Graphical Passwords in the Wild – Understanding How Users Choose Pictures and Passwords in Image-based Authentication Schemes

Florian Alt¹, Stefan Schneegass², Alireza Sahami Shirazi³*, Mariam Hassib¹, Andreas Bulling⁴

¹University of Munich – Group for Media Informatics (Amalienstrasse 17, 80333 München, Germany)
²University of Stuttgart – VIS (Pfaffenwaldring 5a, 70569 Stuttgart, Germany)
³Yahoo Labs (701 First Avenue, Sunnyvale, CA 94089, United States)
⁴Max Planck Institute for Informatics (Campus E1 4, 66123 Saarbrücken, Germany)

ABSTRACT
Common user authentication methods on smartphones, such as lock patterns, PINs, or passwords, impose a trade-off between security and password memorability. Image-based passwords were proposed as a secure and usable alternative. As of today, however, it remains unclear how such schemes are used in the wild. We present the first study to investigate how image-based passwords are used over long periods of time in the real world. Our analyses are based on data from 2318 unique devices collected over more than one year using a custom application released in the Android Play store. We present an in-depth analysis of what kind of images users select, how they define their passwords, and how secure these passwords are. Our findings provide valuable insights into real-world use of image-based passwords and inform the design of future graphical authentication schemes.

Author Keywords
Graphical passwords; images; security

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation: User Interfaces—Input devices and strategies; K.6.5 Computing Milieux: Security and Protection—Authentication

INTRODUCTION
Secure user authentication is essential in light of the increasing amount of sensitive information available on today’s smartphones. Image-based password schemes offer a good trade-off between security and password memorability and recently received considerable attention in research [4, 3, 5, 22] and industry [12]. These schemes require the user to select a number of password points either across a sequence of multiple or within a single image (for example the church door, the top of the small palm, the top of the left large palm, and the front left window in the bell tower, see Figure 1). The underlying image helps users to remember their password while at the same time offering a reasonably large theoretical password space compared to, for example, lock patterns [16].

As of today, most previous works on image-based authentication schemes were conducted in the lab or through field studies albeit with limited ecological validity. This causes a considerable knowledge gap with respect to the properties of image-based passwords during real-world use. Most notably, the underlying image can be expected to have a strong influence on how users select their passwords. The number of features as well as the content itself may lead to so-called hotspots, i.e., locations within the image that are more frequently selected than others [19]. These hotspots reduce the theoretical password space and thereby provide valuable hints for attackers as to what the password of the user is.

To bridge this gap, we present the first study on the use of image-based passwords in the wild. In October 2013 we released an image-based authentication application in Google Play [16]. Within 12 months we gathered data from 2318 unique devices on which the application has been installed. The data allows us to answer the following questions:
We found that a significant number of people (41%) prefer to use custom images and that the vast majority of passwords consist of 3 or 4 password points (68.4%). In contrast to other graphical password schemes, such as lock patterns, the distribution of password points is strongly influenced by the underlying image. With regard to security, our results show that 31% of the password points fall into salient regions of the image and could thus be predicted by computational models of visual attention. Human attackers can predict hotspot areas even more accurately: in a small study, participants were able to predict 51% of the password points. Our findings suggest directions for future work and can help to inform the design of future image-based authentication mechanisms.

RELATED WORK
Our work builds on previous studies on image-based graphical passwords and in-the-wild studies on usable security.

Image-based Passwords for Authentication
This work investigates image-based passwords that, like lock patterns, belong to the class of graphical passwords. Graphical authentication schemes have long been subject of investigation in usable security research [2, 17] due to their ability to leverage the large capacity and capabilities of the human visual system [4, 5]. Prior research showed graphical passwords to increase security without compromising usability [13, 15].

Of particular interest to our work are locimetric password schemes in which users are presented a single image. Users define passwords by selecting points of the image sequentially. An example is Passpoints, where users can select an arbitrary image and then define a password through click points within the image [22]. More recently, Microsoft presented an authentication scheme in Windows 8 where combinations of shapes need to be drawn onto a picture to create a password.

These approaches allow users to choose both the image and the password locations. This is problematic since users tend to select similar password points, the so-called hotspots [19]. One solution to this problem was presented by Bulling et al. who proposed to identify and mask potential hotspots automatically using a computational model of human visual attention, thus increasing security [3]. Another problem on touch-enabled surface are smudge traces from fingers that can be used to infer the underlying password. Schneegass et al. proposed to transform the underlying image each time the user authenticates to create different smudge traces [16].

