Proceedings of the 2013 Federated Conference on
Computer Science and Information Systems pp. 837–844

Are Graphical Authentication Mechanisms As Strong As Passwords?

Karen Renaud*, Peter Mayer†, Melanie Volkamer† and Joseph Maguire*
*School of Computing Science, University of Glasgow
†Center for Advanced Security Research Darmstadt, Technische Universität Darmstadt
E-mail: karen.renaud@glasgow.ac.uk, {peter.mayer, melanie.volkamer}@cased.de

***Abstract*—The fact that users struggle to keep up with all their (textual) passwords is no secret. Thus, one could argue that the textual password needs to be replaced. One alternative is graphical authentication. A wide range of graphical mechanisms have been proposed in the research literature. Yet, the industry has not embraced these alternatives. We use nowadays (textual) passwords several times a day to mediate access to protected resources and to ensure that accountability is facilitated. Consequently, the main aspect of interest to decision-makers is the strength of an authentication mechanism to resist intrusion attempts. Yet, researchers proposing alternative mechanisms have primarily focused on the users' need for superior usability while the strength of the mechanisms often remains unknown to the decision makers. In this paper we describe a range of graphical authentication mechanisms and consider how much strength they exhibit, in comparison to the textual password. As basic criteria for this comparison, we use the standard guessability, observability and recordability metrics proposed by De Angeli *et al.* in 2005. The intention of this paper is to provide a better understanding of the potential for graphical mechanisms to be equal to, or superior to, the password in terms of meeting its most basic requirement namely resisting intrusion attempts.**

I. Introduction

ONE OF the most basic, everyday tasks of computer usage is authentication. Every user will, sooner or later, have to authenticate themselves. Their ability to do this effectively will impact on their ability to do their daily jobs and on their personal lives. The failure of the mechanism to resist intrusions will potentially have an impact on the user personally (e.g., in terms of ID theft or financial losses) or in professional environments on the organisation he or she works for.

Nowadays, the most widely used authentication mechanism is the textual password. However, it is well known, that most users are frustrated by their experiences with these traditional passwords in general [1]. Even if they want to behave securely, they often do not understand what constitutes a "secure" password since guidelines for the creation of secure passwords are seldom adequate [2]. Even with good guidelines in place, human nature will lead users to prefer the path of least resistance e.g. choosing weak passwords, writing them down, storing them in plain text on their mobile phones or reusing them [3], [4]. This is understandable considering the findings of Ives, Walsh and Schneider [5]: users are expected to recall an average of 15 different passwords on a daily basis. Due to human cognitive limitations, four or five is normally the maximum a typical user can handle [3].

Password managers can help users to manage an unlimited number of passwords. However, they constitute a single point of failure and systems cannot be easily accessed from a device that does not have the manager installed. Thus, password managers are no substitution for a secure and usable authentication solution [6]. The same holds for single sign on solutions.

To address the human inability to deal with large amounts of passwords, a new type of authentication system was conceived. The graphical password, first proposed by Blonder [7], required the person to verify their identity by clicking on positions within a picture. This is called a *locimetric* system. Other common types are *searchmetric* (pick a picture from a grid of images) and *drawmetric* (draw your secret) [8]. The most important motivator behind the use of a graphical authentication mechanism is that their memorability is superior to that of textual passwords. In the first place, there is what is called a "picture superiority effect", as described by Paivio [9]. Paivio explained that pictures were encoded using a dual mechanism. So, a password, being textual has only one route whereby the human can reach it. If that route decays, and is forgotten, the password cannot be accessed. If the memory item is visual, there will potentially be multiple routes to access it, and the decay of one access route does not render the item unreachable.

Numerous studies regarding the usability of graphical authentication schemes have been conducted. Yet, many of these sweep security concerns aside or deal with them in a desultory fashion [10]. However, very few graphical mechanisms are used in practice. Notable exceptions are the Windows 8 picture password and the Android lock-screen pattern. A number of reasons could be advanced. In this paper we will consider the elephant in the room: do these mechanisms provide the basic requirement namely resisting intrusion attempts at an equal or higher level than textual passwords do? In order to answer this question, we need to be clear about exactly what variation of the amorphous password we consider, since this impacts the resulting security-level. Hence, we consider textual passwords with a length of at least 8 characters and which are used in a system with a three-times-lockout technique.

We will use the different categories of attacks proposed by De Angeli, Coventry, Johnson and Renaud [11] in 2005 as a starting point. In each case the mechanism will be compared to textual passwords. The remainder of this paper is structured as follows. First, we present the three types of attacks we

978-1-4673-4471-5/$25.00

use to evaluate whether the same security level as textual passwords are provided, and show how the most common attacks fit into this framework. Next, we describe the general ideas behind the different classes of graphical authentication mechanisms. Then, we report on efforts that have been made to strengthen these mechanisms. We then compare the security properties of graphical authentication schemes to the security offered by textual passwords. Finally, we summarise, discuss and conclude.

II. Vulnerabilities

To evaluate the security of different authentication mechanisms, the resistance against certain attacks is reviewed. Table I classifies the set of attacks proposed in [12] into the evaluation categories which we use throughout this paper, namely Guessability, Observability, Recordability, and Memorability; which are explained in the following paragraphs.

