SemanticLock: An authentication method for Mobile devices using semantically-linked images

ILESANMI .A OLADE, Xi'an Jiaotong-Liverpool University, China
DR. HAI-NING LIANG, Xi'an Jiaotong-Liverpool University, China
DR. CHARLES FLEMING, Xi'an Jiaotong-Liverpool University, China

We introduce SemanticLock, a single factor graphical authentication solution for mobile devices. SemanticLock uses a set of graphical images as password tokens that construct a semantically memorable story representing the user's password. A familiar and quick action of dragging or dropping the images into their respective positions either in a *continous flow* or in *discrete* movements on the the touchscreen is what is required to use our solution.

The authentication strength of the SemanticLock is based on the large number of possible semantic constructs derived from the positioning of the image tokens and the type of images selected. Semantic Lock has a high resistance to smudge attacks and it equally exhibits a higher level of memorability due to its graphical paradigm.

In a three weeks user study with 21 participants comparing SemanticLock against other authentication systems, we discovered that SemanticLock outperformed the PIN and matched the PATTERN both on speed, memorability, user acceptance and usability. Furthermore, qualitative test also show that SemanticLock was rated more superior in like-ability. SemanticLock was also evaluated while participants walked unencumbered and walked encumbered carrying "everyday" items to analyze the effects of such activities on its usage.

Additional Key Words and Phrases: Authentication-Mechanisms; Pattern; PIN; Mobile Devices; Security; User Studies

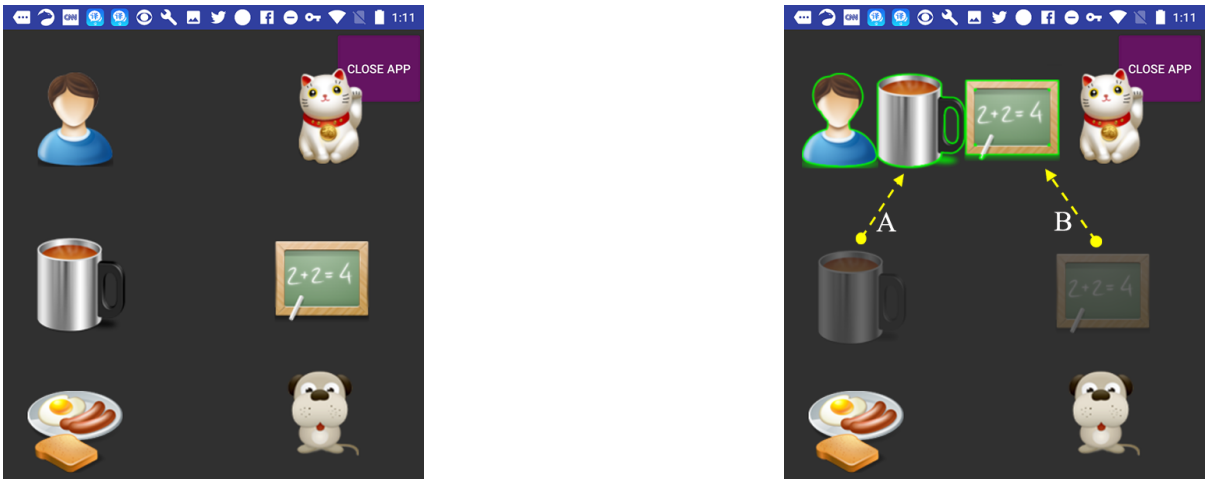
ACM Reference Format:
Ilesanmi .A Olade, Dr. Hai-ning Liang, and Dr. Charles Fleming. 2018. SemanticLock: An authentication method for Mobile devices using semantically-linked images. 1, 1 (July 2018), 21 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Mobile phones, being the de facto personal communication device, are ubiquitous within our society [28] . We depend on these devices to store substantial amounts of confidential information and perform activities such as emailing, social networking, personal banking and entertainment. All mobile devices manufactured in the last decade come with a default set of authentication or login mechanisms. Existing research shows that a large percentage (64% based on assertions by [17]) of users chose not to secure or use an authentication system on their mobile devices [14]. While it has been suggested that users may not assign significance to the information existing on their mobile devices[2], other arguments, such as that made by Micallef et al[17], suggest that users dislike the inconvenience of repeatedly unlocking their mobile devices. Even those who choose to engage in the use of their mobile device locking mechanisms are discouraged by the time and effort it takes to log in and the frustrating login failure errors that they observe during those attempts [24]. Users prefer to sacrifice security for access speed and efficiency [6]. The popularity of touch-enabled mobile devices has reduced the use

Authors' addresses: Ilesanmi .A Olade, Xi'an Jiaotong-Liverpool University, 111 Ren ai Street, Suzhou, China, 23185, China, ilesanmi.olade@xjtlu.edu.cn; Dr. Hai-ning Liang, Xi'an Jiaotong-Liverpool University, 111 Ren ai Street, Suzhou, China, 23185, China, HaiNing.Liang@xjtlu.edu.cn; Dr. Charles Fleming, Xi'an Jiaotong-Liverpool University, 111 Ren ai Street, Suzhou, China, 23185, China, Charles.Fleming@xjtlu.edu.cn.

of alphanumeric textual passwords which has a larger password space and better security [7] but suffer from usability issues [16]. Because of these usability issues and the popularity of touchscreen devices, graphic-based authentication approaches such as PIN [13, 15, 27] and PATTERN [9, 13, 26, 30] have recently become prominent [27]. The PIN authentication system Fig.3(c), which is a numeric display of numbers inputted by discrete touches on the screen and PATTERN authentication system Fig.3(b),which is a grid like display of nodes whose password pattern is selected by a continuous movement across the screen to connect the secret password nodes are both plagued with numerous usage and security issues [1, 3, 13, 18, 30].



(a) (b)

Fig. 1. **SemanticLock**: (a) Default view for login and setup. (b) Login: the user drags two images to meet a third image. In this case, Cup is dragged to right side of Person ***(movement "A")***, then Blackboard is dragged to right side of Cup ***(movement "B")***.Login can be done with *two quick* movements.

In this paper we present SemanticLock, a single factor graphical authentication method for touchscreen mobile devices. Our solution works by providing the user with a way to unlock their mobile devices by joining images via discrete or continuous gestures to create a semantically memorable story that represent a password (see Fig. 1). SemanticLock can establish a strong memorable password with just two discrete movements allowing the user to construct a semantically meaningful password quickly (see Fig. 1(b)) from the provided images. We show that while SemanticLock can be as secure as PIN and PATTERN, its performance is significantly better than PIN and similar to PATTERN under normal circumstances but exceedingly better in ideal scenarios. Our three weeks user study positively shows that our users found SemanticLock easily discoverable and easier to remember through out the study period. Furthermore as previously mentioned, users drag and join at least three of the provided images to construct their password (see Fig. 1(b)), the semantic story could be *"I drink coffee and study"*. The location of each image in the group constitutes part of the password algorithm. Other password patterns are displayed in Fig. 2(a) and Fig. 2(b). These pictures are just examples and they can be customized by users. Fig. 2(c) shows how easy it is to create a new password.

1.1 Challenges and proposed design approach

In designing the SemanticLock system we set out to develop a system that was easy to use and quick to login, therefore our primary focus was speed, ease of use, and memorability. In addition we expect our solution to

perform consistently across all usage environments and situations our users may find themselves. Our study involved scenarios such as sitting, walking unencumbered and encumbered. We ensured that SemanticLock requires only two distinct swipes or finger movements to construct a login password, and we implemented a close proximity "sticky" feature that visually highlights the two images that are in close proximity to each other while the user is actively dragging one of the images and if the user releases this image it automatically "glides" towards the closest image and "sticks" to it. This feature greatly reduces errors caused by unsteady finger movements and increases overall login speeds. The SemanticLock also inherits the discrete and continuous finger movement properties of the PIN and PATTERN authentication system respectively. However, in contrast to PATTERN authentication system, SemanticLock can rely on two short swipes rather than one continuous and long swipe often from one side of the screen to the other. Similarly, SemanticLock requires lifting the finger only once in between the swipes, rather than 3 or more distinct finger movements required to use the PIN authentication system. SemanticLock is inherently resistant to smudge attacks because the location of its passwords tokens on the screen is irrelevant to the creation of the password, whereas the PIN and PATTERN authentication systems are susceptible to smudge attacks[13, 30].