In summary, the review of related work shows that this authentication scheme is well established and the potential being recognized by the research community. At the same time, research so far was conducted almost exclusively in the lab. As a result, little is known about which images users select and how they define passwords during daily use. Answering these questions is the focus of the current work.

Related Work

Most closely related to our research is, again, the work by Schneegass et al. who released an image-based authentication app in a mobile app store [16]. However, this work focused on usability, investigating how successful users are when performing login attempts with regard to errors and login time. No information is provided on image and password selection.

Further examples involving large number of users in the field include the work by von Zezschwitz et al. who recruited 298 participants to investigate how easily observable grid-based unlock patterns are [20]. Further work evaluated the usability of pattern and PIN-based authentication mechanisms over the duration of 21 days [21]. Despite being conducted outside the lab, both studies are limited with regard to ecologic validity. The former study was conducted on the web while the latter one did not embed the login procedure in the users’ daily authentication routine. However, we deemed this crucial for our research to obtain ecologically valid results.

USE OF IMAGE-BASED PASSWORDS IN THE WILD
Compared to PINs or lock pattern, only little is known about users' choice of images and passwords in image-based schemes. Knowledge about user behavior can inform the design of future authentication mechanisms and increase security and usability of such schemes. Our research is guided by three main objectives: 1) reveal what kind of images users select, 2) understand how users define passwords on these images, 3) assess the security of these passwords.

Image Selection
Prior work suggests that the underlying image has a strong influence on how users select passwords [19, 22]. Hence, at the outset of creating and implementing an image-based authentication mechanism, the fundamental question of which images are to be used needs to be answered. The first objective of this work is to better understand how images influence password selection to inform decisions whether or not to allow users to select their own images. There is an obvious trade-off: allowing users to select their own image may lead to higher memorability and thus better user experience. Given that there is still a significant number of users not protecting their phone at all [18], this could be a crucial motivating factor for protecting the phone in the first place. However, user-chosen images may lead to lower security, for example if images are too simple and hence allow (parts of) the password to be predicted.

Password Selection
Today, no ecologically valid data exists how users choose passwords in image-based authentication schemes. Similar to other schemes, there is a trade-off between memorability and security, i.e. generally, the more points a password consists
of, the more secure it is. In addition, user strategies may compromise security. For example, prior work found that for lock patterns, about 40% of the users select the top left point in the grid as their first password point [1]. Revealing similar strategies for image-based passwords could help to develop policies with the goal to motivate or even force users to select more secure passwords. Our second objective is to investigate the password length, in which areas they are defined, and how password points are spatially chosen relative to each other.

**Security Assessment**

Our third objective is to understand how secure user chosen passwords are. We evaluate different methods for “predicting” vulnerable areas, i.e. areas that are likely to become hotspots. We study computational methods to find hotspots in an automated way and we compare our data to findings from a small user study where human attackers proposed hotspots.

**Summary**

To answer the aforementioned questions, we opted to design an authentication mechanism with as few constraints as possible. Firstly, users could freely choose the image to use. Still, we delivered a set of (default) images with the app for users to chose from. This allows us to obtain knowledge about which images users select. At the same time, providing pre-defined images led to that multiple users chose the same images. This allows us selection strategies to be identified and hotspots to be revealed. Secondly, we did not constrain users when choosing the passwords. Passwords could have arbitrary length and be defined on arbitrary spots. Thus, we aimed to understand how users chose passwords easy enough to remember but, at the same time, considered to still be sufficiently secure.

**DATA COLLECTION**

To increase ecological validity, we implemented an Android application that fully replaces the user’s login screen. We use Android’s device policy manager to set a password and lock the phone. This way, our application can be used in the same way as, for example, the Android lock pattern. We provide a wizard to inform users about our research and to help them setup their image-based password. On first launch, users are asked to participate in our research by allowing us to analyze any data generated while using the application in an anonymized way. Note, that users who opt out can still use the full functionality of the app. In addition they can opt in and opt out at any time by changing the respective setting.

The user can then either select one of the 18 default images delivered with the application or select an image already on the smartphone. Afterwards, users are prompted to define a password on the selected image by swiping over their preferred password points (we refer to the connection between two password points as a stroke). Then, the user is asked to confirm the password by entering it once again. In addition, users need to define a PIN as a fallback in case they fail to correctly remember their password. Pressing the home button forces the system to show the Android PIN input mask.