TABLE I
Common Attack Types

Vulnerability	Attack Type	Tool
Guessability	Brute Force Attack	Offline
	Dictionary Attack	On & Offline
Observability	Shoulder Surfing	Human Observer
	Spyware	Technology
Recordability	Social Engineering	Deception
	Theft	Unsecured Record
Memorability	Forgetting	Coping Techniques

Guessability: Brute force and dictionary attacks are the two types that have to be considered in this category. For dictionary attacks to be possible, passwords have to be predictable so that an attacker can create a dictionary with the most widely used passwords. Obviously having personal knowledge about the user can make it even more predictable. These attacks can in particular be carried out if the database or password file is obtained (without in our setting the account will be blocked after three trials). In this case the attacker can perform an offline attack to test all possible combinations of the password-space and thus has no limitations regarding aspects like a lockout policy etc. One way to resist this kind of attack is to impose password strength requirements when a password is chosen. Another way is for the system to issue strong passwords. This ensures that passwords are unpredictable and that all possible elements of the password space are evenly likely to occur. To be resistant against offline attacks, the password space has to be sufficiently large, where values of 2^{20} to 2^{28} seem to be commonly used on the Internet [13]. To be secure, the password space should be greater than 2^{80}, which is the lowest security strength NIST allows for government applications [14].

Observability: Observing the user while authenticating can be performed either by a human (shoulder surfing), by a human with technical equipment (filming the action of authentication) or using technical means (e.g. spyware). The goal is to collect information that allows an attacker to reproduce the authentication with an as high as possible probability. Shoulder surfing is performed by observing the user while he is authenticating himself with the given implementation. Lately, the ubiquity of mobile phones means that a camera could be used to capture the user's authentication secret without their knowledge. Spyware or malware in general can be installed on the user's system by the attacker. It can monitor input peripherals or obtain screenshots during authentication.

Resisting this kind of attack is challenging. One can resist this to a certain extent by making it necessary for an attacker to capture multiple authentication attempts before they obtain the full authentication secret [15].

Recordability: Attacks exploiting the recordability of an authentication mechanism are always performed through the human factor. The first way a password can leak to an adversary is that the user records their password in some way and it is then stolen (theft). The second way is for the person to be fooled into disclosing their secret to another person (social engineering). Social engineering includes all attacks, which do not target the system as such, but the user.

Both attacks rely on the relative ease with which a user can record his/her authentication secret. An implementation is resistant if it is hard for the secret to be recorded or disclosed. In the era of ubiquitous mobile phones with built-in cameras it is very difficult to resist this kind of vulnerability.

Memorability: While this is not an attack type, it can be a vulnerability if attackers exploit the consequences of a user's coping strategies [3]; e.g. writing down has an impact on observability as someone might not observe the user authenticating but the note with the information about the password. Similar re-usage has an impact on the dictionary attack. Note, with respect to memorability it is important to consider also situations where one might only authenticate once in a while, like once in a semester for students to register or once in a year to file a tax return.

III. Graphical Authentication

In general, graphical authentication works like any other knowledge-based authentication mechanism. The user has to verify knowledge of a secret he or she shares with the system. Contrary to the abstract nature of textual passwords, graphical authentication relies on visual memory. In both cases he or she has to access that secret in stored memory. Memory can be accessed in three ways, as described below, using the password as an example in each case to explain approaches for each way.

Recall: Information is extracted from memory when requested. This is the paradigm adopted by the traditional textual password authentication. Recall is a cognitively difficult task. Therefore, users tend to resort to coping strategies. Graphical passwords that rely on this kind of memory are the drawmetric based schemes like Android screen unlock and the searchmetric grIDsure [16].

Cued-Recall: Information is extracted from memory when cued. One can also ask for a password framed as a response task, similar to Zviran & Haga's associative passwords [17]. In their scheme users provide a number of associations

at enrolment which they are prompted for at authentication. Most of the graphical authentication schemes relying on this memory are locimetric, but exceptions such as the drawmetric BDAS exist [18].

Recognition: Information is presented and the individual is able to identify the correct item. One could conceivably display a number of passwords on the screen and ask a user to identify theirs, but this scheme is so obviously weak that it has not been trialled. Many graphical authentication mechanisms do rely on this mechanism, since it is the least cognitively demanding and particularly suitable for use with images. These are the searchmetric mechanisms. They display a succession of challenge sets, with one "target" image and a number of distractor images. The user identifies the target image in each challenge set by clicking on it.

The following sections provide examples of graphical authentication mechanisms that rely on each of these memory types. The sections are intended to be illustrative and examples of mechanisms have been chosen because they were the first of their kind. The inclusion of a particular mechanism does not suggest that it is in any way superior to others which are not mentioned. There is, unfortunately, no room for an exhaustive review of all mechanisms.

A. Recall

Draw-a-secret or DAS, proposed by [19], is a recall-based graphical authentication mechanism. The approach expects an individual to draw their authentication secret to access an application. These should be more memorable than passwords because they rely on visual, lexical and kinaesthetic memory [8]. DAS does not rely on drawings from a semantic perspective but on the underlying grid sectors.

