and Smith, J. M. Smudge attacks on smartphone touch screens. In *USENIX Conference on Offensive Technologies (WOOT)* (2010).
2. Bao, P., Pierce, J., Whittaker, S., and Zhai, S. Smart phone use by non-mobile business users. In *ACM MobileHCI* (2011).
3. Biddle, R., Chiasson, S., and Van Oorschot, P. Graphical passwords: Learning from the first twelve years. *ACM Computing Survey 44*, 4 (Sept. 2012), 19:1–19:41.
4. Chiasson, S., Forget, A., Biddle, R., and van Oorschot, P. C. Influencing users towards better passwords: Persuasive Cued Click-Points. In *British HCI Annual Conference (BCS-HCI)* (2008).
5. Dunphy, P., Heiner, A. P., and Asokan, N. A closer look at recognition-based graphical passwords on mobile devices. In *ACM Symposium on Usable Privacy and Security (SOUPS)* (2010).
6. Dunphy, P., and Yan, J. Do background images improve "Draw a Secret" graphical passwords? In *ACM Computer and Communications Security (CCS)* (2007).
7. Florencio, D., Herley, C., and Coskun, B. Do strong web passwords accomplish anything? In *USENIX Workshop on Hot Topics in Security (HOTSEC)* (2007), 1–6.
8. Henze, N., Rukzio, E., and Boll, S. 100,000,000 taps: analysis and improvement of touch performance in the large. In *ACM MobileHCI* (2011).
9. Hlywa, M., Biddle, R., and Patrick, A. S. Facing the facts about image type in recognition-based graphical passwords. In *Annual Computer Security Applications Conference (ACSAC)* (2011).
10. Jermyn, I., Mayer, A., Monrose, F., Reiter, M. K., and Rubin, A. The design and analysis of graphical passwords. In *USENIX Security Symposium* (1999).
11. Lee, S., and Zhai, S. The performance of touch screen soft buttons. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, ACM (New York, NY, USA, 2009), 309–318.
12. Passfaces Corporation. The science behind Passfaces. `www.passfaces.com/published/The%20Science%20Behind%20Passfaces.pdf`. June 2012.
13. Sae-Bae, N., Ahmed, K., Isbister, K., and Memon, N. Biometric-rich gestures: a novel approach to authentication on multi-touch devices. In *ACM CHI* (2012).
14. Schaub, F., Deyhle, R., and Weber, M. Password entry usability and shoulder surfing susceptibility on different smartphone platforms. In *ACM Conference on Mobile and Ubiquitous Multimedia (MUM)* (2012).
15. Tao, H., and Adams, C. Pass-Go: A proposal to improve the usability of graphical passwords. *International Journal of Network Security 7*, 2 (2008).
16. van Oorschot, P. C., and Thorpe, J. On predictive models and user-drawn graphical passwords. *ACM Transactions on Information System Security 10*, 4 (2008).