Furthermore to assess our system and get a comparative evaluation, we selected the two most widely used graphical authentication systems as a control and conducted a user study over a three weeks period to compare our SemanticLock authentication system against PATTERN, PIN, and PIN-Shuffled authentication systems (see Fig.3). As such we have formulated the following hypotheses:

Hypothesis H1: We expect that the results obtained from the SemanticLock system will be comparatively similar or relatively close to those of the PATTERN authentication system, despite the fact that most of our users have prior inherent knowledge and familiarity with the PATTERN authentication system.

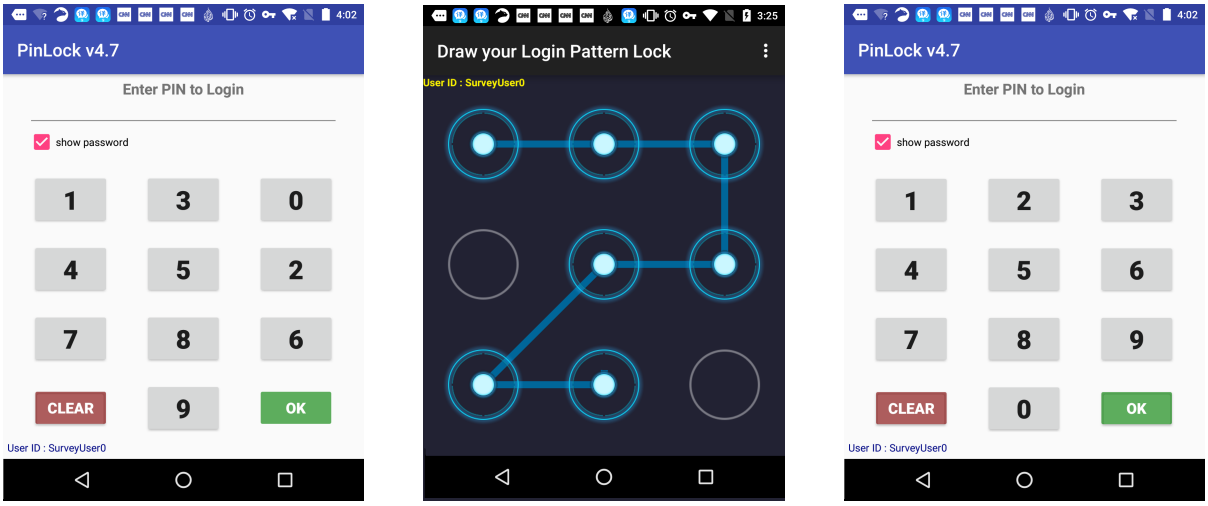
Hypothesis H2: We also expect the PIN system to firmly supersede the SemanticLock system and be closer in performance to the PATTERN authentication system

Hypothesis H3: We predict that the SemanticLock will supersede the PIN-SHUFFLE in regards to unlock speed. Both authentication systems are inherently resistant to smudge attacks.

(a) (b) (c)

Fig. 2. **SemanticLock:** (a) (person-breakfast-coffee: *"I eat breakfast with coffee"*). (b) (cat-breakfast-dog: *"cat shares meal with dog"*)

Before we describe our research methodology, the next section presents work related to authentication methods.



(a) PIN Shuffled (b) Pattern (c) PIN

Fig. 3. Prominent mobile device authentication systems

2 RELATED WORKS ON AUTHENTICATION METHODS

SemanticLock is a single factor graphical password system that is suited for touch screen mobile devices. It combines properties from both the *Recall* and *Recognition* aspects of the graphical password system. Our research focuses on the popular draw-metric category that includes the PATTERN authentication system. Although PIN authentication system do not fit into any definable taxonomy, it is applicable to all mobile device form factors [2, 7, 11] and remains very popular with users. As such, we want to also examine various authentication methods and compare them with SemanticLock. Any new authentication methods should be evaluated for their levels of usability and security [8].

2.1 Graphical Passwords

Graphical authentication methods fit into 3 main categories; namely: *Recall*, *Recognition* and *Cued-recall*. The *Recall* graphical authentication system gets its origin from works done on Draw-a-Secret[30] and other similar systems[4, 9**? ?**]. The advantage of *Recall* is that it benefits from the inherent motor memory of the users and our superior ability to recall shapes and patterns[11, 26]. PATTERN authentication and SemanticLock belong to this class, but SemanticLock is resistant to Smudge attacks while PATTERN authentication is not.

Graphical passwords, like the PATTERN authentication system, have also been used as a baseline measurement for new methods. von Zezschwitz et al. [28] reported the study of three custom graphical methods to gauge their performance against the PATTERN authentication system. Their aim was to study their prototypes' susceptibility to smudge attacks, level of memorability, usability and user acceptance. Their results confirmed that PATTERN authentication system was superior to their proposed prototypes in all aspects except for resistance to smudge attacks. The PIN authentication system was not included in their study. More recently Aly et al. introduced SpinLock [2], a technique that is based on a physical combination lock, and requires users to rotate a dial both

counter-clockwise and clockwise alternatively to select a password token. The rotating mechanism is meant to delete previous smudge traces on the screen, thereby making it very difficult to replicate a user's password via smudge attacks. Also, its design is meant to make it usable but without sacrificing security. Their study with 21 participants using SpinLock in 63 trials with various degrees of password complexity show that it has led to significantly lower time performance than Pattern Lock and only achieved the similar performance with PIN. Their participants thought that SpinLock was more usable and enjoyable to use.

Harbach et al. [13] performed a month long study with 134 participants using the PhoneLab1 panel instrumentation consisting of LG Nexus 5 phones to collect detailed data on login speed, error counts, types of errors and the effect of use scenarios. They compared three authentication systems (PIN, Pattern, slide-to-unlock). The slide-to-unlock required no secret token, but users needed to do a"sliding" action on a button to move it from left to right across the screen. Their quantitative results shows that users spent longer "preparation time" when using the PIN authentication system, suggesting that this was due to the recall process associated with the PIN passwords. Predictably the slide-to-unlock had the shortest preparation time. The PATTERN authentication system had a shorter unlock time when compared with PIN authentication system, while the slide-to-unlock had the shortest unlock time. The PATTERN authentication system led to a higher occurrence of unlock errors when compared to PIN authentication system; the authors argued that it was probably attributed to problems with the users interacting with the touchscreen to perform the continuous swiping gesture. Their subjective data suggested that research on lock screens needs to focus more towards real life usability factors such has authentication times, positive user experiences and error rates rather than more secure factors.

Although various drawmetric based studies has attempted to improve memorability while trying to have reasonable login speed, only a few, such as Spinlock [2], have performed real world experiments over a reasonable period of time. We want to assess how well SemanticLock will perform in somewhat natural environments and scenarios that are representative of how mobile phones are actually used.

2.2 Text Based Passwords

There are numerous variants of PIN based authentication. Scrambled PIN proposed by [9] is similar to our PIN-Shuffle by shuffling the keys. Using a Cyanogen Mod android version, the authors investigated how these passwords can overcome attacks that affect the standard PINs. Their results show that the shoulder-surfing attack achieves an ASR (Attack Success Rate) of 45.83% on a standard PIN keypad and an ASR of 21.39% when using scrambled PINs. However, the improvement in security led to increase in time needed to enter the passwords the standard PINs were 2.5 times faster than the scrambled version. In addition, the latter led users to a slightly higher number of failed attempts. That is, although the scrambled PINs resulted in better security, it had a lower level of usability.

2.3 Effects of Mobility and Activity on Authentication Experience

Not much research is available on the effects of mobility and encumbrance while using mobile devices, especially to unlock them. Users of Mobile devices rarely focus all their attention on their mobile devices, but their attention is distributed [19]. Ng et al [22] in their initial study, discovered that mobile phone users simultaneously carry or hold other items while interacting with their devices in public tend to carried shopping bags and boxes often. Additional studies by Ng et al. [20], had users clicking on"crosses" or target points that randomly appear on the screen to study the effects of encumbrance and walking on user's targeting accuracy while using walking on a treadmill or on normal ground determined that targeting error rates increased by 112% compared to when participants were standing still. Walking on the treadmill led to greater number of errors than walking on the ground. In [22], it has been reported that users performed better using rotational actions while walking or encumbered, while tapping and dragging (or sometimes swiping) caused increases in the number of errors.

Wilson et al.[29] and others [10, 12, 25] found users were markedly less accurate at targeting on mobile devices and selection time would increase significantly when encumbered or walking while interacting with a mobile phone regardless of input hand posture (similar findings were reported in [23]).

In the next section we describe our research methodology to investigate performance, error recovery, memorability, usability and users' perception of SemanticLock as an alternative authentication mechanism.