Thorpe and van Oorschott [20] postulated that *symmetrical graphical secrets* are a real concern, as symmetrical drawings have superior recall. They argue that an attacker could craft a dictionary of symmetrical secrets and use it to compromise DAS which would take only 6 days to crack if the password is symmetrical. Nali and Thorpe conducted an informal user-study with 16 individuals to determine if user-generated DAS secrets exhibited any patterns [21]. Their participants' drawings did exhibit patterns: 45% created symmetrical drawings, 56% of the drawings were centred and 80% of the drawings used fewer than 3 strokes. The authors argue if these results are symptomatic of a larger user-base, then DAS has a much smaller, practical password-space.

The latest drawmetric mechanism is the common Android lock-screen pattern authentication. Such pattern based authentication mechanisms are also vulnerable to attacks based on observations of smudges on the device touchscreen [22].

Another recall-based approach is grIDsure, an authentication scheme that relies on knowledge of a secret pattern [16]. It requires an individual to create a sequenced pattern on a 5x5 grid. The user is presented the same 5x5 grid during authentication, except each one of the 25 cells contains a random value between 0 and 9. The values are randomly generated for each authentication attempt and are not unique to a cell. The secret pattern, generated by the user, is applied to the grid to generate the authentication secret, i.e. a 4-digit PIN. Bond identified some severe security flaws in this scheme [23]. He was able to identify the user's secret using only two forged authentication grids.

B. Cued-Recall

There is some concern that users forget their drawings with recall-based mechanisms such as DAS, or at least the stroke order [24]. To address this, Dunphy and Yan proposed a grid superimposed over a background image Background DAS (BDAS) to act as a cueing mechanism to improve memorability. Unfortunately users still created weak passwords [18].

Building on the ideas of Blonder, Wiedenbeck [25] proposed a mechanism called PassPoints. In PassPoints the user is expected to select five click points on an image. The sequence of click points is the authentication secret. Each position has a small tolerance radius as perfect replication is not expected. The password-space of PassPoints is vast, even with the addition of the tolerance radius, as a single image can contain a large amount of possible click points. The image can be selected from a library or provided by the user, the only requirement being that the image is complex enough to inspire users and protect the secret.

The apparent strength of the PassPoints approach is the large theoretical password-space afforded by the pixel-rich images. Thorpe and Van Oorshot argue the practical password space of PassPoints is reduced because of 'hot-spots', i.e. popular click points, as well as patterns within secret generation [26]. They investigated both human-based attacks and automated attacks. They investigated two highly-detailed images for popular positions. They discovered that 5 points in both images proved popular with individuals, between 24-31% for the first image and 20-24% for the second. Similarly, Dirik, Memon and Birget [27] developed a model that they claim can identify popular regions for points. They cautiously report that they were able to extract 70 - 80% of points. Furthermore, Thorpe and Van Oorshot [26] also suggest that predictable patterns exist in sequence selection.

C. Recognition

Dhamija and Perrig [1] propose Déjà Vu, a recognition-based graphical authentication mechanism. Each image is abstract in nature and the collection is generated using a mathematical formula, the output depends on an initial seed. The beauty of this design is that the actual images do not need to be stored, just the small initial seed. Déjà Vu performed well against competing recall-based approaches such as passwords and PINs. Indeed, Dhamija and Perrig reported that more individuals were unable to recall their username than were unable to recognise the images within their secret sets. Individuals using Déjà Vu felt that it was overall easier to use but at the expense of time and security. However, they reported an interesting insight in regards to the image-type used in Déjà Vu. When using semantic images, i.e. photographic scenes, some individuals selected the same images. One specific image

was selected by 9 out of 20 individuals. Furthermore, these images are far easier to explain and describe, thus, as a consequence an individual's secret set of images is easier to convey to someone else. For example, the aforementioned popular semantic image contained the Golden Gate bridge. Conversely, abstract images rarely overlapped and descriptions of them rarely, if ever, matched. In theory this strengthened the practical password space of the approach as there was no real pattern or popular images. Naturally, further investigation will be required but Dhamija and Perrig highlight the impact the image type can have on the memorability of images.

PassFaces is one of the few commercial graphical authentication mechanisms. The authentication approach assigns an individual a collection of faces as their authentication secret. The user is then presented a series of challenge stages that are each comprised of a nine image grid. In each grid the user has to identify one image from his or her password (*target*) among eight *distractors*. One property which might severely impact guessability (offline attacks) of searchmetric schemes, is that they usually need to store some password information in the clear [28].

IV. Strengthening The Graphical Password

A couple of improvements and also combinations of different approaches have been proposed in order to overcome weaknesses with respect to one or several of the vulnerabilities mentioned in Section II. These are proposed and discussed in this section.

A. Guessability Resistance

Cued-Recall: PassPoints authentication exhibits popular positions or hot-spots which are problematical in terms of guessability. Chiasson, van Oorschot and Biddle [29] propose Cued Click Points or CCP. The approach is a variation on PassPoints, in the sense that an individual selects a position from an image. However, the main difference is that an individual is required to repeat this action over several images. Therefore, the secret is a sequence of click points selected from a series of images with one click point on each image. The images are intended as cues.

There are concerns about the predictability of CCP, primarily those inherited from PassPoints, such as popular click points. Chiasson, Forget, Biddle and van Oorschot tackle this specific problem with Persuasive Cued Click Points or PCCP [30]. This approach uses CCP with the addition of a *persuasive viewport*. During the registration phase a viewport is randomly positioned over the image. The viewport is emphasised by reducing the brightness of the rest of the image. The individual is only allowed to select a click point from within the viewport, which they can shuffle if they do not like the position. Their lab-based experiments showed that the viewport is successful in reducing hot-spots and increases the spread of click points and their web-based trials were equally positive [30]. However, a more longitudinal evaluation is required to confirm whether the findings are replicated in the wild.

