Measuring the Revised Guessability of Graphical Passwords

Rosanne English
School of Computing Science
University of Glasgow
Glasgow, Scotland
Email: rose@dcs.gla.ac.uk

Ron Poet
School of Computing Science
University of Glasgow
Glasgow, Scotland
Email: ron@dcs.gla.ac.uk

Abstract*—There is no widely accepted way of measuring the level of security of a recognition-based graphical password against guessing attacks. We aim to address this by examining the influence of predictability of user choice on the guessability and proposing a new measure of guessability. Davis *et al.* showed that these biases exist for schemes using faces and stories, we support this result and show these biases exist in other recognition-based schemes. In addition, we construct an attack exploiting predictability, which we term "Semantic Ordered Guessing Attack" (SOGA). We then apply this attack to two schemes (the Doodles scheme and a standard recognition-based scheme using photographic images) and report the results. The results show that predictability when users select graphical passwords influence the level of security to a varying degree (dependent on the distractor selection algorithm). The standard passimages scheme show an increase on guessability of up to 18 times more likely than the usual reported guessability, with a similar set up of nine images per screen and four screens, the doodles scheme shows a successful guessing attack is 3.3 times more likely than a random guess. Finally, we present a method of calculating a more accurate guessability value, which we call the revised guessability of a recognition-based scheme. Our conclusion is that to maximise the security of a recognition-based graphical password scheme, we recommend disallowing user choice of images.

I. Introduction

Users need to prove they are who they claim to be (authenticate) for many services, such as online banking, e-mail and e-commerce. Currently, alphanumerical passwords are the authentication mechanism of choice. However, passwords have a number of well documented problems associated with them, such as use of weak passwords and lack of memorability ([1],[7]). As an alternative, graphical passwords are often discussed (e.g. [9]). A graphical authentication mechanism is one in which the user is asked to select an image, points on an image, or draw an image in order to authenticate. Such schemes can be separated into three categories; recognition-based, recall-based and cued-recall [2].

This work concerns recognition-based schemes, in these schemes the user selects a number of "passimages". In order to authenticate they are presented with a number of "challenge screens". On each challenge screen, the user is shown one of their chosen images and a number of alternative "distractor" images. The user must then select their image from the distractor images for each challenge screen presented to successfully authenticate.

Whilst the security of cued-recall ([12],[13]) and recall-based schemes ([14]) have been researched in terms of predictability, the security of recognition-based schemes remains to be examined in equivalent depth. This work aims to be a step towards examining the level of security of such schemes.

User choice influences the security of passwords by making them susceptible to dictionary attacks and hence may also influence the security of recognition-based graphical passwords. This is highlighted by De Angeli *et al.* [5] who noted that the issue of predictability in user selected passimages still requires evaluation. This research assesses whether predictability of user choice affects the security level of recognition-based schemes and constructs an attack based on this predictability.

II. Background and Related Work

Researchers often report a chance of guessing (or guessability) for recognition-based graphical passwords as shown in Equation 1

$$\frac{1}{X^n} \tag{1}$$

where X is the number of images displayed on a challenge screen and n is the number of challenge screens (e.g. [4] and [6]).

Whilst the majority of work reports the guessability as described in Equation 1, Davis *et al.* progress further. In "On User Choice in Graphical Password Systems" [4] the authors implement two recognition-based schemes, Face and Story. Face was based on the PassFaces scheme created by Id-Arts Ltd.. The Story scheme let the user select a sequence of images to construct a "story" password, where each image selected was from a distinct category (e.g. cars, landscapes etc.). In both the schemes, the images are categorised into non-overlapping subsets of images (e.g. typical male, typical female etc. for the Faces scheme and cars, landscapes etc. for the Story scheme). The assumption was then made that the images in any given category are equivalent .

The authors then examined user choice of passimages in each category to determine the probability order (most chosen category to least chosen category). The authors used information resulting from the experiment to calculate guessing entropy as determined by Massey (closely related to Shannon's entropy) [8] to examine the security.

The work showed that if an offline exhaustive attack of the passimage space were possible, the attack would not take very long. They showed that in the Faces scheme, the inclination of users' choice towards attractive young females means it is far less secure by calculation of guessing entropy. However, as our work later shows, this is believed to be a result of selection of each of the distractors from the remaining distinct categories. The work we present in this paper provides further contributions as it makes explicit the type of attack which can be constructed to exploit preferences in user choice. It also attempts the attack against different schemes and reports the results, which provides concrete evidence to back up the calculations of guessing entropy reported by Davis *et al.*. In addition we show that the bias exists in other schemes, not just in the Faces and Story schemes.