3 METHODOLOGY

We conducted a three weeks user study involving 21 participants within an indoor environment to collect both qualitative and quantitative data to gain insight into our participant's perception of the likeability, usability, memorability and login speed of the 5 authentication approaches (with two being shuffled, variants): (1) SemanticLock, (2) SemanticLock-shuffled, (3) Pattern Lock, (4) PIN, and (5) PIN-shuffled. Pattern and PIN are treated as our baseline for this test due to their popularity.

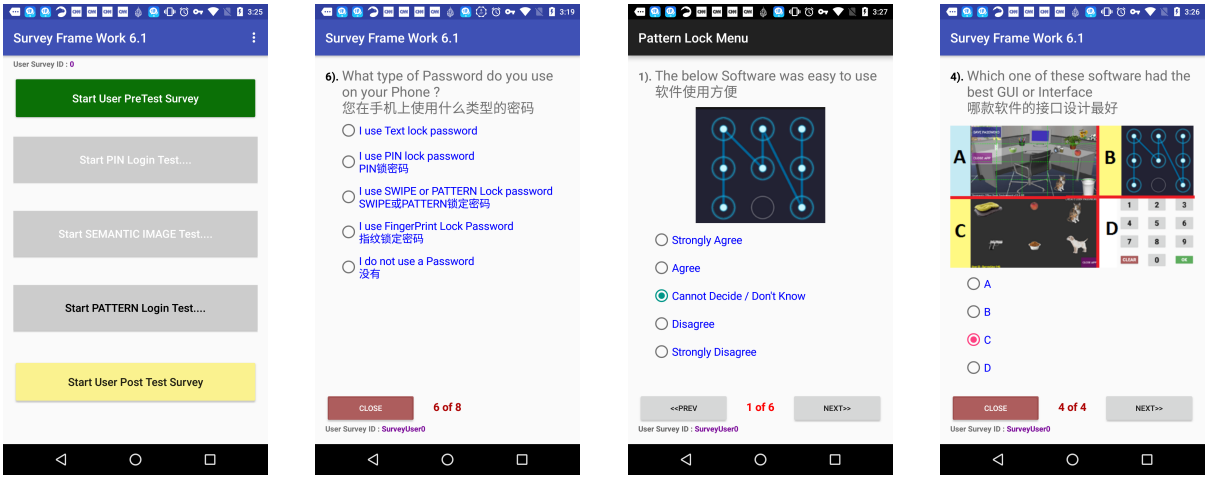
3.1 Apparatus

The prototypes, shown in (Fig. 2 and Fig. 3), were developed using Android studio ensuring compatibility with Android 6.0 and above. All prototypes were adapted to work on the phone and tablet form-factors. The Training mode option allowed users to receive adequate training and practice before the actual testing. During the testing, a participant's activities such as touches, password tokens, strokes, pauses, timings, aborts and errors were logged for later analysis. We developed an additional application to help us convey the testing and survey to our participants in a uniform and predictable way. It allowed participants to view an initial training video, assigned a unique participant ID that allowed us to correlate data across Login techniques on participant basis and also presented the pre-survey and post-survey questionnaires in the proper sequences while implementing the Latin square approach to counterbalance the order of the techniques (see Fig. 4).

3.2 Experimental Design

Our goal was to compare three main techniques and two shuffled variations and their interactions with other independent variables. To do this, we followed a within-participants design. The *independent variables* were (1) Technique, (2) Device Form-Factor, (3) Physical Posture, and (4) Hand Posture. Our *dependent variables* are (1) Login Speed, (2) Pre-Login Delay Time, (3) Error Rate, and (4) User usability and acceptance.

3.2.1 Technique. Our experiment compared three techniques and two shuffled variations. The task required of each participant was to enter the password tokens as fast as possible during each session, whereby we implicitly collected and tracked data and meta-data for further empirical analysis. We assigned password tokens for each technique so that each participant would use a sufficiently strong password properly distributed within the space of possible passwords. We attempted to ensure that the password tokens given for each technique had relatively the same password strength. For the PIN Technique we issued a series of 4 and 6 digit password tokens, having a possible theoretical password space of 10,000 to 1,000,000 (see Table 1). 4-6 digit passwords represent the range of what most people would use in mobile devices and other platforms (e.g., ATM PINs). For the Pattern Technique, we assigned a series of irregular and widely distributed patterns with a minimum of 5 connected nodes, giving us a theoretical space range of 7000 to 140,000 possibilities (Table 1). For the SemanticLock Technique we issued a series of semantically meaningful passwords that would enhance memorability, with a minimum of three and a maximum of six password images required to form a password token, thereby giving a theoretical password space of 11,520 to 2,764,800 possibilities (Table 1).



(a) User Study Admin App (b) User Survey Question (c) Post Test In-App Survey (d) Post Test Survey

Fig. 4. **Survey App Framework** : The survey framework app allowed us to provide a consistent process to all participants. (a) The Main menu used Latin Square to present the Test options. (b) Pre-Test Survey collected user demographics and preferences. (c) Post-Test survey specific to the system just tested. (d) Post-Test general Survey, to collect user's overall opinions

Table 1. Authentication System Password Space.

	Number of Possibilities		
Nodes,Numbers,Images	**PATTERN**	**PIN**	**SemanticLock**
3		1,000	11,520
4	1624	10,000	814,464
5	7152	100,000	1,915,776
6	26016	1,000,000	2,764,800
7	72912	10,000,000	
8	140704	100,000,000	
9	140704	1,000,000,000	

3.2.2 Device Form-Factor. Mobile devices come in various dimensions. We used two different form-factors: (1) A 5.2" LG Nexus 5X phone, and (2) 10.2" Google Pixel C tablet. The tablet was only used during the Seated session (Fig, 5 (a)) of the experiment, while the LG phone was used for all sessions(Fig.5 (b),(c)).

3.2.3 Physical Posture. Studies show that the physical posture of users has an effect on the way they use the devices [19, 20, 22]. Therefore in this study we included 3 physical postures (seated, walking, walking-encumbered). The *Seated* condition required participants to sit on a comfortable chair and operate the mobile device on a table and could use one or two hands (see Fig. 5(a)). Walking unencumbered implied that the person operating the mobile device was also walking but without carrying any other objects with their hands or arms (see Fig 5(b)). Walking encumbered took place when participants would operate a mobile device while carrying other items such as books or bags with their hands or arms (see Fig 5(c)). Recent studies have shown that walking

(a) (b) (c)

Fig. 5. **Participants in the Study:** (a)Participant performing a Seated Test using the Tablet. (b) User in walking unencumbered. (c) encumbered posture while using the single hand main thumb input posture.

encumbered or unencumbered and operating a mobile device had shown significant effects on the usage pattern of mobile devices [21, 23, 25]

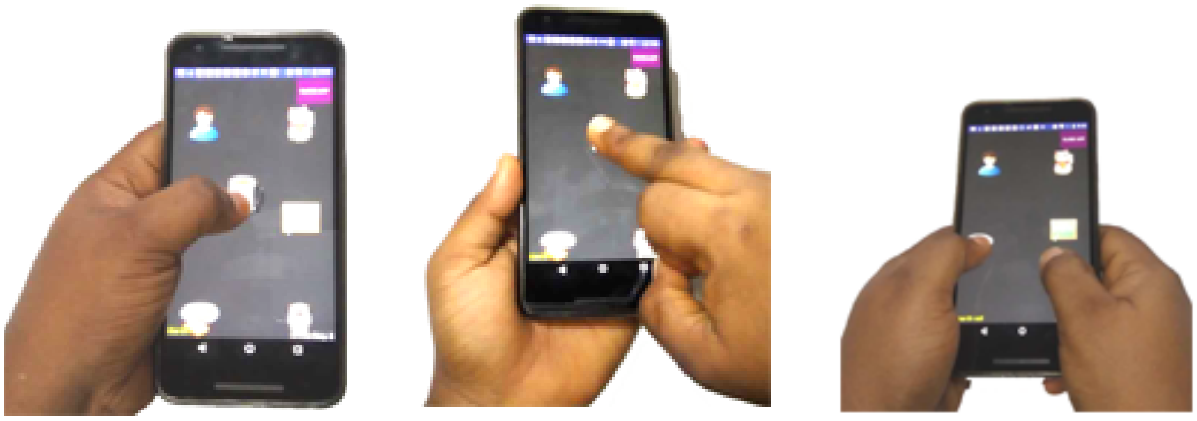
3.2.4 Hand Posture. Hand Posture defines how a mobile phone is held when in use by the user. There are 3 prominent input postures: *one-handed preferred thumb, two-handed index finger* and *two-handed both thumbs* (see Fig. 6). With the advent of larger mobile phone screens, many users have had to change from the one hand input posture to the two-handed input posture [21, 23].