Recognition: Graphical authentication relying on recognition requires people to use images. There are three ways of associating people with images: (1) let them supply the images themselves, or (2) allow them to choose from a range of images or (3) assign images randomly to them. Unfortunately, the first two options can cause severe security concerns. If users are allowed to supply their own images they tend to choose predictable images [31]. The same holds, if users select their images from a set of supplied ones. As long as users have a choice they will behave predictably. Davis, Monrose and Reiter [32] discovered that individuals make predictable choices when they are required to select images for use in graphical authentication utilising facial images. Individuals are influenced by attraction, race and familiarity [33]. Even with everyday representational images, humans tend to make predictable choices [34]. To minimise guessability, images should be assigned to users randomly. In this case the theoretical and the practical password-space are the same.

B. Observation-Resistance

The first reaction to graphical authentication is often that it will be too easy for a human observer to gain knowledge of the authentication secret. Thus, a variety of attempts have been made to make it more difficult for observers to do this. Essentially, variants focussing on the observation resistance have been proposed for all three approaches to graphical authentication.

Recall: A DAS secret is easily exposed to onlookers. If an attacker is able to observe entry of a DAS secret, he or she may be able to authenticate using the same drawing (this holds for the cued-recall BDAS as well). Lin, Dunphy, Olivier and Yan proposed Qualitative DAS (QDAS) to tackle the problem of observation [35]. Chakrabarti, Landon and Singhal argued that rotating the canvas which the user draws on could improve the resiliency of DAS to observation [36]. Yet, neither of these has been tested in the wild so only lab-based results about the impact of the scheme on ease of observation have been reported.

Cued-Recall: While originally intended to be used with recognition-based authentication systems, the approach of Dunphy, Heiner and Asokan proposed in [15] can easily be translated to the CCP and the PCCP scheme. They propose to use a portfolio of target images of which only a different random subset is needed for the authentication process each time. Likewise CCP and PCCP could be implemented using a click point portfolio with only a random subset of the images (and the according click point) being needed to authenticate. To the authors' knowledge this has not yet been attempted, but might prove an interesting subject for future research.

Recognition: A range of observation resilient systems have been proposed. Dunphy, Heiner and Asokan [15] tested redundancy with users of a searchmetric system with 8 distractors and one target image. Each user had a portfolio of 6 images, of which only 4 were used at each authentication attempt. Attackers needed 7.5 observations, on average, in order to be able to reconstruct the password.

TABLE II
COMPARISON OF GRAPHICAL PASSWORD SCHEMES TO TEXTUAL PASSWORDS. G=GUESSABILITY, O=OBSERVABILITY, R=RECORDABILITY, M=MEMORABILITY, ?=NO DATA AVAILABLE.

	Scheme	Guessability	Observability	Recordability	Memorability
Recall	QDAS [35]	G ?	O ?	R ?	M ?
	BDAS [18]	$G\downarrow$	O ?	$R=$	M ?
Cued-recall	PassPoints [37]	$G=$ [26], [38]	O ?	$R=$ [39]	$M\uparrow$ [40], [6], [41]
	PCCP [30]	$G\uparrow$ [30]	O ?	$R=$ [39]	$M=$ [30]
Recognition	Passfaces [42]	$G\downarrow$	$O\uparrow$ [43], [15]	$R=$ [39]	$M\uparrow$ [44]
	Passfaces (system issued passwords)	$G\uparrow$	$O\uparrow$ [43], [15]	$R=$ [39]	M ?
	Dynahand [45]	$G\downarrow$ [45]	$O=$ [45]	R ?	$M=$

Wiedenbeck [46] proposes an authentication approach, framed as a game with a cognitive trapdoor, that relies on users generating a convex hull. The cognitive trapdoor is knowledge of 5 icons or pass-icons. The game comprises of a series of challenges, in each challenge the user needs to click within a specific region to 'win'. The specific region is revealed by uncovering a convex hull. During authentication, the user is presented a canvas containing several icons, including *at least* 3 pass-icons. The user is required to envisage the convex hull spanned by the pass-icons and needs to click within the convex hull to complete the challenge. Regardless of whether users click within the correct region or not, they progress to the next challenge when they click on the canvas. Wiedenbeck states that authentication times for the proposed Convex Hull Click approach are lengthy, they are the necessary expense of increased resilience to observation. Moreover, the author argues that energy has been spent to delight and excite the user to maintain interest with the approach.

Searchmetric mechanisms can be morphed into *limited disclosure* searchmetric mechanisms to foil shoulder surfing and key-logging software, since they rely on the use of arrow keys or a mouse to manipulate sets of pictures and the user does not click on their actual image. Most limited disclosure searchmetric mechanisms have some redundancy built in so that the observer is not able to deduce the key from casual observation but has either to observe a number of authentications or carry out an error-prone deduction of the key based on a few observations.

Tetrad [47] was proposed by Renaud and Maguire. Tetrad displays a grid (9x5) of facial images. Users line up their secret images by manipulating rows and columns instead of clicking on the images themselves. This introduces a level of indirection which means that casual observation is less profitable to an attacker.