III. Methodology and Results

A. Recognition-Based Graphical Password Schemes

In order to provide a basis of comparison, two recognition-based schemes are used. The first is the "Doodles" scheme [11] where black and white user drawings (doodles) are used for authentication. This was selected as an example of a recognition-based graphical password scheme where the user is involved in the creation of the passimage. In the authentication process, the user is presented with a screen which shows 16 doodles of which one is their passimage and the other 15 are distractors. A total of four such screens are presented in order to complete the authentication. In [10] the authors assert that the "guessability" is as shown in Equation 2

$$\frac{1}{16^4} = \frac{1}{65536} \tag{2}$$

The distractor selection algorithm for the doodles scheme is based on a measure of similarity between any two doodles (which is calculated based on the number of joins, black regions and white regions of a doodle), images dissimilar to the users doodle can be used as distractors.

The second scheme is a standard recognition-based scheme which offers the user 144 potential images to select as their "passimages" . This scheme was selected to represent examples of schemes where the user selects from a collection of predetermined images.The scheme uses a selection of the remaining images as the distractor images. To authenticate, the user is presented with 4 screens, each with 9 images giving a guessability as shown in Equation 3 .

$$\frac{1}{9^4} = \frac{1}{6561} \tag{3}$$

B. Experiments

Three stages are proposed:

1) Categorisation of the passimages of two recognition-based schemes into non-overlapping semantic categories;
2) Examine the distribution of passimage choices between these categories to determine the extent of user bias;
3) Exploiting the bias as extrapolated from the second step by construction of an attack

The first stage involves splitting the passimages and pass-doodles into non-overlapping semantic categories based on their content. The process will be carried out separately for both schemes resulting in a different set of categories for each scheme. For the images scheme images were categorised based on the prominent semantic content by the experiment conductor. However the doodles scheme presented more of a challenge, since the doodles used in this experiment were collected years previously and it was not possible to ask the users to categorise their own drawings.The categorisation of the doodles reflects the user choice of doodle passimages. The process of categorisation is described for both schemes in Section III-C and Section III-E. Users will be asked to select four images from the collection to create their graphical password, these choices will be examined for bias towards semantic categories. In the final experiment the potential for an attack which uses the most commonly selected images as the guess for a challenge screen will be examined. The rest of Section III discusses the specifics of these experiments and the corresponding results.

C. Categorisation of Image Passwords

144 digital images of objects are used to provide a selection of images for users to select a portfolio (set) of four images as their password. There are 12 categories for the images and are as follows: Food and Drink, Cartoon and Fictional Characters, Scenery, Animals, Faces and Body Parts, Transport, Clothing, Entertainment, Trees plants and flowers, Skyscapes, Buildings tools and devices, People. Each category has 12 images to remove potential bias in the selection due to more images being in any particular category.

D. Passimages User Selection Results

A total of 64 individuals participated, each selecting a combination of 4 images, resulting in 256 passimage selections. As expected, there was a bias in user selection. Of the 144 images, 42 images were not selected at all. Only the Scenery category had every image selected by a user at least once. The most unpopular category was "people" which showed images of non-famous individuals. Food and drink was the most popular category having 37 selections; cartoon characters came a close second with 30 selections. The distribution of user choice of images is shown in Figure 1 .

E. Categorisation of Doodle Passwords - Name That Doodle!

In order to obtain categories for the doodles, an initial experiment was carried out which asked users to classify or name doodles. It is proposed here that human classification is appropriate for understanding how users select their doodles. Since the users who created the doodle collection were not contactable to ask what they drew, participants were asked to categorise the doodles in collection. The purpose was to minimize any ambiguity over doodle content.