3.3 Participants

We recruited 21 participants (15 females) from a local university. Data from our pre-testing survey told us that 51% of the participants were between the ages of 17 to 22 and all our participants were right-handed. All were active iPhone (31%) and Android (66%) users. 55% of them used a phone with fingerprint sensor, while 17% used the PIN, and 14% Pattern. The remaining 14% did not use authentication. 50% of our participants claim the input hand posture they preferred to use depended on the situation and the app in question; 27% claimed they preferred to use two hands to operate their mobile devices. All participated voluntarily without any financial remuneration.

3.4 Task and Procedures

Our first step was to inform the participants about the confidentiality of their supplied information and to explain the purpose of the project and the tasks they would need to do. We provided a three minute training video to each participant (see Fig. 4a), after which they were allowed to practice each technique a couple of times. They



(a) One-handed thumb (b) two-handed index finger (c) two-handed both thumbs

Fig. 6. **Input Hand Postures:** The most common hand postures when using mobile devices. These postured were tested during the Study.

practiced the creation of a password and the use the password to log in into the mobile device. We emphasized the need for a speedy login during the actual testing phase, accuracy, and conform.

3.5 Week 1 (First Phase)

Each participant was required to answer a Pre-test questionnaire before commencing the test (see Fig. 4b). We allowed each participant to choose password tokens for each technique from our supplied list. If the participant entered a wrong password, the application alerted them to enter the correct password again. The average time for participants to complete all techniques (including questionnaires) was 4 minutes. The experiment finished with a Likert questionnaire (see Fig. 4c) that collected qualitative data about how the participants' perceived usability, error-handling, security and likeability of each technique. This week's session was applied as a seated session and the participants used the Techniques on the LG mobile phone and the Google tablet. The main independent variable was *Techniques* (PIN, Pattern and SemanticLock) and *Mobile form factor* (Phone and Tablet). The PIN and SemanticLock have a *"shuffled"* variation within the main technique as explained previously in this paper. Each participant has to enter a total of 15 passwords per session, 3 for each Technique and participants are allowed a 60 second rest in between techniques to avoid fatigue.

3.6 Week 2 (Second Phase)

In the second phase, we explore the memorability of the techniques where we asked the same participants to recall the passwords they had used for each technique the week before. During this session we tracked error-rates, type-of-error, action-delay times, login speed required for our future analysis.

3.7 Week 3 (Third Phase)

We recalled the participants for a third session that required them to perform log in activities while walking around a predefined path within an indoor environment. We followed certain practices and insights from [21, 22] in which they examined the effect of mobility and encumbrance on participants using both one and two-handed

Fig. 7. **Pacing the user:** A Pacesetter (right) keeping the participant (left) at a steady walking pace with the help of a metronome software during a login test.

interactions on touchscreen mobile devices. The walking speed was paced by a researcher who used a metronome to ensure a proper walking speed was maintained (see Fig. 7). After the walking test (see Fig 5(b)), each participant undertook the Encumbrance test, which required each participant to walk along a path at a paced speed carrying two nylon bags containing a 100cl plastic bottle of Coca-cola while unlocking the device using each technique (see Fig 5(c)). The decision to use nylon bags was informed by the research done by Ng et al. [22]. In this phase, we sought to investigate the effect of mobility and encumbrance on the login speed and input errors while assessing the techniques with the 3 commonly used input postures as discovered in a research by [23]: *onehanded preferred thumb, two-handed index finger*, and *two-handed both thumbs*.

4 MEASUREMENT

We collected data for a number of dependent variables. However, we used the hypotheses we presented previously to give a focused analysis of our data.

4.1 Pre-Login Delay time

The elapsed time between when the participant indicated that they were ready to start to unlock the device and the actual entry of the initial password token is called *pre-login delay time*. This period of time provided the data that would be used to derive information about how well the participants were familiar or comfortable with each technique. It would also give us an insight into participants' recall of the passwords, the longer the delays, the more difficult it was for them to remember their password.

4.2 Login Speed

The time period used to complete each trial of the Login process for a technique was recorded. This measurement only recorded successful trials; failed trials were recorded as singular failure events. Login speed was tracked from the moment a participant starts password token entry until the entry was completed successfully.

4.3 Error Rate

The error rate was measured as a percentage of failed login attempts to the total number of attempts required to complete the technique's session. The number of failed login attempts during a trial did not affect the number of trials that constituted a complete session. That is, some techniques required 3 successful trials to constitute a session while other techniques, such has the PIN and SemanticLock; required 6 successful trials to complete a session.

4.4 Subjective Data

We collected Pre-Test, In-Test and Post-Test surveys via an electronic questionnaire (see Fig.4 (b,c,d)). The questions focused on ease of use, perception of speed, likelihood of adoption, error recovery, and interface usability. We implemented the questionnaire in electronic form and used a 5-point Likert questions for some aspects of the questionnaire.

5 RESULTS

5.1 Login Speed

The mean values of the Login speed of each technique and other independent factors are shown in Table 2. The results show that the Pattern performed better than the other techniques and across device form factors and postures. SemanticLock was superior in performance to PIN across all independent variables. There was a statistically significant difference between the techniques Login Speed as determined by one-way ANOVA ($F(4,535) = 170.44$, $p = .000$). A Tukey post hoc test revealed that Pattern (807.06 ± 167.23 ms, $p = .000$) was significantly faster than SemanticLock and PIN (both $p < .000$). Furthermore, SemanticLock (1200.10 ± 315.12 ms, $p < .000$) was significantly faster than PIN, SemanticLock-Shuffle and PIN-Shuffle. There was no significant difference between the SemanticLock-Shuffle and PIN ($p = .593$).

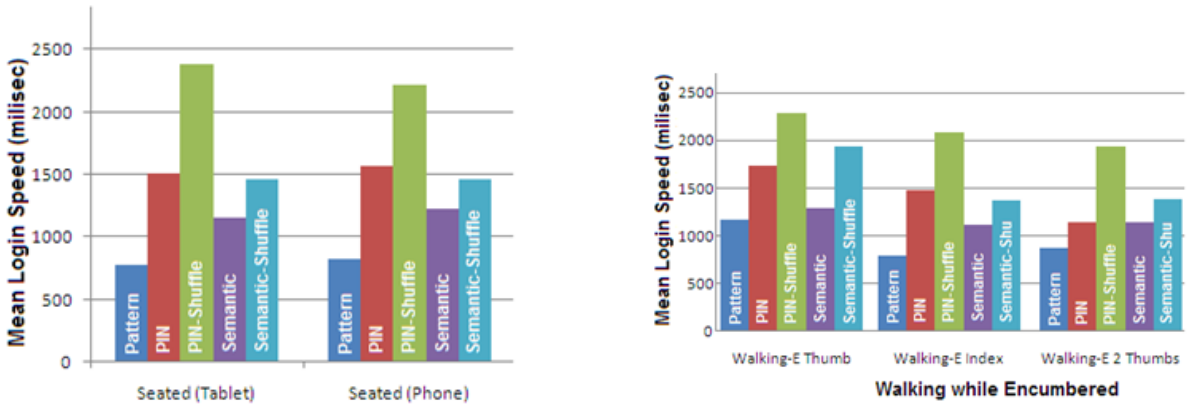
Table 2. Average Login speed across posture and technique

	Pattern	**PIN**	**PIN-Shuffled**	**SemanticLock**	**SemanticLock-Shuffled**
Seated (Tablet)	785	1516	2380	1161	1460
Seated (Phone)	825	1570	2224	1234	1460
Walking Thumb	1135	1885	2480	1543	1625
Walking Index	916	1395	2030	1163	1238
Walking 2 Thumbs	945	1208	1982	1378	1591
Walking-E Thumb	1175	1736	2290	1288	1942
Walking-E Index	800	1474	2089	1110	1374
Walking-E 2 Thumbs	873	1147	1935	1144	1386

Note: *Walking-E* = Walking Encumbered

5.2 Difference across Device Form Factors

As stated earlier, we used two different types of device form-factors during the "seated" session (a Nexus 5 phone and a Google Pixel C tablet (see Fig. 8 (a)). Results of a two-way ANOVA show that there was no significant effect of device form-factor ($F(1,530) = .003$, $p = .995$) on Login speed across techniques. Furthermore there was no significant interaction effect between device form-factor and Login technique ($F(4,530) = 1.208$, $p = .306$), (see Fig. 9 (a))