C. Recordability Resistance

When graphical passwords were originally released, one of the their most touted strengths was the fact that they would be harder to record than passwords. This was naïve, in hindsight. The ubiquity of mobile phones with cameras makes it trivial to record anyone doing anything, including using the mouse to enter the graphical password. Even without the use of additional electronic gadgets incorporating cameras, the computer user can record the graphical password easily using the ever-available screenshot facility. Storing such an image on the hard drive or printing it out is equivalent to an unencrypted textfile or the infamous post-it for the textual password.

V. COMPARISON

Before providing the comparison and discussing it, we discuss some of the requirements of textual passwords relevant for the comparison.

A. Some information about textual passwords

The following information has been considered for the comparison in the next subsection.

Guessability: The susceptibility of textual passwords to guessing attacks has been shown again [2] and again [48]. User-chosen textual passwords are not uniformly distributed in the password space. Malone and Maher showed that Zipf's Law is a relatively good model for password distributions [49]. Consequently, adversaries can easily create a dictionary of the most commonly used passwords. Weir, Aggarwal, Collins and Stern used such dictionaries to conduct an analysis of large sets of revealed textual passwords [2]. They were able to crack at least about one fifth of 8 character passwords in those sets and only slightly less of the passwords with a length of 9 and 10 characters.

Observability: As described above in Section II it is important to differentiate between human and technical observers. Using technical means it is possible to reconstruct textual passwords from video footage or even from the sound of the keyboard input [50]. Human observers have to rely on visual input. Tari, Ozok and Holden [51] discovered that when users type long and obscure passwords, entry is more easily observed by shoulder-surfers than when typing simple and familiar words. Unfortunately, textual authentication secrets generated by users to protect bank accounts and tax records are likely to exhibit exactly these characteristics, so efforts spent by a user to be "secure" might actually backfire. Even so, most users are fairly confident that observers cannot guess their password with any degree of accuracy [52] even though such confidence is probably misplaced.

Recordability: Passwords are trivial to record and users can and will write down their passwords when the burden of recovering a lost password is too high [53]. Security professionals realise that this is an inherent vulnerability of

textual passwords, so they deal with it by instructing people not to take this action. However, their instructions are mostly ignored [3].

B. Overview of the comparison

Table II presents a comparison between the textual password and some graphical authentication schemes. To give a better idea of how secure an authentication scheme is, we introduce the following four levels.

1) A scheme is considered equal to textual passwords in our setting (=), when it offers roughly the same resistance to common attacks as does the password.
2) A scheme is considered worse than textual passwords in our setting (↓), when it offers even less resistance to common attacks than the password.
3) A scheme is considered better than textual passwords in our setting (↑), when it is better than the password in this particular respect.
4) A scheme might not allow a rating (?) if no data regarding the aspect is available or the available data does only allow a very rough estimation instead of a real assessment (e.g. very small sample).

These levels are based on the literature which often reports findings that are extremely difficult to compare, so the comparison should not be considered definitive, but rather based on an understanding of whether the approach is prone to show vulnerabilities. Moreover, it becomes apparent that there are many aspects that do not allow a rating due to missing data or data that only allows a very rough estimation instead of a real assessment.

In terms of *guessability*, graphical mechanisms generally can be as strong as textual passwords in practise are. Most schemes have a variable password space and can therefore easily be adopted to be resistant to pure brute force attacks. This, however, requires special configurations to strengthen them, which have not yet been tested in the wild. Dictionary attacks remain a severe concern for many schemes, but examples such as PCCP show that guiding the user during password choice using persuasive technology can mediate this issue. Recognition-based schemes in which the password is issued by the system can even avoid this problem entirely. *Observability* is a problem across the board and attempts to introduce redundancy or indirection into the process tend to increase authentication times unacceptably (eg. up to 180 seconds [54]). However, it must be acknowledged that comparisons between textual passwords and graphical schemes regarding the vulnerability to shoulder surfing are hard to find, even among those schemes specifically designed to be observation resistant. It could be that they are equally resistant or even better as sometimes appraised by their proponents, but in the absence of hard data we cannot judge. Also, whether the approach of Dunphy, Heiner and Asokan can be successfully applied to a wider range of schemes than initially proposed might be an interesting topic for future research. In terms of *recordability* there does not seem to be much difference between graphical and textual passwords. Recognition-based secrets are mostly more *memorable*, but the other types can display the same memorability as textual passwords.

C. Discussion

The previous section has shown that many of the proposed graphical authentication schemes exhibit advantages and disadvantages in one area or another. Most have advantages regarding their memorability. This is no surprise, as the intention behind graphical passwords was to relieve users of the cognitive burdens of textual passwords.

The predictability of the users' drawings and the complete recreation of the secret during authentication in the drawmetric approaches (both recall and cued-recall based) cause severe concern and it is unclear or open for future research whether they perform good enough with respect to their memorability if the system would set the drawing for the user. However, the iterative processes the locimetric appraoches (i.e. the remaining cued-recall based approaches) and the searchmetric/recognitaion-based approaches have gone through has resulted in a more robust set of mechanisms with respect to guessability. For example, PCCP is very guessability resistant in particular compared to the textual passwords in our setting. The same holds for recognition-based schemes with system issued passwords.

