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

ILESANMI A OLADE, Xi’an Jiaotong-Liverpool University
HAI-NING LIANG, Xi’an Jiaotong-Liverpool University
CHARLES FLEMING, Xi’an Jiaotong-Liverpool University

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. Passwords are entered via the familiar and quick action of dragging and dropping the images into their respective positions on the touchscreen. It is well known that for locking mechanisms such as PIN or PATTERN that users tend to pick memorable passwords such as dates or simple geometric shapes, significantly reducing the effective password space for these mechanisms. 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. While graphical passwords have been shown in some cases to have lower entropy than other password types, we avoid this problem by performing a series of studies to understand which images and image pairs users prefer and selecting images that avoid any type of explicit or implicit bias, resulting in an effective password space that is essentially the same as the total password space.

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

CCS Concepts: • Human-centered computing → Smartphones; Mobile devices;

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

ACM Reference format:
Ilesanmi A Olade, Hai-ning Liang, and Charles Fleming. 2016. SemanticLock: An authentication method for Mobile devices using semantically-linked images. 1, 1, Article 1 (January 2016), 24 pages.
DOI: 10.1145/nnnnnn.nnnnnnn

1 INTRODUCTION

Mobile devices, being the de facto personal communication device, are ubiquitous within our society [50]. We depend on these devices to store substantial amounts of confidential information and perform activities such as emailing, social networking, personal internet banking, and entertainment. All mobile devices manufactured in the last decade come with a default set of authentication or login mechanisms. Research by Micallef et al[35],
shows that over 64% of users chose not to secure or use an authentication system on their mobile devices [28]. However, 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 [33], suggest that users dislike the inconvenience of repeatedly unlocking their mobile devices. Moreover, those who choose to use their mobile device unlocking mechanisms are discouraged by the time and effort it takes to unlock these devices or the frustrating unlock failure errors observed [41]. In general, research has shown that the behaviour, engagement, and interest of the users have a major impact on the effective security level of their mobile devices, with many users preferring to sacrifice security for convenience [13]. Studies by [14, 32], indicate that the distribution of text passwords chosen by users effectively have very low entropy, meaning that the actual space of passwords most users choose from is much smaller than the total space available. Prominent authentication systems such as PIN [27, 30, 48] and PATTERN [27, 46, 53] have being extensively studied; and have a large body of existing literature [48, 49]. The PIN authentication system (see Fig.1(a)), which is a numeric display of numbers inputted by discrete touches on the screen, and the PATTERN authentication system (see Fig.1(b)), which is a “grid-like” display of nodes whose password pattern is selected by a continuous finger movement across the screen to connect the secret password nodes, are both plagued with numerous usage and security issues [1, 3, 34, 53]. The popularity of touch-screen based mobile devices allows for graphical authentication techniques that offer possibilities of providing passwords that are effectively stronger than text passwords.

Recently, researchers have developed and studied various graphical authentication systems [2, 5, 19, 43, 50] that take advantage of the inherent human memorability properties and have attempted to mitigate factors such as low password distribution; low unlock speed, medium-to-low entropy and other biases, without much success. Our technique strives to improve on memorability [16, 27] while significantly increasing unlock speed, password distribution and password entropy. 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 finger movements to create a semantically

![Fig. 1. Prominent mobile device authentication systems](image-url)
memorable story that represents a password (see Fig. 2). SemanticLock can create a strong memorable password with just two discrete finger movements allowing the user to construct a semantically meaningful password quickly (see Fig. 2(b)) from the provided images. In the SemanticLock scheme, a password is a sequence of $k$ images selected by the user to make a "story", from a single set of $n \geq k$ images, each non-intrinsically related and placed in position $p$; in one of four locations around a pre-existing image. For two pairs selected from $n$ images this yields $4n(n-1)(4(n-2)(n-3) + 6(n-2))$ choices, where the second term in the inner sum accounts for one of the images for the first pair being selected as the target image for the second pair. For the mobile devices such as phones, six images allows for comfortable use, yielding 8640 passwords, so the practical implementation will rely on login false attempt limitation counter-measures.