(a) Login speed while seated with different device form

(b) Login Speed while Encumbered

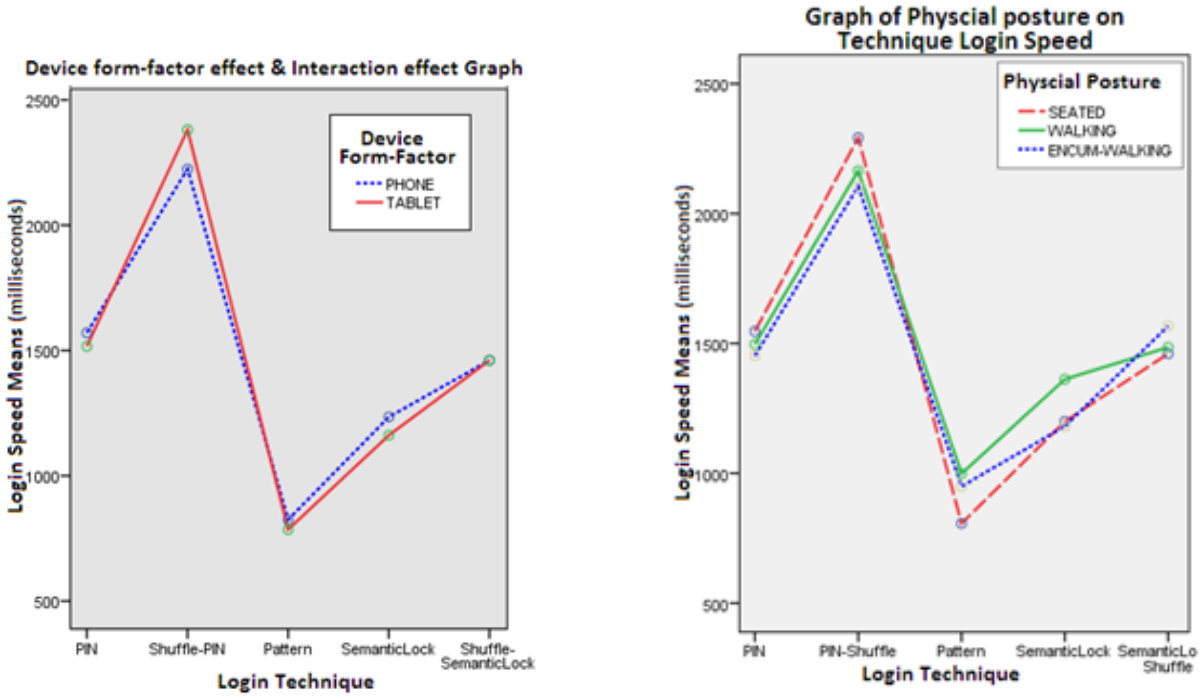
Fig. 8. **Charts of Login Speed** .

5.3 Differences across Physical Postures

Our participants assumed three different physical postures (seated, walking and walking-encumbered). Results of a two-way ANOVA show that there was no significant effect of posture ($F(2,1485) = 1.189$, $p = .305$) on Login speed across Login techniques (see Fig. 9 (a and b)). However, there was a significant interaction effect between physical posture and Login technique ($F(8,1485) = 3.302$, $p = .001$), with participants having a faster speed using the Pattern while seated. Further analysis of the data with the seated posture data excluded, and using a two-way ANOVA to examine the effect of walking posture (unencumbered or encumbered) and Login technique on Login speed show that there was no significant effect of walking posture ($F(1,950) = 1.757$, $p = .185$) on Login speed across Login techniques (see Fig. 8 (b)). Furthermore there was no significant interaction effect between walking posture and Login technique ($F(4,950) = 1.660$, $p = .157$).

5.4 Differences across Input Hand Postures

Our participants while walking either unencumbered or encumbered assumed three different input hand posture (*OneHandThumb, TwoHands2Thumbs, OneHandOtherIndex*) during the testing of the Login Technique (see Fig. 6). Results of a two-way ANOVA conducted to examine the effect of Input Hand posture and Login technique on Login speed shows that there was a significant effect of Input Hand posture ($F(2,945) = 59.318$, $p = .000$) on Login speed across Login techniques (see Fig. 10 (a)). Furthermore there was a significant interaction effect between Input Hand posture and Login technique ($F(8,945) = 2.973$, $p = .003$). A Tukey post hoc test revealed that the *TwoHand2Thumb* posture (1357 ms, $p = .000$) was statistically significantly faster than *OneHandThumb*, but there



(a) Login Speed compared on Device Form-factor for each Technique.

(b) Login Speed compared on Physical Posture for each Technique.

Fig. 9. **Login Speed vs Technique:** Login Speed compared on (a) Device Form-factor or (b) Physical Posture independent variables.

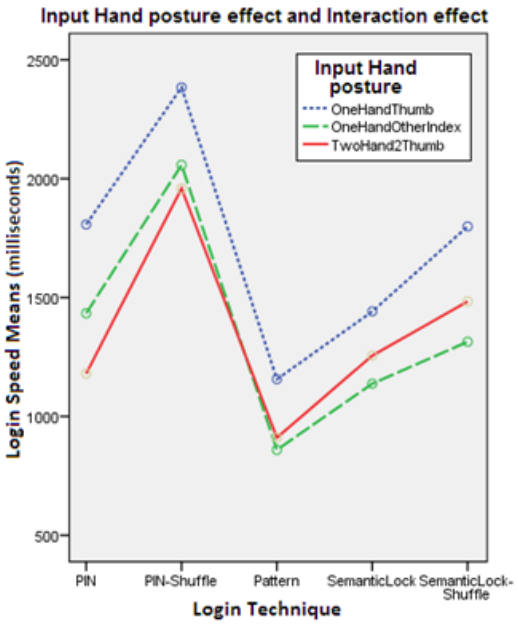
was no statistically significant difference between the *TwoHand2Thumb* and *OneHandOtherIndex* posture (1360 ms, p = .965).

5.5 Pre-Login Delay Time

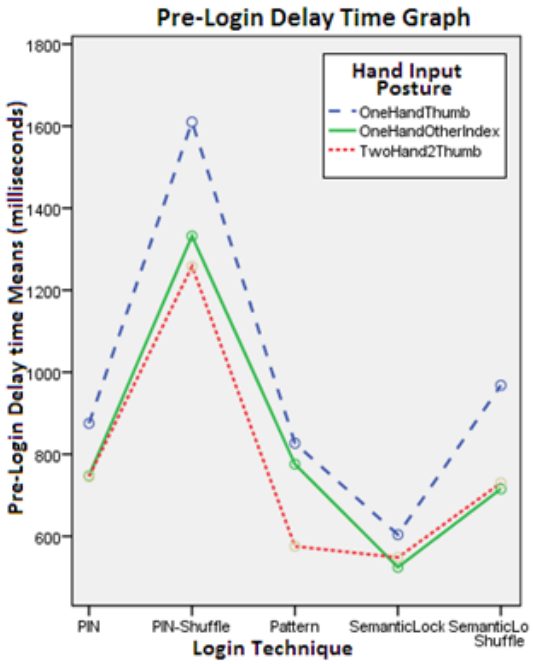
Our participants experience a time delay between when the trial started and when an initial action or interaction was made. This Pre-login delay time would give an indication of familiarity, memorability or ease of use of the techniques. SemanticLock had the lowest pre-login delay time across all Hand Input posture (see Fig. 10 (b)). ANOVA test showed a significant main effect for Login technique ($F(4,930) = 53.864$, $p < 0.05$), where pre-login delay time took significantly longer for PIN-Shuffle (1766 ms) technique while walking unencumbered and using the *OneHandThumb* input posture. The ANOVA test results also showed a significant main effect for Hand Input posture, ($F(2,930) = 9.877$, $p < 0.05$), where the *Twohand2Thumb* had a significantly lower pre-login time than the *OneHandThumb* but there was no significant difference with the *OneHandOtherIndex* ($p= 0.624$).