While, the focus of graphical password research was on improving memorability they raised new usability issues: It is clear from the literature that it takes at least as long [6] for users to authenticate using these mechanisms and mostly even longer [30], [55], [1]. Thus, depending on how regularly one needs to authenticate some of the graphical alternatives might not be an alternative although security wise at least as secure as textual passwords. At creation time this makes them inconvenient but also extends the window for observation. Whether the timings achieved by systems more resilient to observation or predictability are acceptable might be up for debate, as the difference decreases the longer the period between two logins with the same password [1].

However, when considering the time it takes to login one should also consider the time and effort it takes to reset passwords. Thus, if it is much less often needed to replace a secret for a particular type of graphical password due to its superior memorability, a longer login time might become acceptable. The resulting benefits in time expenditure and convenience might well be worth the offset. To make a decision here it is very important to have detailed knowledge about the situation and application for which an alternative is considered.

Graphical authentication has its strengths and its weaknesses. Where authentication timings are of the essence other solutions might be a better choice. However, when retention times are high (consider e.g. a task that has to be carried out a few times a year), graphical passwords, with their superior memorability, can mediate.

Another important aspect is that one can never look at authentication in isolation. The context of deployment has to be considered. Different devices impose different constraints

on the authentication process. The two most prominent are desktop computers and mobile devices such as smartphones or tablets. Desktop computers and laptops normally offer high resolution screens and diverse input devices. The variety of techniques that can be applied is thus larger. For recognition-based schemes, however, the available data suggest, that the chosen bitstrength has a severe influence on the efficiency of the systems. Traditionally, mobile devices offer fewer resources such as smaller screens and especially a limited physical keyboard (if at all). Hence, keyboard-based solutions are mostly infeasible in the mobile environment. The most important factor here is eavesdropping through shoulder-surfing, as usage of such devices often occurs in public places [15]. As on mobile devices PIN-equivalent security is the de facto standard, recognition-based authentication schemes using shoulder-surfing resistant variants, as e.g. proposed by Dunphy, Heiner and Asokan, come to mind. They offer an alternative to textual passwords comprising both, high usability and security equal to the PINs they ought to replace.

VI. Conclusion

The base question behind this work was whether graphical authentication can be as strong as textual authentication. Based on our analysis the answer must be: "it can be, if you design it properly." So, what constitutes a proper design? Regarding the three metrics guessability, observability and recordability the following three considerations can serve as a first quick assessment:

(1) Unguided user choice translates to predictable choice. A resistant scheme should specifically encourage or even better force users to choose random passwords or have the system issue passwords.

(2) Obfuscation techniques, such as the asterisks routinely used to obscure textual password entry or portfolio-based approaches are examples of observation resistance.

(3) The scheme should generate secrets that are hard to describe or record.

These three considerations should serve only as a starting point for an evaluation. In some situations not all of the three aspects above might be important. For example, the ubiquitous textual password does not conform to statement (3). Whether a scheme is appropriate for a certain situation depends on the context of use, the risk associated with the asset the authentication mechanism controls access to, the time constraints, the device being used and the frequency of use. If the mechanism is low risk and used infrequently, graphical authentication might well be better than textual authentication.

Graphical schemes have the potential to be as secure as textual systems. Yet, the jungle of diverse graphical authentication schemes easily explains decision-makers' reluctance to adopt graphical authentication. They are rightly sceptical about the strengths of the mechanisms and also still in a one-authenticator-for-everything mindset. If we want this mindset to change, we, as researchers, will have to provide a way for decision makers to start deploying a wider range of mechanisms, in a more nuanced and discriminating fashion than a one-size-fits-all password. What we should be striving for is more diversity of authentication mechanism usage, deploying the wide variety of mechanisms that have been trialled in situations where we can match them to the risk mitigation requirement and deployment context. This way, users would finally benefit from superior memorability of available alternatives and organisations from less secret reuse across systems.

Therefore, we want to encourage the research community to try and fill the gaps in the knowledge about the promising graphical authentication schemes already in existence, rather than proposing even more. Unification of testing methodologies and comparability of approaches should be the goal. This can facilitate decisions to integrate one of the schemes in a particular application or environment and could give insights important not only for the domain of graphical authentication, but for authentication as a whole.

References

[1] R. Dhamija and A. Perrig, "Déjà vu: a user study using images for authentication," in *Proc. SSYM '00*, 2000, pp. 4–4.

[2] M. Weir, S. Aggarwal, M. Collins, and H. Stern, "Testing metrics for password creation policies by attacking large sets of revealed passwords," in *Proc. CCS '10*, 2010, pp. 162–175.

[3] A. Adams and M. A. Sasse, "Users are not the enemy," *Comm. of the ACM*, pp. 40–46, 1999.

[4] N. Ben-Asher, N. Kirschnick, H. Sieger, J. Meyer, A. Ben-Oved, and S. Möller, "On the need for different security methods on mobile phones," in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '11, 2011, pp. 465–473.

[5] B. Ives, K. R. Walsh, and H. Schneider, "The domino effect of password reuse," *Comm. of the ACM*, vol. 47, pp. 75–78, 2004.