To increase the entropy of the selected password distribution, we ensured that we reduced password image bias by performing a two week preliminary study with the goal of eliminating disproportionately popular images and image pairs. In that study, our participants were required to match intrinsically related password images from a set of 40 images that were initially selected from diverse categories (see Fig. 3). We subsequently obtained 6 “least intrinsically” related images from that study and used them during another 7 days SemanticLock password creation study (see Fig. 4). This user study positively reveals that our SemanticLock significantly has a highly uniform password distribution, signifying a high password entropy.

Fig. 2. Semantic-lock: (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 (A,B).

In designing the SemanticLock system; we set out to develop a system that was easy to use and quick to login. Therefore our primary focuses were 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. 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 only requires two short swipes rather than one continuous long swipe.

For our study, we utilized the dataset we collected during a three week period and, we show that while SemanticLock can be practically more secure than the PIN and PATTERN authentication systems, its performance is significantly better than the PIN and similar to PATTERN under normal circumstances but exceedingly better in ideal scenarios.

The rest of this paper is structured as follows. We describe related work in Section 2. In Section 3, we describe our methodology in more detail, such as the preliminary web-based studies, the graphical password schemes that we evaluated during the mobile device study, and experimental design. In Section 4, we introduced our data sources and data collection models. In Section 5, 6, 7, 8 we discuss issues and findings and present our results for them, including study limitations in Section 9. Finally, we conclude in Section 10.

2 RELATED WORKS ON AUTHENTICATION METHODS
User authentication and access control are very important in today’s electronic world. The advent of personal computing and mobile devices has made security a foremost consideration in the design and usage of these devices. While authentication can exist in many forms, there are three core types of authentication categories with which a user can be identified by a system. These categories are namely: What you know, What you have and What you are. The practical implementation of these categories are the text and graphical passwords, token based passwords, and biometric passwords. We shall examine the history and various studies pertaining to text and graphical password implementations.

2.1 Text Based Passwords
Alphanumeric text-based passwords have dominated human-computer authentication since the 1960s [20], where keyboards were used to input user passwords. With the emergence of mobile devices with 10 digit keypads [48], the use of numeric-based PIN passwords became mainstream. First generation touch-screen based smart-phones featured numerous variants of PIN-based password systems [16, 26], and has been used by all mobile device form factors [2, 14, 23] and remains very popular with users. Although the text-based and the PIN passwords have high theoretical password spaces, numerous studies, such as those by Bonneau et. al [11], and Melicher et.al [32] show that the practical password spaces and entropy are very low due to user security behaviours. In a study by [11], it was shown, based on available large public dataset of PINs, that 29% of the selected 4-PIN and 6-PIN passwords correspond to a date based sequence. This reduces the practical password space of PIN passwords.

For many years the security literature lacked sound methodology and ecological validity [24] to answer elementary questions about practical password distribution, or the effects of demographic properties on their outcome, and it remained an open question as to the extent to which passwords are weak due to a lack of motivation or inherent user limitations [9, 10]. The massive disclosure of millions of real-life user passwords in hacked password databases [8, 29, 47] from several websites such as RockYou, Yahoo, Hotmail, Flirtlife and Computerbits, exposed an enormous gap between a real password distribution and the theoretical space of passwords. Furthermore, analyses by Malone et.al [31] observed that security motivations such as registering payment cards or supplying sensitive financial information did not affect the users tendency to create weak passwords. In the final analysis, practical user passwords distribution is skewed to provide low password entropy and protection. Additionally, studies by Melicher et.al [32] confirm that this pattern of skewed password
distribution and low password entropy is worse with mobile devices users due to additional restrictive factors inherent with using mobile devices.