6 ERROR RATE

A two-way ANOVA was conducted to examine the Error rate for each technique. There was no significant effect of interaction by these independent variables on the Error rate. Furthermore analysis showed that error rate was lowest for all Hand Input postures when using SemanticLock and there was no significant difference in the error



(a) Login Speed compared on Input Hand Posture for each Technique



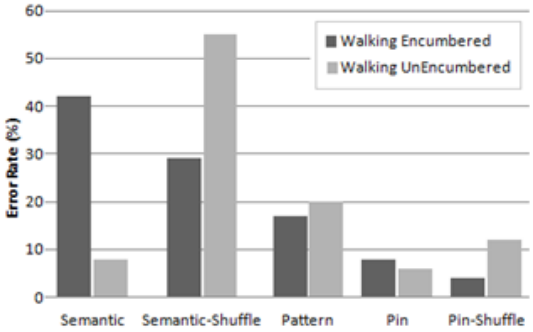
(b) Pre-Login Delay Time compared on Input Hand Posture for each Technique

Fig. 10

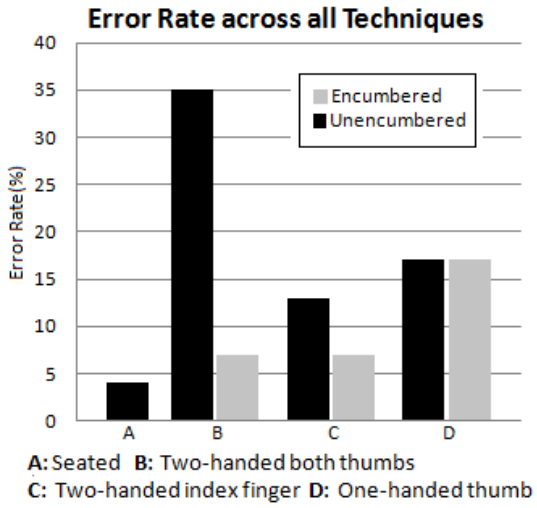
rate of the Pattern technique (p = .925). Additionally, results shows that the error rate while walking encumbered was the lowest when using PIN-Shuffle, while PIN had the lowest error rates when walking unencumbered (see Fig. 11 (a)). It should be noted that data from the participants "seated" sessions was excluded from this walking analysis due to certain inconsistencies in the fidelity of the data. Error rates across all techniques indicates that participants in the seated position had the lowest error rates while the participants using Two-handed both thumbs while walking unencumbered had the highest error rates (see Fig. 11 (b)). Error rates classified by techniques show that SemanticLock-Shuffle (42%) had the highest error rates, followed by Pattern (27%), SemanticLock (19%), PIN-Shuffle (8%) and PIN(4%) (see Fig. 12 (a)).

7 QUALITATIVE RESULTS

The results are based on a 5-point Likert scale questionnaire and subsequent user rankings of the three techniques. Each participant prior to the experiment answered an electronic pre-test survey which we used to obtain demographics, personal information and mobile device usage experience. The Likert scaled questions were answered after the trial of each technique to collect their subject preferences. At the end a user ranking of all techniques was collected (see Fig. 12 (b)). The data we collected was analyzed using Friedman Test and we performed post hoc analysis with Wilcoxon signed-rank test with Bonferroni correction (p= 0.05/3 = 0.017) of those that are statistically significant. In the questionnaire we probed aspects of the users experience with the three Login techniques.



(a) Error Rate compared on walking across all techniques



(b) Error Rate compared across all techniques grouped by posture

Fig. 11. **Error Rates**

7.1 Speed

Our participants experience with technique's speed shows there was a statistically significant difference in speed depending on Technique ($\chi^{2(2)}$ = 18.321, p = 0.000) (see Fig. 13(a)). Post hoc analysis indicated that there were no significant differences between PIN and Pattern trials (Z = -2.101, p = 0.036) or between PIN and SemanticLock trials (Z = -1.560, p = 0.119). However, there was significant difference in speed between Pattern and SemanticLock trials (Z = -3.573, p = 0.000).

7.2 Good Feedback

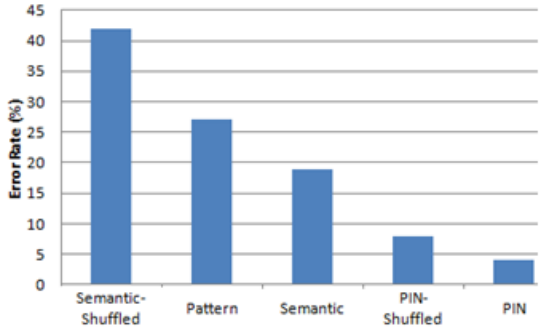
Participants experience with the feedback for each technique also showed that there was a significant difference ($\chi^{2(2)}$ = 17.179, p = 0.000) (see Fig. 13(c)). There were significant differences between Pattern and SemanticLock as well as PIN and Pattern; Pattern were ranked favorably in both cases.

7.3 Likeability

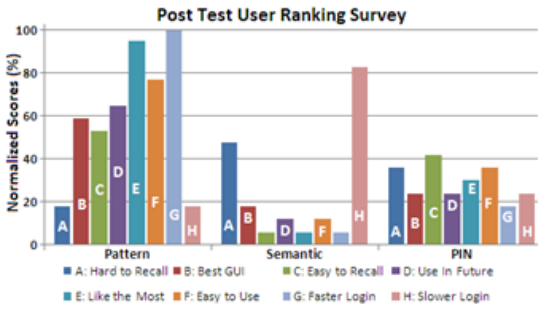
Post hoc analysis indicated that there was no significant difference in how well participants liked the techniques (see Fig. 14 (b)).

7.4 Usability

There was a significant difference in perceived ease of use on Technique ($\chi^{2(2)}$ = 14.22, p = 0.001). Post hoc analysis indicated that there were no significant differences between the PIN and Pattern (Z = -1.672, p = 0.94) or between the PIN and Semantic (Z = -1.628, p = 0.103) (see Fig. 13(b)). However, there was a significant increase in perceived ease of use between Pattern and SemanticLock (Z = -3.140, p = 0.002).



(a) Error Rates for each Technique

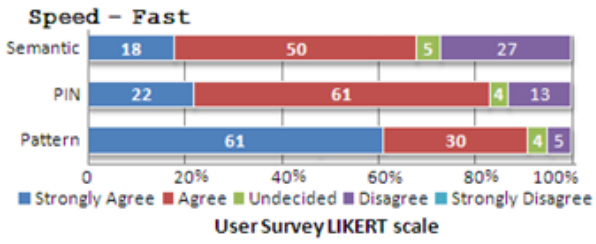


(b) The normalized User ranking.SemanticLock has a strong performance

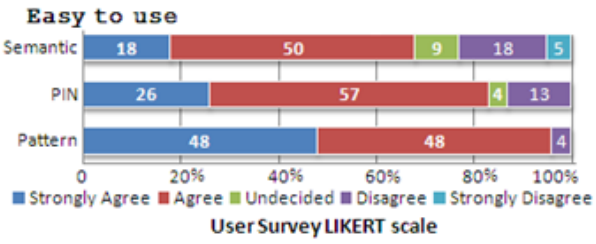
Fig. 12

7.5 Error Recovery

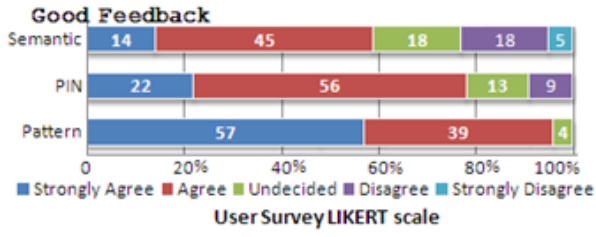
There was a significant difference in error recovery based on Technique ($\chi^{2(2)}$ = 12.667, p = 0.002). Significant differences were found between Pattern and Semantic as well as PIN and SemanticLock. In both cases, Pattern and PIN were ranked favorably in regards to ease of error recovery. There was no significant difference in how participants liked interacting with the techniques (see Fig. 13(d)).