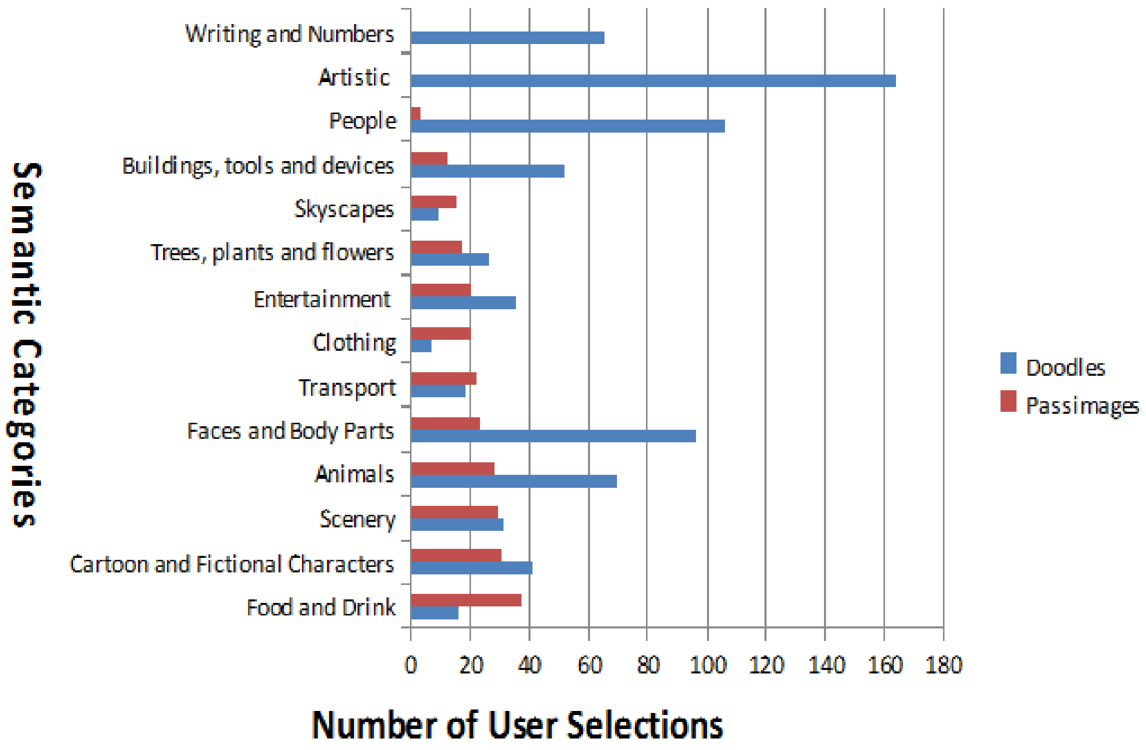


Fig. 1. Doodle and Passimages Collection Semantic Categories Distribution

The following descriptions were used to define equivalence (taken from the categories of classification as discussed by Poet and Renaud [11]) same word; same concept, different word ; different levels of category, same concept.Doodle descriptions were deemed distinct if they fell into the following categories: whole picture vs. detail; different perspective; completely different concepts. Each doodle then had multiple classifications, these were then further refined into categories by applying a majority ruling. If the majority of classifications fell into category X, then that doodle was in category X. The results of this experiment are detailed in Section III-F.

F. Doodle User Selection Results

A total of 44 users (24 male and 20 female) participated in the study providing amongst them 5012 classifications for the collection of 735 doodles.

14 major semantic categories emerged, with 96% of classifications falling within these categories, 4% being un-namable. 12 categories matched those of the images, with an additional two categories of "writing and numbers" and "artistic", both of which related specifically to drawing and writing. Writing and numbers included doodles of letters, words, logos and numbers whilst artistic contained squiggles, random art and shapes . Classifications shown as examples are provided verbatim, ignoring spelling and grammatical errors.

People (which included classifications such as "man"," female", "people",), animals (which included classifications such as " cat", "dog" and "animal") and faces & body parts (which included classifications such as "eye", "ear", "face") were by far the most popular classifications accounting for 33% of the classifications between them.

After the classifications had been analysed as detailed in Section III-E the final dispersion of doodles between the categories was as shown in Figure 1, which also shows passimages choice distribution. It should be noted that passimages had two less categories (writing & numbers and artistic were not included) since these categories did not apply to any images in the collection. These results established the bias in user choice towards images in particular categories.

G. Examining User Choice- Semantic Ordered Guessing Attack (SOGA)

In order to exploit the established bias in user choice, we propose an attack which we call a "Semantic Ordered Guessing Attack" (SOGA). In this attack the attacker has knowledge of the most likely categories for graphical password selection. This information could be obtained in a similar method to that described in Sections III-C and III-E. The attacker is then presented with an authentication challenge screen which has the passimage or doodle and the selection of distractors. The attacker then attempts to authenticate by selecting the image from the most likely category given the screen presented.