2.2 Graphical Passwords

A graphical password, a term introduced by Blonder [7], is an authentication system that is presented to the user via a graphical user interface (GUI), and from a smart mobile device perspective, this GUI includes a touch-screen system that enables easy interaction with the objects displayed on the GUI. Graphical passwords provide a promising alternative to traditional alphanumeric passwords. They are attractive and intuitive since people usually remember shapes and images better than random words or text. In recent years, various studies have categorized graphical authentication methods into 3 main categories, which are:

**Recall:** The Recall graphical authentication system gets its origin from works done on Draw-a-Secret [53], Pass-Go [44] and other similar systems. It is shown to be a memory intensive task [6] due to the fact that the secret diagram or pattern initially drawn by the user has to be entirely remembered and reproduced, but 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 [23, 46]. The Android Pattern password system is recall-based.

**Recognition:** The recognition graphical authentication systems revolves around the ability of the user to 'recognize' sets of images from among decoys, that had been selected earlier during the initial creation of the passwords. Recognition based systems such as Passfaces [12, 40], Déjà vu [21] have been extensively studied.

**Cued-recall:** Cued-recall based systems exploit various studies that conclude that the human memory holds information that may be available yet inaccessible for retrieval without the proper trigger or catalyst [18]. This system based on the idea that pictorial indicators can simplify the task of recall for a user [46]. Cued-recall based systems such as PassPoint [51] and Cue Click Points (CCP) [17] have been extensively studied.

In this paper, our study compares the performance metrics of the SemanticLock authentication system when compared with the Android Pattern and PIN authentication systems. We therefore examined existing literature to find studies similar to ours.

In a study by von Zezschwitz et al. [50], three custom graphical authentication systems were compared against the PATTERN authentication system. The aim was to study their prototypes’ unlock speed, level of memorability, usability and user acceptance. Results confirmed that PATTERN authentication system was superior to their proposed prototypes in regards to unlock speed, and performed comparatively similar in regards to usability, user acceptance and memorability but was considered less secure by the users. The PIN authentication system was not included in their study. The effective password distribution or password space was not evaluated in this study. In a later study, von Zezschwitz et al. [49] compared the PIN and Pattern authentication system, and the results indicated that PIN had a faster unlock speed and smaller error rate, but the Pattern was more usable, memorable and likeable. However, studies of user Pattern password creation by [4, 45, 46], found empirically that there is a high bias in the Pattern selection process resulting in low entropy and a practical effective security of less than a three digit randomly-assigned PIN.
More recently Aly et al [2] introduced SpinLock, 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. This 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.

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 divided [35]. Ng et al [38] in their initial study, discovered that mobile phone users simultaneously carry or hold other items while interacting with their devices in public and these users tend to carry shopping bags and boxes often. Additional studies by Ng et al. [36], had users clicking on “crosses” or target points that randomly appear on the screen to study the effects of encumbrance and walking on the user’s targeting accuracy while walking and compared to when the user stood still. They determined that targeting error rates increased by 112% for the those walking. Wilson et al.[52] and others [22, 25, 42] 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[39].

3 METHODOLOGY

We employed two strategies in an attempt to achieve the desired features and functions of our previously described SemanticLock system. Pre-system development analysis and experiments are required in order to derive initial icon sets. Therefore two studies were conducted, a web-based study and an Android-based study. Both studies are discussed below.

3.1 Web-based Study : Password icon selection and practical space evaluation

For this study, we utilized a web-based interface that was designed using HTML5, PHP and MySQL database back-end technologies. This allowed us to implement icon drag-n-drop actions that are common on touch-screen based mobile devices. This web-based approach allowed us to collect a large amount of data from our participants in various locations and use this data to determine various types of icon selection and password space evaluations. Although web-based experiments are harder to control than laboratory or supervised field studies [6], this channel of data collection meets our requirements and offers numerous advantages. The below sections offer further details of our experiments.