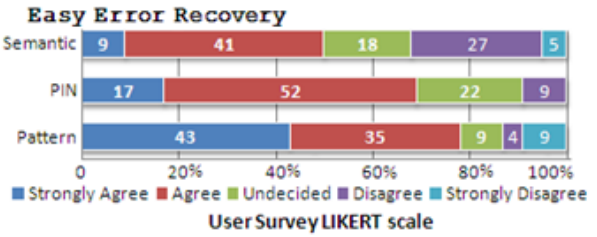
(a) Fast Speed



(b) Easy to use



(c) Good Feedback



(d) Error Recovery

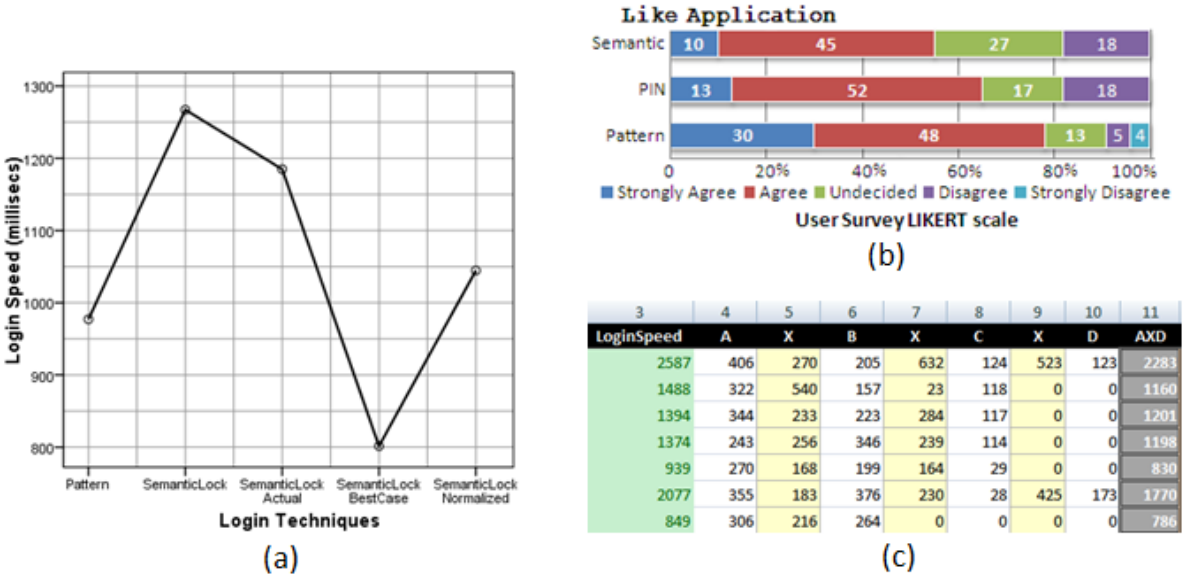
Fig. 13. Quantitative Results

8 DISCUSSION

The results of our research confirms and debunks some of our hypothesis, while exposing other factors that were important to users.

8.1 Login Speed

Our participants performed better, but not significantly, in Login Speed using Pattern than SemanticLock. This would support **H1** (*that SemanticLock will have similar performance as Pattern*).



3	4	5	6	7	8	9	10	11
LoginSpeed	A	X	B	X	C	X	D	AXD
2587	406	270	205	632	124	523	123	2283
1488	322	540	157	23	118	0	0	1160
1394	344	233	223	284	117	0	0	1201
1374	243	256	346	239	114	0	0	1198
939	270	168	199	164	29	0	0	830
2077	355	183	376	230	28	425	173	1770
849	306	216	264	0	0	0	0	786

Fig. 14

The subjective data indicated that our participants ranked SemanticLock as the slowest, but this was contrary to the quantitative results (see Fig. 13(a)). On the other hand, SemanticLock performed significantly better than the PIN and this is contrary to **H2** (*PIN system to firmly supersede the SemanticLock system*). Lastly **H3** (*SemanticLock will supersede the PIN-SHUFFLE in regards to unlock speed*) was confirmed as SemanticLock outperformed significantly PIN-Shuffle in speed. Our initial impression was that SemanticLock was faster than Pattern based on our observations. We were surprised when we saw the results.

Upon closer evaluation of the Login Speed data for the SemanticLock, we discovered a minor discrepancy between the final recorded Login Speed (see Fig. 14 (c), the ***LoginSpeed*** column) time and the accumulative sum ($AXD = A + X + B + X + C + X + D$) of the time used to actually move each image that make up a password token (see Fig. 14 (c)). "X" is the interval between the conclusion of moving the previous object and starting the movement of the next object. Columns A,B,C,D represent the time in milliseconds to move an object on the screen. The differences (LoginSpeed minus AXD) ranged from 63 to 2400 milliseconds. The value of AXD is shown as *SemanticLock Actual* in (see Fig. 14 (a)). Upon further examination, we found that this additional time was in fact used by the underlying Android Operating System and graphic rendering processes. The same effect was also discovered by Harbach et al. [13]. The effect is caused by temporal resolution issues of Android's threading and process scheduling. This discrepancy was only found within the SemanticLock data due to its

heavy reliance of graphics and the movement of images across the screen. We explored the possibilities of better performance (see Fig. 14 (c)) from the SemanticLock by generating results based on the actual time of movement (AXD) of all the objects that the user selected for the password token. We argue that this inherent temporal issue affecting the Android system may be the cause of the discrepancy between the quantitative and qualitative result of the SemanticLock LoginSpeed. Then additionally we attempted to extrapolate what the Login Speed of the SemanticLock might be possible if participants were familiar with the technique and had a shorter intra-movement delay time. For this simulation we found the statistical mode value for the intra-object movement delay time X, which is 243 milliseconds and made all intra-object delay time greater than 0 to become 243, this is shown as *SemanticLock Normalized* in (see Fig. 14 (a)). We also evaluated a *SemanticLock BestCase* scenario where all intra-object delay time will be zero. As you can observe from (see Fig. 14 (a)), the "best case" SemanticLock result indicates that we could improve on our system and have it supersede Pattern authentication in Login Speed.

8.2 Error rates

Our participants experienced the lowest error rate when seated and using their preferred Hand Input posture. Interestingly we also discovered that during the walking session PIN-Shuffled had the lowest encumbered error rate, while PIN had the lowest error rate across all techniques. Participants ranked the techniques based on how easy was to recover from errors in this order: Pattern (43%), PIN (17%), and SemanticLock (9%).

8.3 Memorability Test

Our participants displayed varying levels of difficulty in recalling their passwords. 70% of the participants did not recall their Pattern passwords, 50% did not recall their PIN passwords while 10% did not recall their SemanticLock password. This is an indication that the SemanticLock was more memorable to the participants.

8.4 Key Lessons

Our study has provided us data from which we have learnt the following:

- Graphical authentication systems based on discrete and continous movements out performs other authentication systems solely based on either. The potential of SemanticLock to be faster that the PATTERN is attributed to these dual movement properties.
- Authentication systems based on core graphical tokens with mnemonic properties result in higher memorability values.
- The SemanticLock performed excellently during the walking test. The results for both walking encumbered and unencumbered were satisfactory.
- Error recovery is high subjected to system design. We determined that graphical user interactivity and user familiarity greatly reduces the error rates
- The Semantic lock had the smallest pre-login delay time, which means that the participants found it easy to recall their password than with the other techniques.
- The Semantic lock performed better than the PIN lock while having a slightly similar performance with the Pattern lock.
- The type of device used by the participants i.e. phone or tablet had no effect on their performance.
- The hand posture used by participants affected their Login speed performance.

9 STUDY LIMITATIONS

Since we only had 3 weeks to perform this study we were not able to evaluate the long term memorability effects and also training effects on the techniques. We also believe that the SemanticLock performance would have

benefited from a longer term study period. In regards to generalization, it is important to know that the sample size may have had effect on the results, but due to adequate planning of the study and large numbers of trials we can maintain that our data is valid.

10 CONCLUSION AND FUTURE WORK

In this study, we explored a new screen lock concept based on semantic constructs; we used a set of graphical images as password tokens, this also enhances password memorability. The user is able to create a password using a quick action of dragging and dropping image tokens into their respective positions either as a discrete movement or in a continuous flow on the touchscreen. The large number of possible semantic constructs derived from the positioning of the image tokens and the varieties of images to choose from gave our system a theoretically large password space while providing high resistance to smudge attacks, a weakness the other authentication systems were prone to.

During our three weeks user study we engaged 21 participants and provided them our SemanticLock and other authentication systems to run a range of comparative test, whose results have been discussed earlier in this paper. In summary the SemanticLock was able to perform superiorly to the PATTERN and PIN authentication techniques, our results encouragingly warrants future improvements of this concept.Our future work is to continue and expand on the areas of the technique where improvements can be achieved, it is also equally important to expand the size of the participants and the period of the research while further integrating the SemanticLock system as the default authentication mechanism; so that our future participants can enjoy a more natural usage experience during the next future study.