In order to examine the potential of a SOGA an experiment was carried out to select distractor images and perform the attack each of the passimages (doodles and images) multiple times. A program was written to select each passimage/doodle, along with its distractors (using the algorithm created by Poet and Renaud for doodles, three variants on random selection for images) .It was then recorded whether the password was in the most likely category for the collection of distractors selected and recorded the result. The results are presented in Sections III-H and III-I.

H. Doodle SOGA

Each attack consisted of selection of a passdoodle and 15 distractor images (due to 15 distractors per challenge screen). The attack then selected the doodle from the most common category (using the ordering which is shown in Figure 1) and checked to see if it was the passdoodle. The attack was carried out for each of the 735 doodles 4 times (due to the 4 challenge screens for successful authentication).

The attack was performed for each of the 735 doodles 4 times (due to the use of 4 challenge screens in the doodles scheme). The results are displayed in Figure 2 where it can be clearly seen that the SOGA was successful for approximately 15% of the 2940 screen attacks.

Taking the example of 15% of attacks on individual screens being successful for the doodles scheme, the guessability is given as $\frac{1}{16^4} = \frac{1}{65536}$ when using Equation 1 (4 screens with 16 doodles per screen). The guessability using the SOGA for the doodles scheme is reported as 15% per screen, which is equivalent to roughly 6.67 doodles per screen (obtained by solving for x given the equation $\frac{1}{x} = 0.15$) which results in a guessability of $\frac{1}{6.67^4}$ for the four screens. This is approximately 33 times higher than the guessability (for random guessing) calculated using Equation 1 with 16 doodles for 4 screens.

I. Passimages SOGA

The approach taken was similar to that of the attack when performed on the doodles scheme. However, in contrast to the doodles scheme, each user selected image is not unique and so the number of iterations per passimage was dependent on the number of times it was selected by users. Each image was used as a passimage (for each time a user selected the image) and had distractors selected using each of three algorithms .The results were then analysed by ordering the collection of selected distractors categories and passimage category by popularity as discovered in our study. The number of instances where the passimage was in the most common category given the user choice bias (displayed in Figure 1) was then counted.

The results of the SOGA for the passimages scheme are shown in Figure 2. It can be seen from this figure that the most resistant distractor selection algorithm was random selection from categories other than the passimage category, which resulted in the attack working against 20% of the 256 passimage challenge screens (64 users who selected 4 images each). Performing the calculation (as we did for the doodle scheme) to determine the increase on chances of guessing when compared to random guessing probability (as presented in Equation 1) this meant a guessing attack using this approach is 10.5 times more probable than random guessing.

The second most effective algorithm was random selection, which resulted in the attack working for 21% of the 256 passimage challenge screens. This related to an increase on the probability of guessing of 12.8 times. The worst performance came from selection of distinct categories which resulted in the attack working for 23% of the 256 passimage challenge screens. this meant a guessing attack using this approach is 18 times more likely to be successful than random guessing. This was as expected, since there is a user bias in image selection towards more popular images, it follows that if you select from the remaining categories when a user has selected from a highly popular category, the distractors will be from less likely categories giving perfect conditions for the proposed attack.

IV. Computing the Guessability

We now propose a way of calculating the "revised guessability" for recognition-based graphical password schemes (where the images can be separated into distinct categories). This value is intended to be a more accurate reflection of the guessability when compared to those calculated without considering user bias. This is done by calculating the "revised" number of passimages per challenge screen and calculating the guessability using Equation 1 with this value. The steps are as follows:

1) Collect a sample of user selected images, a larger sample is better
2) Establish the bias in user choice by examining the categories of the passimages selected and ordering the categories from most to least popular
3) Calculate the revised number of images per screen by solving Equation 4 for x, the revised number of images per challenge screen.