3.1.1 Goals.

As part of our goals in the design of our Semantic-lock system, our initial intention is to avoid any implicitly induced biases in the researcher’s selection of the password icons that may lower the entropy or reduce the achievable password space[20]. In general, security experts have observed that an authentication systems theoretical password space is never optimally achieved during practical usage[15], and there is a need to determine the actual practical password space that determines the ecological validity of such an authentication system. We define two stages of experiment to achieve the above stated objectives, and implemented these stages with two different groups of participants. The output of the analysis of the dataset collected in the first stage was utilized during the second stage. No demographic data was collect from the participant during these stages.

3.1.2 Participants.
For Stage 1 study, we engaged 372 participants, a large number of them were students at the university campus, but we ensured that 40% non-student persons, of all age groups, also participated in this study. Our participant group included 45% female users; we did not collect any further demographic information such as academic background, computer skills or their experience with mobile devices or authentication systems. The participants were not informed of the final purpose of the study, they were only told to 'pair' icons they felt were related, the reason or logic of this relationship was based on their discretion.

For Stage 2 study, we engaged 184 participants, 70% were students within the university campus and the rest were non-students. Our web portal included a 3 minute training video, and each participant was encouraged to watch the video before attempting to create passwords. We advised our participants to create at least 3 passwords each. Our participant group included 18% female users; we did not collect any further demographic information such as academic background, computer skills or their experience with mobile devices or authentication systems.

3.1.3 Experiment Design [Stage 1]: Acquisition of Independent password Icons. Our initial process was to provide a set of 40 icons that were drawn from various categories and genres. We pointedly avoided icons that had major gender significant colors, cultural, national or religious relevance. Our participants were then presented with a web-based interface that displayed these icons on a 10 by 4 grid; with each icon randomly positioned in different grid-cells during every selection session to prevent locational bias. Participants were required to create 10 sets of "icons-pairs" that they believed were related by dragging these icons into the provided ‘pairboxes’ (see fig. 3). Each participant was allowed multiple iterations.

We analyzed the 1039 collected pair-datasets to extract 6 icons that were the least intrinsically related. These "non-intrinsically" related icons were used in the next stage of the experiment.
3.1.4 Experiment Design [Stage 2]: Evaluation of practical password space.

Our primary goal was to quantify the effect of a participant’s choice on the security of passwords chosen. Every authentication scheme has an entropy and the strength of such entropy is determined by the probability distribution associated with the password space. Ideally this distribution is approximately uniform. At this stage of our experiment we presented a web-based interface displaying the six derived non-intrinsically related password icons on a 9 by 6 celled grid to our participants (see fig. 4(a) ). Our participants were required to create several semantic passwords with the password icons by dragging a chosen icon to the left, top, right or bottom position of an associated stationary icon (see fig. 4(b) ) .

![Fig. 4. Semantic-lock Web-based Password Creator: (a) Default view of icon placement. (b) Creating Password: the user drags the “cheese” to meet the stationary “bottle” icon. In this case, “cheese” is also dragged to right side of “bottle”. Lastly a three-icon password is shown (see black circle)](image)

3.1.5 Data Collection and Analysis.

The data collected from the participants during stage two of the web-based study has been analyzed to derive the below information:

**Password Icon distribution:** Frequency analysis was performed on the semantic password data sets collected during stage two of the study. Each semantic password is made up of unique icons selected from the 6 initial password icons. From our data set of 785 semantically created passwords, our analysis suggests that the choice of each of the six password icons is uniformly distributed (see fig. 5(a)), which is a strong indication that our semantic-lock system does not suffer any bias that may affect its practical password space.

**Password Icon pair distribution:** As each semantic password is composed of two or more sets of password icons, we pre-processed the collected data sets and decomposed semantic passwords that consist of more than two password icons into two pairs of password icons and performed frequency analysis on
these password icon pairs. Our analysis shows a uniform distribution of password icon pairs (see fig. 6). Our participant’s selections were not biased or skewed.