More information on the design and implementation of the SmudgeSafe application used for this research can be found in the work of Schneegass et al. [16].

---

**Figure 2.** Default images contained sufficient opportunities to easily select password points and objects were well distributed.

**Image Selection**

For the default images we chose a variety of creative commons images from flickr.com¹, depicting landscapes, animals, and artwork. We made sure that images contained a large number of features, i.e., a number of objects significantly larger in comparison to the expected number of password points. Furthermore, features were equally distributed (see Figure 2).

**Data Logging**

The application anonymously logs password points and login attempts. In addition, users can choose to share their own (custom) images. Collected data is encrypted and transferred to our server each time users connect to WiFi to not infer any connection costs.

**Distribution**

After testing the application with colleagues and ourselves for two weeks, we uploaded it to the Google Play Store in October 2013. We initially advertised the application to colleagues and friends via Facebook, through bulletins at University and through word-of-mouth. Some weeks into deployment, a post on the application appeared in a popular online blog, leading to a significant increase in the number of users. During the past year we maintained the application, in particular making sure that new Android OS releases did not compromise the functionality of the application. We released several updates, none of which changed the main functionality. The application can be downloaded from the Google Play store².

**Dataset**

In the following analysis we consider data gathered over 12 months from November 2013 until October 2014. During this time, 2318 unique devices contributed to the dataset. For later analysis we excluded all devices that did not perform at least one successful login event. Overall, users defined 1793 passwords, 697 on predefined images and 1096 on custom images. Users shared 146 of their custom images. Overall, active users performed on average 47.5 logins per day (Avg. per hour = 2.97; Max = 24.72; SD = 3.48; assuming users are awake for 16 hours a day). This is in line with findings from Harbach et al. who reported that people using lock patterns log in 3 times per hour on average [6]. According to information from Google Play, users mainly came from Brazil, Spain, France, and Russia.

¹Flickr website: [http://www.flickr.com](http://www.flickr.com)

We first analyzed which images users selected by grouping all images into categories. Table 1 provides an overview of the categories, number of images, and number of passwords.

### People
This category includes images of celebrities (mainly singers), babies, children, and (groups of) adults, presumably the users themselves as well as partners, friends, or relatives.

### Comic & Animation
The second largest category consists of images depicting scenes / characters from comics or animation films (Snoopy, Mickey Mouse, Spongebob, etc.).

### Scenery
Many users selected images of places they presumably visited and/or liked. Particularly popular were skylines, beaches, and famous sites (e.g., the Eiffel tower).

### Phone Screenshots
Some used a screenshot of their former login or main screen. We assume users did this to deceive attackers or as app icons seemed good for selecting a password.

### Art
This category includes images showing pieces of art as well as poems or song texts.

### Cars
Nine images depict cars or motorcycles. The images included both professional pictures as well as pictures of (presumably) vehicles of the user.

### Brands
Some users selected brand logos (such as the NIKE logo) or high-quality advertising images.

### Animals
Animal images were not particularly popular. Only four custom images showed animals. Three were cats laying in their beds, one depicted a duck posing in the water.

The mean number of passwords defined on each default image was 39 (SD = 46.73). Overall, users mostly preferred the Scenery category (512 defined passwords = 57 per image) followed by Animals (103 defined passwords = 27 per image) and Art (82 defined passwords = 17 per image).

In addition we made the following observations.

### Leveraging Contextual Information
We found the idea of using a screenshot as password image interesting. Though we did not evaluate this further, this seems a clever way of creating an easy-to-remember, yet secure password. A user could, for example, select the most frequently used apps as a password. This would allow the password to be remembered more easily while at the same time not providing any hint to the attacker, unless he has knowledge about this strategy and the habits of the owner. Additionally, attackers could be confused by the effect that the smartphone seems to not be protected but at the same time not be reacting to the attacker’s input.

### Customized Images
In some cases images were created particularly for the application. For example, one user created a collage consisting of eight images that showed a boy, a girl, and two sports stars. Over the 12 months he updated the collage several times by replacing some of the photos. Such collages could be used in a similar manner as above.

### Advertising Opportunities
The fact that some user chose brand logos as their background creates opportunities for advertisers. We assume that these images represent (one of the user’s) favorite brand(s).