[6] S. Chiasson, A. Forget, E. Stobert, P. C. van Oorschot, and R. Biddle, "Multiple password interference in text passwords and click-based graphical passwords," in *Proc. CCS '09*, 2009, pp. 500–511.

[7] G. Blonder, "Graphical password," Sep. 24 1996, uS Patent 5,559,961.

[8] K. Renaud and A. De Angeli, "Visual passwords: cure-all or snake-oil?" *Communications of the ACM*, vol. 52, no. 12, pp. 135–140, 2009.

[9] A. Paivio, *Mental representations. A dual coding approach*. New York: Oxford University Press, 1986.

[10] R. English and R. Poet, "Towards a metric for recognition-based graphical password security," in *Network and System Security (NSS), 2011 5th International Conference on*. IEEE, 2011, pp. 239–243.

[11] A. De Angeli, L. Coventry, G. Johnson, and K. Renaud, "Is a picture really worth a thousand words? exploring the feasibility of graphical authentication systems," *International Journal of Human-Computer Studies*, vol. 63, no. 1, pp. 128–152, 2005.

[12] X. Suo, Y. Zhu, and G. S. Owen, "Graphical passwords: A survey," in *Computer Security Applications Conference, 21st Annual*. IEEE, 2005, pp. 10–19.

[13] S. Komanduri, R. Shay, P. G. Kelley, M. L. Mazurek, L. Bauer, N. Christin, L. F. Cranor, and S. Egelman, "Of passwords and people: measuring the effect of password-composition policies," in *Proc. CHI '11*, 2011, pp. 2595–2604.

[14] E. Barker, W. Barker, W. Burr, W. Polk, and M. Smid, "Recommendation for key management - part 1: General," in *NIST Special Publication 800-57, NIST*, 2005.

[15] P. Dunphy, A. P. Heiner, and N. Asokan, "A closer look at recognition-based graphical passwords on mobile devices," in *Proc. SOUPS '10*, 2010, pp. 3:1–3:12.

[16] S. Brostoff, P. Inglesant, and M. Sasse, "Evaluating the usability and security of a graphical one-time pin system," in *Proceedings of the 24th BCS Interaction Specialist Group Conference*. British Computer Society, 2010, pp. 88–97.

[17] M. Zviran and W. J. Haga, "Cognitive passwords: the key to easy access control," *Computers & Security*, vol. 9, no. 8, pp. 723–736, 1990.

[18] P. Dunphy and J. Yan, "Do Background Images Improve "Draw a Secret" Graphical Passwords?" in *Proceedings of the 14th ACM Conference on Computer and Communications Security*. ACM, October 2007, pp. 36–47.

[19] I. Jermyn, A. Mayer, F. Monrose, M. Reiter, and A. Rubin, "The Design and Analysis of Graphical Passwords," in *Proceedings of the 8th USENIX Security Symposium*. Washington DC, 23-26 August 1999, pp. 1–14.

[20] J. Thorpe and P. Van Oorschot, "Graphical dictionaries and the memorable space of graphical passwords," in *13th USENIX Security Symposium*, 2004, pp. 135–150.

[21] D. Nali and J. Thorpe, "Analyzing user choice in graphical passwords," *School of Computer Science, Carleton University, Tech. Rep. TR-04-01*, 2004.

[22] A. J. Aviv, K. Gibson, E. Mossop, M. Blaze, and J. M. Smith, "Smudge attacks on smartphone touch screens," in *Proceedings of the 4th USENIX conference on Offensive technologies*. USENIX Association, 2010, pp. 1–7.

[23] M. Bond, "Comments on gridsure authentication," 2008, http://www.cl.cam.ac.uk/ mkb23/research/GridsureComments.pdf.

[24] J. Goldberg, J. Hagman, and V. Sazawal, "Doodling our way to better authentication," in *CHI'02 extended abstracts on Human factors in computing systems*. ACM, 2002, pp. 868–869.

[25] S. Wiedenbeck, J. Waters, J. Birget, A. Brodskiy, and N. Memon, "PassPoints: Design and Longitudinal Evaluation of a Graphical Password System," *International Journal of Human-Computer Studies*, vol. 63, no. 1, pp. 102–127, 2005.

[26] J. Thorpe and P. van Oorschot, "Human-Seeded Attacks and Exploiting Hot-Spots in Graphical Passwords," in *Proceedings of the 16th USENIX Security Symposium*. USENIX Association, 06-10 August 2007, p. 8.

[27] A. Dirik, N. Memon, and J. Birget, "Modeling user choice in the PassPoints graphical password scheme," in *Proceedings of the 3rd Symposium on Usable Privacy and Security*. ACM, 18-20 July 2007, pp. 20–28.

[28] R. Biddle, S. Chiasson, and P. van Oorschot, "Graphical passwords: Learning from the first twelve years," *ACM Computing Surveys 44(4)*, 2011.

[29] S. Chiasson, P. van Oorschot, and R. Biddle, "Graphical Password Authentication Using Cued Click Points," *European Symposium on Research in Computer Security*, vol. 4734, pp. 359–374, September 2007.

[30] S. Chiasson, E. Stobert, A. Forget, R. Biddle, and P. C. Van Oorschot, "Persuasive cued click-points: Design, implementation, and evaluation of a knowledge-based authentication mechanism," Carleton University, Ottawa, Canada, Tech. Rep., 2011.