**Password Icon-pair position distribution:** As explained previously, all semantic passwords, consists of two or more sets of password icons. As such, these password icons are used to create semantic passwords by dragging a selected password icon to a "resting position" next to the stationary password icon. This "resting position" could either be the left, top, right or bottom of a stationary password icon. We analyzed the collected positional data sets to determine if our participants displayed a bias in their choice of "resting positions". Our analysis indicated that the participant selection of "resting positions" was fairly uniform with a small bias against the left, which is somewhat expected from predominantly right handed users.

3.2 Mobile Device Study

Our mobile device study made use of the Android platform. We developed a mobile version of the interface that was used during our web-based study. We also developed Android versions of the Pattern and PIN lock authentication systems since these authentication systems will be our baseline or control for this study due to their popularity and large body of research literature about their performance. 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. 7).

3.2.1 Goals. Our goal during this three week study, which involved 21 participants in an indoor environment, was to collect both qualitative and quantitative data which would provide insight into our participant’s perception of the likeability, usability, memorability and login speed of the 3 authentication approaches:

- SemanticLock
- Pattern Lock

![Chart showing analysis of password Icon usage.](a)

![Chart showing analysis of dragged Icon resting or drag-to position on the stationary Icon.](b)
Fig. 6. **Icon-Pair selection Analysis**: The distribution of "Icon-Pair" selection within the password icon data sets. The chart shows a "uniform" distribution, indicating a strong password entropy.

**• PIN**

The prototypes, shown in (Fig. 1 and Fig. 2), meet our goal of ensuring compatibility with Android 6.0 and above, while meeting the requirements of working on 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 further analysis.

**3.2.2 Participants.** We recruited 21 participants (15 females) from a local university. The data from our pre-testing survey reveals that 51% of the participants were between the ages of 17 to 22 and all our participants were right-handed. All were active users of iPhone (31%) and Android (66%) mobile phones. 55% of them used a phone with fingerprint sensor, while 17% used the PIN, 14% Pattern and, 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.

**3.2.3 Experimental Design.**

Our goal was to compare three main techniques and their interactions with other independent variables. To do this, we followed a within-participants design. Below are the variables we are tracking:
SemanticLock: An authentication method for Mobile devices using semantically-linked images

The independent variables are:
- Technique
- Device Form-Factor
- Physical Posture
- Hand Posture

The dependent variables are:
- Login Speed
- Pre-Login Delay Time
- Error Rate
- User usability and acceptance

Technique: Our experiment compared three techniques. 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 digit password tokens, having a possible theoretical password space of 10,000. 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 5 connected nodes, giving us a theoretical space range of 7000. For the SemanticLock technique we issued a series of randomly generated passwords along with some semantic interpretation that would enhance memorability, with a theoretical password space of 8640 possibilities.

Device Form-Factor: Mobile devices come in various dimensions. We used two different form-factors:
- 5.2” LG Nexus 5X phone
- 10.2” Google Pixel C tablet

The tablet was only used during the Seated session (Fig. 8 (a)) of the experiment, while the LG phone was used for all sessions(Fig. 8 (b),(c)).
Physical Posture: Studies show that the physical posture of users has an effect on the way they use the devices [35, 36, 38]. Therefore in this study we included 3 physical postures:

Seated: This posture 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. 8(a)).

Walking Unencumbered: This posture implied that the person operating the mobile device was also walking but without carrying any other objects with their hands or arms (see Fig 8(b)).

Walking Encumbered: This posture 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 8(c)).

Recent studies have shown that walking encumbered or unencumbered and operating a mobile device had shown significant effects on the usage pattern of mobile devices [37, 39, 42]

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. 9). 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 [37, 39].
3.2.4 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. 7a), after which they were allowed to practice each technique a couple of times. They 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.

**Week 1 (First Phase):** Each participant was required to answer a pre-test questionnaire before commencing the test (see Fig. 7b). 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. 7c) that collected qualitative data about how the participants’ perceived usability, error-handling, security and likeability of each technique. The next week’s session was a seated session and the participants used the techniques on the LG mobile phone and the Google tablet. The main independent variable was technique (PIN, Pattern and SemanticLock) and mobile form factor (phone and tablet). Each participant had to enter a total of 9 passwords per session, 3 for each Technique and participants were allowed a 60 second rest in between techniques to avoid fatigue.