### Password Forms
On some occasions, users selected forms as passwords. For example one user defined her password in the form of a heart and then fitted it into the image. We assume that this made remembering the password for her even easier.

### Password Selection

#### Distribution of Password Points
To analyze the distribution of password points, we divided the screen into a 3x3 grid. Data from all default passwords showed that users tend to use the right part of the image (46%) more often than the left and middle part (Figure 4c) In contrast, there is no clear tendency for the vertical distribution. Each part of the screen contains between 31% (top) and 36% (middle) of the password points. Comparing the location of the first password point to the other points (Figure 4a) shows that users tend to start in the top right corner. This is in contrast to lock patterns where more than 40% of the passwords start in the top-left corner [1]. In 38% of the cases the end point lies in the middle or lower right area of the image (Figure 4b). Though we do not know how many users were right-handed, this skew might be explained anatomically, i.e. it is more comfortable to end password in the lower right corner.

We also looked at the distribution for custom images. The location of the first password point is rather on the top left (Figure 4d-f). Password points are more centered around the middle of the screen. A closer examination of the custom images reveals, that in many cases, the subject of the photo is at the center. This suggests that the composition of an image can influence anatomic considerations.
Figure 4. Distribution of Password Points: The distribution for default images (left) and custom images (right) differed, presumably as a result of different positions of ‘interesting’ objects within the image. This suggests objects of interest to be more important than ergonomic factors.

Figure 5. Directions of strokes performed by the users. Each direction includes all strokes that are performed within ±22.5 degrees. The diagram shows a tendency to draw password from downwards and to the right.

Length of Passwords
Second, we identified two metrics for the length of the password. First, we calculated the number of password strokes (i.e., distance between two consecutive password points) of each password (cf., Figure 3 – left). Most passwords consist of two or three password strokes ($Med = 2$). Second, we calculated the length of each stroke in pixels. Thereby, we differentiated between passwords with different numbers of strokes (1–5 strokes). We found that the average length of the strokes decreases with an increasing number of strokes (ranging from 380 to 218 pixels). We also found that independent of the number of password strokes, the stroke length decreases from 360 pixels (1st stroke) to 247 pixels (5th stroke). This is interesting because it suggests that with increasing password length users tend to look for features that are closer together. A reason may also be that users try to avoid overlapping strokes.

Direction of Password Strokes
Third, we evaluated the stroke direction by defining eight directions (each including strokes within an angle of 45 degrees, Figure 5). For default passwords, we found that most first strokes were performed downwards, the second ones downwards or right. Overall, more strokes were performed downward (48.3%) than upward (30.6%). Interestingly, the percentage of strokes to the right (38.8%) is higher than to the left (30.4%), given that top-right was the most popular first point.

For custom passwords the majority of first strokes was performed to the right and to the bottom. This is likely to be a result of initial password point lying at the top-right of the screen. The trend is less pronounced for subsequent strokes. Again, the percentage of downward strokes is higher (44.8%) compared to upward (28.7%) and also rightward strokes occur more frequently (42.8%) than leftward strokes (28.1%).

Password Security
Finally, we were interested in how secure the chosen passwords were. We used two methods to “predict” from which image area password points are most likely selected: a computational method and a human attacker based approach.

Computational Approach
For the computational approach, we created the saliency maps for both the default and custom images. To do so, we used a Graph-Based Visual Saliency (GBVS) model (see [8] for details on GBVS and [7] for the MATLAB toolbox we used). GBVS was shown to predict human fixations on natural images with superior performance to the original visual saliency algorithm [11]. The saliency maps were calculated using the toolbox’s default parameters. The greyscale heatmaps returned by the GBVS algorithm were normalized and a threshold applied at the 0.5 level so as to separate salient and non-salient areas. The salient areas were used as saliency masks that were overlaid onto the original images.

We then looked into (a) how many password points fall into this region and (b) how large this region is. The results for default images show that 34.2% of the password points are located within the salient region ($SD = 20.0\%, Min = 1.6\%, Max = 71.7\%$). This is in line with findings from Bulling et al. who found 41 out of 119 images (34.4%) to be defined within the salient regions. At the same time, saliency maps cover on average 21.8% of the image ($SD = 9.6\%, Min = 0.9\%, Max = 33.8\%$). The results for custom images reveal that 21.0% of the password points are located within the salient region. Saliency maps cover on average 9.15% of the image.