[31] K. Renaud, "On user involvement in production of images used in visual authentication," *Journal of Visual Languages & Computing*, vol. 20, no. 1, pp. 1–15, 2009.

[32] D. Davis, F. Monrose, and M. Reiter, "On User Choice in Graphical Password Schemes," in *Proceedings of the 13th USENIX Security Symposium*, 2004, pp. 151–164.

[33] J. N. Maguire, "An ecologically valid evaluation of an observation-resilient graphical authentication mechanism," Ph.D. dissertation, Computing Science, 2013.

[34] R. English and R. Poet, "Measuring the revised guessability of graphical passwords," in *Network and System Security (NSS), 2011 5th International Conference on*. IEEE, 2011, pp. 364–368.

[35] D. Lin, P. Dunphy, P. Olivier, and J. Yan, "Graphical Passwords & Qualitative Spatial Relations," in *Proceedings of the 3rd Symposium on Usable Privacy and Security*. ACM, 18-20 July 2007, pp. 161–162.

[36] S. Chakrabarti, G. Landon, and M. Singhal, "Graphical passwords: drawing a secret with rotation as a new degree of freedom," in *The Fourth IASTED Asian Conference on Communication Systems and Networks (AsiaCSN 2007)*. Citeseer, 2007, pp. 561–173.

[37] S. Wiedenbeck, J. Waters, J.-C. Birget, A. Brodskiy, and N. Memon, "Passpoints: design and longitudinal evaluation of a graphical password system," *Int. J. Hum.-Comput. Stud.*, vol. 63, pp. 102–127, 2005.

[38] P. C. Van Oorschot, A. Salehi-Abari, and J. Thorpe, "Purely automated attacks on passpoints-style graphical passwords," *Trans. Info. For. Sec.*, vol. 5, pp. 393–405, 2010.

[39] P. Dunphy, J. Nicholson, and P. Olivier, "Securing passfaces for description," in *SOUPS '08: Proceedings of the 4th symposium on Usable privacy and security*. ACM, Jul. 2008.

[40] S. Chiasson, R. Biddle, and P. van Oorschot, "A Second Look at the Usability of Click-based Graphical Passwords," in *Proceedings of the 3rd Symposium on Usable Privacy and Security*. ACM, 2007, pp. 1–12.

[41] S. Wiedenbeck, J. Waters, J. Birget, A. Brodskiy, and N. Memon, "Authentication using graphical passwords: Basic results," in *Proc. HCI International '05*, 2005.

[42] Real User Corporation, "The science behind passfaces," 2004. [Online]. Available: http://www.realuser.com

[43] F. Tari, A. A. Ozok, and S. H. Holden, "A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords," in *Proc. SOUPS '06*, 2006, pp. 56–66.

[44] S. Brostoff and M. A. Sasse, "Are passfaces more usable than passwords? a field trial investigation," in *Proc. HCI '00*, 2000. [Online]. Available: http://www.cs.ucl.ac.uk/staff/S.Brostoff/index_files/brostoff_sasse_hci2000.pdf

[45] K. Renaud and E. Olsen, "Dynahand: Observation-resistant recognition-based web authentication," *Technology and Society Magazine, IEEE*, vol. 26, no. 2, pp. 22–31, 2007.

[46] S. Wiedenbeck, J. Waters, L. Sobrado, and J. Birget, "Design and evaluation of a shoulder-surfing resistant graphical password scheme," in *Proceedings of the working conference on Advanced visual interfaces*. ACM, 2006, pp. 177–184.

[47] K. Renaud and J. Maguire, "Armchair authentication," in *BCS-HCI '09: Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*. British Computer Society, Sep.

[48] C. Herley and P. van Oorschot, "A Research Agenda Acknowledging the Persistence of Passwords," *Security & Privacy, IEEE*, vol. 10, no. 1, pp. 28–36, 2012.

[49] D. Malone and K. Maher, "Investigating the distribution of password choices," in *WWW '12: Proceedings of the 21st international conference on World Wide Web*. ACM, Apr. 2012.

[50] L. Zhuang, F. Zhou, and J. D. Tygar, "Keyboard acoustic emanations revisited," in *Proceedings of the 12th ACM conference on Computer and communications security*, ser. CCS '05, 2005, pp. 373–382.

[51] F. Tari, A. Ozok, and S. Holden, "A Comparison of Perceived and Real Shoulder-surfing Risks between Alphanumeric and Graphical passwords," in *Proceedings of the 2nd Symposium on Usable Privacy and Security*. ACM, 12-14 July 2006, pp. 56–66.

[52] D. Weirich and M. Sasse, "Pretty Good Persuasion: A First Step towards Effective Password Security in the Real World," in *Proceedings of the 2001 Workshop on New Security Paradigms*. ACM, 2001, pp. 137–143.

[53] P. G. Inglesant and M. A. Sasse, "The true cost of unusable password policies: password use in the wild," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '10, 2010, pp. 383–392.

[54] D. Weinshall, "Cognitive authentication schemes safe against spyware," in *Security and Privacy, 2006 IEEE Symposium on*, 2006, pp. 6 pp.–300.

[55] K. M. Everitt, T. Bragin, J. Fogarty, and T. Kohno, "A comprehensive study of frequency, interference, and training of multiple graphical passwords," in *Proc. CHI '09*, 2009, pp. 889–898.