**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.

**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 [37, 38] in which they examined the effect of mobility and encumbrance on

---

**Fig. 9. Input Hand Postures:** The most common hand postures when using mobile devices. These postures were tested during the Study.
participants using both one and two-handed 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. After the walking test (see Fig 8(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, while unlocking the device using each technique (see Fig 8(c)). The decision to use nylon bags was informed by the research done by Ng et al. [38]. In this phase, we sought to investigate the effect of mobility and encumbrance on the login speed, memorability and input errors while assessing the techniques with the 3 commonly used input postures as discovered in a research by [39].

4 DATA COLLECTION & MEASUREMENT

We collected data for a number of dependent variables and used this data to compare techniques.

4.1 Pre-Login Delay time: Memorability & Usability

Pre-login delay is the elapsed time between when the participant indicated that they were ready to start unlocking the device and the actual entry of the password. This data provides a view into evaluating the memorability and usability of the system. Studies by Stobert et al.[43] defined a direct relationship between memorability and pre-login delay time. We analyze this data to quantify the level of memorability and usability.

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 three successful trials to constitute a session while other techniques, such as 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.7 (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 SemanticLock performed better than the other techniques 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 Semantic-lock (807.06 ± 167.23 ms, p = .000) was significantly faster than Pattern and PIN (both p < .000).
SemanticLock: An authentication method for Mobile devices using semantically-linked images

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. 10 (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. 11 (a)).

![Login speed while seated with different device form](image1)

![Login Speed while Encumbered](image2)

(a) Login speed while seated with different device form
(b) Login Speed while Encumbered

Fig. 10. 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. 11 (a and b)). However, there was a significant interaction effect between

\[\text{Average login speed (milliseconds) across posture and technique}\]

| Pattern            | PIN  | Semantic-lock |
|--------------------|------|---------------|
| Seated (Tablet)    | 785  | 590           |
| Seated (Phone)     | 825  | 652           |
| Walking Thumb      | 1135 | 853           |
| Walking Index      | 916  | 708           |
| Walking 2 Thumbs   | 945  | 768           |
| Walking-E Thumb    | 1175 | 917           |
| Walking-E Index    | 800  | 910           |
| Walking-E 2 Thumbs | 873  | 655           |

Note: Walking-E = Walking Encumbered
physical posture and login technique ($F(8,1485) = 3.302, p = .001$), with participants having a faster speed using the Pattern method 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. 10 (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 postures (OneHandThumb, TwoHands2Thumbs, OneHandOtherIndex) during the testing of the Login Technique (see Fig. 9). 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. 12 (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 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 should give an indication of familiarity, memorability or ease of use of the
SemanticLock: An authentication method for Mobile devices using semantically-linked images

Techniques. SemanticLock had the lowest pre-login delay time across all hand input postures (see Fig. 12 (b)). The ANOVA test results 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 \((p = 0.624)\).

![Graph showing login speed and delay for different techniques](image)

(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. 12. Login speeds and delay for the three techniques

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 rate of the Pattern technique \((p = .925)\). Additionally, results shows that the PIN had the lowest error rates when walking unencumbered (see Fig. 13 (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. 13 (b)). Error rates classified by techniques show that Pattern \((27%)\) had the highest error rates, followed by SemanticLock \((19%)\), and PIN\((4%)\) (see Fig. 14 (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 subjective preferences. At the end a user ranking of all techniques was collected (see Fig. 14 (b)). The data we collected was analyzed using the 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.

7.1 Speed

Our participants experience with each 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. 15(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. 15(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 of 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. 15(b)). However, there was a significant increase in perceived ease of use between Pattern and SemanticLock ($Z = -3.140, p = 0.002$).

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. 15(d)).