Human Attacker
We invited three people to take part in a small user trial (all male, aged 25, 27, and 29). We opted for a small number of participants to show that there is no huge effort needed to
While the analysis in this work looked at the case where an approach, such as lock patterns, struggle with pre-defined passwords. We found that human attackers performed about 50% better with locations identified by the saliency model. We also qualitatively compared the user-selected locations with locations identified by the saliency model. We found that users seemed to define password points on aesthetically pleasing locations rather than on high-contrast ones. An example is the image with ducks (Figure 2): While the saliency algorithm identifies the feet as salient regions the majority of users selected the low-contrast heads. Users also tended more towards a particular object – even if, from a computational perspective, it was not as salient as, for example, its shadow. For the cups image (Figure 2) this means that users tended to select the bottom of the cups rather than more salient parts.

Conclusion
We found that human attackers performed about 50% better than saliency masks. This calls for research that aims to create more accurate models. One opportunity would be to create a model based on human behavior. Subsequently, images could be automatically checked for security and – similar to [3] – masked on vulnerable areas from password selection.

DISCUSSION AND FUTURE WORK
While approaches, such as lock patterns, struggle with predictable passwords (many users selecting the top-left corner as starting point, drawing similar shapes), the smart choice of password images could help to overcome this. While, overall, the direction of the password still seems to follow the reading direction (right and down) as well as to consider ergonomic constraints (e.g., the top left corner difficult to reach; end of password more comfortable at the bottom-right), our results suggest that sufficiently “interesting” features in the image can overcome this. Future work could, for example, investigate the users’ motivation to select a particular spot. Furthermore, the connection between points of interest and the selected password points (most importantly the starting point) could be investigated.

While the analysis in this work looked at the case where an attacker has no knowledge about the user, it could be interesting to put the user more into perspective. In contrast to approaches like PINs or lock patterns, image-based passwords introduce a personal component, i.e. the selected image and its content. Passwords may be strongly influenced by personal preferences or by the relationship of the user to an object (for example, the partner in a group of people). In such cases, an attack may be easier for people who know the user. Strategies to cope with this could be interesting research directions.

Finally, interviews with users could focus on the acceptance of the approach. In particular, it would be interesting to find out in which cases users decided to stop using image-based passwords and revert back to their previously used authentication mechanism. Note, that assessing the usability of the approach was beyond the scope of this study and has been reported in detail in prior work [16].

LIMITATIONS
While our findings provide valuable insights our study also has limitations. First, our results are limited by the default images we used. While we offered images that provide opportunities for selecting passwords in any area of the image, still some areas may have been more attractive to users than others and hence have impacted on the selection of the password points. However, to the best of our knowledge no commonly agreed upon procedure exists for predicting such attractive areas, which would have allowed us to better control for this. As has been shown above, saliency masks provide a good, though not perfect estimate. We furthermore acknowledge that since we did not counter-balance the presentation of default images, this may have influenced how often particular images were selected.

A second limitation is that we do not have information about which and how many users were right or left-handed. While the distribution of password points (Figure 4) suggests that the reason for the skew in the distribution might be anatomically, we cannot entirely support this suggestion. Future work should look in more detail at the influence of the dominant hand on where users select image-based passwords.

Third, we acknowledge the general limitations of an in-the-wild study, foremost the lack of internal validity. We did not have control over the situations in which users entered their passwords (while walking, driving, biking) and how they chose it. Furthermore, users may have tested the system only a few days, before reverting back to their previously used authentication mechanism.

CONCLUSION
This paper explores the use of image-based passwords in the wild. We released an image-based password app in the Google Play store and collected data from 2318 unique devices over one year. Through investigating aspects, such as the choice of images and passwords, it became apparent, that findings from prior work on the security of other schemes (PINs, lock patterns, etc.) do not easily transfer to image-based passwords. Our initial assessment of security shows that there is a need for further research, to make image-based passwords more secure. Our work reveals weaknesses but also opportunities offered by such authentication schemes.

ACKNOWLEDGEMENTS
We thank Frank Steinmle for his help with maintaining the SmudgeSafe application for the duration of the study. The research leading to these results has partly received funding from the German Research Foundation within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart.
REFERENCES


18. Tapellini, D. Smart phone thefts rose to 3.1 million last year. ConsumerReports.org, May 2014.