REFERENCES

[1] Yomna Abdelrahman, Mohamed Khamis, Stefan Schneegass, and Florian Alt. 2017. Stay Cool! Understanding Thermal Attacks on Mobile-based User Authentication. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3751–3763. https://doi.org/10.1145/3025453.3025461

[2] Yomna Aly, Cosmin Munteanu, Stefania Raimondo, Alan Yusheng Wu, and Molly Wei. 2016. Spin-lock Gesture Authentication for Mobile Devices. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '16)*. ACM, New York, NY, USA, 775–782. https://doi.org/10.1145/2957265.2961863

[3] Panagiotis Andriotis, Theo Tryfonas, George Oikonomou, and Can Yildiz. 2013. A Pilot Study on the Security of Pattern Screen-lock Methods and Soft Side Channel Attacks. In *Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '13)*. ACM, New York, NY, USA, 1–6. https://doi.org/10.1145/2462096.2462098

[4] Adam J. Aviv, Devon Budzitowski, and Ravi Kuber. 2015. Is Bigger Better? Comparing User-Generated Passwords on 3x3 vs. 4x4 Grid Sizes for Android's Pattern Unlock. In *Proceedings of the 31st Annual Computer Security Applications Conference (ACSAC 2015)*. ACM, New York, NY, USA, 301–310. https://doi.org/10.1145/2818000.2818014

[5] Andreas Bulling, Florian Alt, and Albrecht Schmidt. 2012. Increasing the Security of Gaze-based Cued-recall Graphical Passwords Using Saliency Masks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 3011–3020. https://doi.org/10.1145/2207676.2208712

[6] Daniel Buschek, Fabian Hartmann, Emanuel von Zezschwitz, Alexander De Luca, and Florian Alt. 2016. SnapApp: Reducing Authentication Overhead with a Time-Constrained Fast Unlock Option. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 3736–3747. https://doi.org/10.1145/2858036.2858164

[7] Ashley A. Cain, Steffen Werner, and Jeremiah D. Still. 2017. Graphical Authentication Resistance to Over-the-Shoulder-Attacks. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 2416–2422. https://doi.org/10.1145/3027063.3053236

[8] Seunghun Cha, Sungsu Kwag, Hyoungshick Kim, and Jun Ho Huh. 2017. Boosting the Guessing Attack Performance on Android Lock Patterns with Smudge Attacks. In *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security (ASIA CCS '17)*. ACM, New York, NY, USA, 313–326. https://doi.org/10.1145/3052973.3052989

[9] Hsin-Yi Chiang and Sonia Chiasson. 2013. Improving User Authentication on Mobile Devices: A Touchscreen Graphical Password. In *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '13)*. ACM, New York, NY, USA, 251–260. https://doi.org/10.1145/2493190.2493213

[10] David Dobbelstein, Gabriel Haas, and Enrico Rukzio. 2017. The Effects of Mobility, Encumbrance, and (Non-)Dominant Hand on Interaction with Smartwatches. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers (ISWC '17)*. ACM, New York, NY, USA, 90–93. https://doi.org/10.1145/3123021.3123033

[11] Paul Dunphy, Andreas P. Heiner, and N. Asokan. 2010. A Closer Look at Recognition-based Graphical Passwords on Mobile Devices. In *Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS '10)*. ACM, New York, NY, USA, Article 3, 12 pages. https://doi.org/10.1145/1837110.1837114

[12] Shimin Feng, Graham Wilson, Alex Ng, and Stephen Brewster. 2015. Investigating Pressure-based Interactions with Mobile Phones While Walking and Encumbered. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '15)*. ACM, New York, NY, USA, 854–861. https://doi.org/10.1145/2786567.2793711

[13] Marian Harbach, Alexander De Luca, and Serge Egelman. 2016. The Anatomy of Smartphone Unlocking: A Field Study of Android Lock Screens. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4806–4817. https://doi.org/10.1145/2858036.2858267

[14] Marian Harbach, Emanuel von Zezschwitz, Andreas Fichtner, Alexander De Luca, and Matthew Smith. 2014. It's a Hard Lock Life: A Field Study of Smartphone (Un)Locking Behavior and Risk Perception. In *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. USENIX Association, Menlo Park, CA, 213–230. https://www.usenix.org/conference/soups2014/proceedings/presentation/harbach

[15] G. Kovelamudi, J. Zheng, and S. Mukkamala. 2016. Scramble or not, that is the question a study of the security and usability of scramble keypad for PIN unlock on smartphones. In *2016 IEEE/CIC International Conference on Communications in China (ICCC)*. 1–6. https://doi.org/10.1109/ICCChina.2016.7636862

[16] William Melicher, Darya Kurilova, Sean M. Segreti, Pranshu Kalvani, Richard Shay, Blase Ur, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Michelle L. Mazurek. 2016. Usability and Security of Text Passwords on Mobile Devices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 527–539. https://doi.org/10.1145/2858036.2858384

[17] Nicholas Micallef, Mike Just, Lynne Baillie, Martin Halvey, and Hilmi Güneş Kayacik. 2015. Why Aren'T Users Using Protection? Investigating the Usability of Smartphone Locking. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 284–294. https://doi.org/10.1145/2785830.2785835

[18] Keaton Mowery, Sarah Meiklejohn, and Stefan Savage. 2011. Heat of the Moment: Characterizing the Efficacy of Thermal Camera-based Attacks. In *Proceedings of the 5th USENIX Conference on Offensive Technologies (WOOT'11)*. USENIX Association, Berkeley, CA, USA, 6–6. http://dl.acm.org/citation.cfm?id=2028052.2028058

[19] Alexander Ng. 2014. The Effects of Encumbrance on Mobile Interactions. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 405–406. https://doi.org/10.1145/2628363.2634268

[20] Alexander Ng and Stephen Brewster. 2013. The Relationship Between Encumbrance and Walking Speed on Mobile Interactions. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 1359–1364. https://doi.org/10.1145/2468356.2468599

[21] Alexander Ng, Stephen A. Brewster, and John H. Williamson. 2014. Investigating the Effects of Encumbrance on One- and Two- Handed Interactions with Mobile Devices. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1981–1990. https://doi.org/10.1145/2556288.2557312

[22] Alexander Ng, John Williamson, and Stephen Brewster. 2015. The Effects of Encumbrance and Mobility on Touch-Based Gesture Interactions for Mobile Phones. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 536–546. https://doi.org/10.1145/2785830.2785853

[23] Pekka Parhi, Amy K. Karlson, and Benjamin B. Bederson. 2006. Target Size Study for One-handed Thumb Use on Small Touchscreen Devices. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '06)*. ACM, New York, NY, USA, 203–210. https://doi.org/10.1145/1152215.1152260

[24] Oriana Riva, Chuan Qin, Karin Strauss, and Dimitrios Lymberopoulos. 2012. Progressive Authentication: Deciding when to Authenticate on Mobile Phones. In *Proceedings of the 21st USENIX Conference on Security Symposium (Security'12)*. USENIX Association, Berkeley, CA, USA, 15–15. http://dl.acm.org/citation.cfm?id=2362793.2362808

[25] Craig Stewart, Eve Hoggan, Laura Haverinen, Hugues Salamin, and Giulio Jacucci. 2012. An Exploration of Inadvertent Variations in Mobile Pressure Input. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '12)*. ACM, New York, NY, USA, 35–38. https://doi.org/10.1145/2371574.2371581

[26] Sebastian Uellenbeck, Markus Dürmuth, Christopher Wolf, and Thorsten Holz. 2013. Quantifying the Security of Graphical Passwords: The Case of Android Unlock Patterns. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security (CCS '13)*. ACM, New York, NY, USA, 161–172. https://doi.org/10.1145/2508859.2516700

[27] Emanuel von Zezschwitz, Alexander De Luca, Bruno Brunkow, and Heinrich Hussmann. 2015. SwiPIN: Fast and Secure PIN-Entry on Smartphones. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1403–1406. https://doi.org/10.1145/2702123.2702212

[28] Emanuel von Zezschwitz, Anton Koslow, Alexander De Luca, and Heinrich Hussmann. 2013. Making Graphic-based Authentication Secure Against Smudge Attacks. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces (IUI '13)*. ACM, New York, NY, USA, 277–286. https://doi.org/10.1145/2449396.2449432

[29] Graham Wilson, Stephen A. Brewster, Martin Halvey, Andrew Crossan, and Craig Stewart. 2011. The Effects of Walking, Feedback and Control Method on Pressure-based Interaction. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 147–156. https://doi.org/10.1145/2037373.2037397

[30] Nur Haryani Zakaria, David Griffiths, Sacha Brostoff, and Jeff Yan. 2011. Shoulder Surfing Defence for Recall-based Graphical Passwords. In *Proceedings of the Seventh Symposium on Usable Privacy and Security (SOUPS '11)*. ACM, New York, NY, USA, Article 6, 12 pages. https://doi.org/10.1145/2078827.2078835