8 DISCUSSION
8.1 Login Speed
Our participants performed better, but not significantly, in login speed using Pattern than SemanticLock.

The subjective data indicated that our participants ranked SemanticLock as the slowest, but this was contrary to the quantitative results (see Fig. 15(a)). We initially expected that SemanticLock would be faster than Pattern based on our observations, as SemanticLock only involves two swipes. Familiarity with the Pattern unlocking mechanism may explain part of this outcome.

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 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 continuous movements outperform other authentication systems solely on one or the other. 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.
- SemanticLock performed excellently during the walking test. The results for both walking encumbered and unencumbered were satisfactory.
- Error recovery is strongly influenced by system design. We determined that graphical user interactivity and user familiarity greatly reduces the error rates.
- The SemanticLock had the smallest pre-login delay time, which means that the participants found it easier to recall their password than other techniques.
- SemanticLock 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.
SemanticLock: An authentication method for Mobile devices using semantically-linked images

- The hand posture used by participants affected their login speed performance but this effect was uniform across techniques.

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 of the techniques. We 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 and our selection of icons gave it a large practical password space. During our three week user study we engaged 21 participants and provided them our SemanticLock and other authentication systems to run a range of comparative tests, whose results have been discussed earlier in this paper. To conclude, the SemanticLock generally had superior performance compared to PATTERN and PIN authentication techniques on key metrics. Our future work is to continue to 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/10.1145/2957265.2961865

[3] Panagiotis Andriotis, Theo Tryfonas, George Okonomou, and Can Yıldız. 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] Marios Belk, Andreas Pamboris, Christos Fidas, Christina Katsini, Nikolaos Avouris, and George Samaras. 2017. Sweet-spotting Security and Usability for Intelligent Graphical Authentication Mechanisms. In Proceedings of the International Conference on Web Intelligence (WI ’17). ACM, New York, NY, USA, 252–259. https://doi.org/10.1145/3106426.3106488

[6] Robert Biddle, Sonia Chiasson, and P.C. Van Oorschot. 2012. Graphical Passwords: Learning from the First Twelve Years. ACM Comput. Surv. 44, 4, Article 19 (Sept. 2012), 41 pages. https://doi.org/10.1145/2333112.2333114

[7] G. Blonder. 1996.. Graphical Password. In Lucent Technologies, Inc (1996.).

[8] J. Bonneau. 2012. The Science of Guessing: Analyzing an Anonymized Corpus of 70 Million Passwords. In 2012 IEEE Symposium on Security and Privacy. 538–552. https://doi.org/10.1109/SP.2012.49

[9] Joseph Bonneau. 2012. Statistical Metrics for Individual Password Strength. In Proceedings of the 20th International Conference on Security Protocols (SP’12). Springer-Verlag, Berlin, Heidelberg, 76–86. https://doi.org/10.1007/978-3-642-35694-0_10

[10] Joseph Bonneau, Cormac Herley, Paul C. van Oorschot, and Frank Stajano. 2015. Passwords and the Evolution of Imperfect Authentication. Commun. ACM 58, 7 (June 2015), 78–87. https://doi.org/10.1145/2699390
1:22  •  I. Olade et al.

[11] Joseph Bonneau, Sören Preibusch, and Ross Anderson. 2012. A Birthday Present Every Eleven Wallets? The Security of Customer-Chosen Banking PINs. In Financial Cryptography and Data Security, Angelos D. Keromytis (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 25–40.

[12] Sacha Brotstoff and M. Angela Sasse. 2000. Are Passfaces More Usable Than Passwords? A Field Trial Investigation. In People and Computers XIV — Usability or Else?, Sharon McDonald, Yvonne Waern, and Gilbert Cockton (Eds.). Springer London, London, 405–424.

[13] Daniel Buschek, Fabian Hartmann, Emanuel von Zeeschkwitz, 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

[14] 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/10.1145/3027063.3053286

[15] 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 (ASIACCS ’17). ACM, New York, NY, USA, 313–326. https://doi.org/10.1145/3052973.3052989

[16] 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

[17] Sonia Chiasson, Alain Forget, Robert Biddle, and P. C. van Oorschot. 2009. User Interface Design Affects Security: Patterns in Click-based Graphical Passwords. Int. J. Inf. Secur. 8, 6 (Oct. 2009), 387–398. https://doi.org/10.1007/s10207-009-0080-7

[18] Sonia Chiasson, P. C. van Oorschot, and Robert Biddle. 2007. Graphical Password Authentication Using Cued Click Points. In Proceedings of the 12th European Conference on Research in Computer Security (ESORICS’07). Springer-Verlag, Berlin, Heidelberg, 359–374. http://dl.acm.org/citation.cfm?id=2393847.2393880

[19] Darren Davis, Fabian Monrose, and Michael K. Reiter. 2004. On User Choice in Graphical Password Schemes. In Proceedings of the 13th Symposium on USENIX Security Symposium - Volume 13 (SSYM’04). USENIX Association, Berkeley, CA, US, 11–11. http://dl.acm.org/citation.cfm?id=1251375.1251386

[20] Alexander De Luca, Katja Hertzschuch, and Heinrich Hussmann. 2010. ColorPIN: Securing PIN Entry Through Indirect Input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’10). ACM, New York, NY, USA, 1103–1106. https://doi.org/10.1145/1753326.1753490

[21] Rachna Dhamija and Adrian Perrig. 2000. D´eJ`a Vu: A User Study Using Images for Authentication. In Proceedings of the Ninth Conference on USENIX Security Symposium - Volume 9 (SSYM’00). USENIX Association, Berkeley, CA, US, 4–4. http://dl.acm.org/citation.cfm?id=1253006.1251310

[22] 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

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

[24] Sascha Fahl, Marian Harbach, Yasemin Acar, and Matthew Smith. 2013. On the Ecological Validity of a Password Study. In Proceedings of the Ninth Symposium on Usable Privacy and Security (SOUPS ’13). ACM, New York, NY, USA, Article 13, 13 pages. https://doi.org/10.1145/2516670.2516767

[25] 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

[26] S M Taibai Haque, Matthew Wright, and Shannon Scielzo. 2013. Passwords and Interfaces: Towards Creating Stronger Passwords by Using Mobile Phone Handsets. In Proceedings of the Third ACM Workshop on Security and Privacy in Smartphones & Apps; Mobile Devices (SPS ’13). ACM, New York, NY, USA, 105–110. https://doi.org/10.1145/2516760.2516767

[27] Marian Harbach, Alexander De Luca, and Serg 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

[28] Marian Harbach, Emanuel von Zeeschkwitz, 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 Usable Privacy and Security (SOUPS 2014). USENIX Association, Menlo Park, CA, 213–230. https://www.usenix.org/conference/soups2014/proceedings/presentation/harbach

[29] Markus Jakobsson and Mayank Dhaman. 2012. The Benefits of Understanding Passwords. In Proceedings of the 7th USENIX Conference on Hot Topics in Security (HotSec ’12). USENIX Association, Berkeley, CA, USA, 10–10. http://dl.acm.org/citation.cfm?id=2372387.2372397
I. Olade et al.

[49] Emanuel von Zezschwitz, Paul Dunphy, and Alexander De Luca. 2013. Patterns in the Wild: A Field Study of the Usability of Pattern and Pin-based Authentication on Mobile Devices. In Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI ’13). ACM, New York, NY, USA, 261–270. https://doi.org/10.1145/2493190.2493231

[50] 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

[51] Susan Wiedenbeck, Jim Waters, Jean-Camille Birget, Alex Brodskiy, and Nasir Memon. 2005. PassPoints: Design and longitudinal evaluation of a graphical password system. International Journal of Human-Computer Studies 63, 1 (2005), 102 – 127. https://doi.org/10.1016/j.ijhcs.2005.04.010

[52] 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

[53] 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