$$\frac{1}{x} = \frac{\text{percentage of successful attacks}}{100} \tag{4}$$

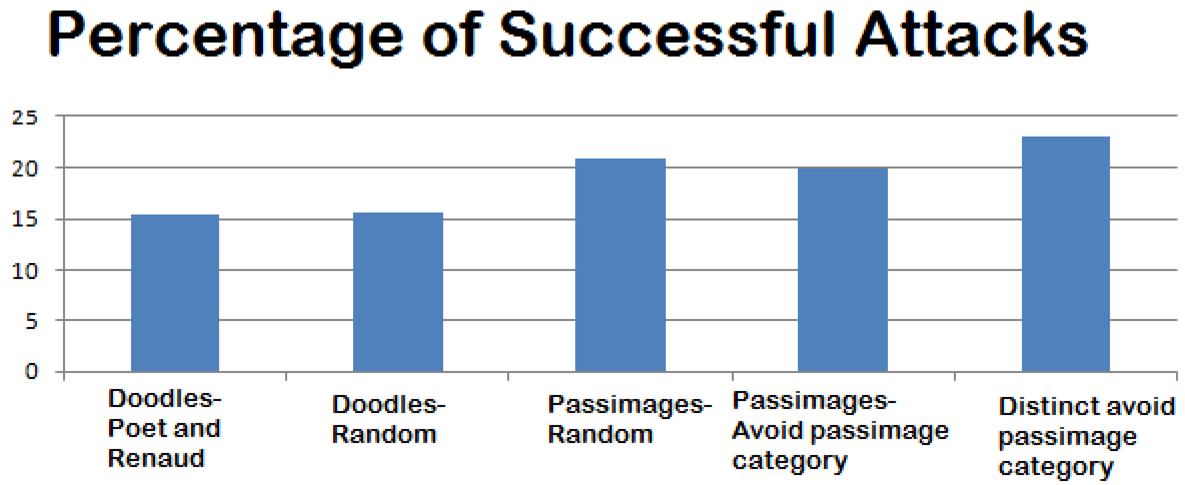
4) Finally, calculate the revised guessability as $\frac{1}{x^n}$ where x is as calculated in Equation 4 and n is the number of challenge screens.

V. Conclusions and Future Work

15% of attacks against doodles challenge screens were successful. Whilst 21% of passimage screens were successfully attacked where distractors were selected randomly (ignoring the semantic categories). 23% of passimage screens were successfully attacked where distractors were selected from distinct passimage categories (excluding the passimage category) and 20% of screens were successfully attacked where distractors were selected from passimage categories (excluding the passimage category).

These figures relate to attacks on individual screens, the attacks are made more difficult by a number of challenge screens. Passimage guessability varied between 10 and 18 times larger than predicted using Equation 1 for 9 images with

Fig. 2. Percentage of Successful SOGA for Each Scheme Variation



4 screens. The guessability is 33 times higher than calculated when using 16 images for 4 screens. To compare the two values, one must calculate the increase in guessability for the doodles scheme where 9 images are used per screen. This figure is approximately 3.3 times higher than predicted using Equation 1 which is noticeably smaller than the passimage results.

These results demonstrate how guessability of recognition-based graphical password schemes can be affected by user choice. This work supports the work of Davis *et al.* [4] and shows that the bias exists in other recognition-based schemes. We also showed how a guessing attack might be constructed to exploit this bias which we called a semantic ordered guessing attack. Following from this, we presented a method to calculate the revised guessability of a recognition-based scheme where the images can be split into distinct categories.

The main conclusion which can be drawn from the results reported here is that the hypothesis that user choice in passimages will have a detrimental effect on the level of security of recognition-based schemes is valid to a varying degree. The extent of the effect is dependent on the distractor selection algorithm ignoring bias in user selection. To minimise attacks, a distractor selection algorithm should not attempt to select less popular images for a challenge screen but instead emmulate random selection and preferably have a large range of categories for passimages. An alternative could be to show images from the same category, though this could reduce memorability.

Given the results we discovered in our studies, we suggest that to maximise potential security of a recognition-based scheme users should not be allowed to pick their own pictures, though it remains to be seen how this might affect memorability of such schemes. This work examines only guessing in terms of a bias in general image choice and does not consider guessing for a specific user, thus there is room to examine the use of social engineering tactics to guess a users pictures.

References

[1] Anne Adams and Martina Angela Sasse. Users Are Not The Enemy. *Communications of the ACM*, 42(12):46, 1999.

[2] Robert Biddle, Sonia Chiasson, and P.C. van Oorschot. Graphical Passwords: Learning from The First Generation, 2009.

[3] Sacha Brostoff and Martina Angela Sasse. Are Passfaces More Usable Than Passwords: A Field Trial Investigation. In *People and Computers XIV-Usability or Else: Proceedings of HCI*, pages 405–424, 2000.

[4] Darren Davis, Fabian Monrose, and M.K. Reiter. On User Choice in Graphical Password Schemes. In *Proceedings of the 13th conference on USENIX Security Symposium-Volume 13*, page 11. USENIX Association, 2004.

[5] Antonella De Angeli, Mike Coutts, Lynne Coventry, Graham Johnson, David Cameron, and Martin H. Fischer. VIP: a visual approach to user authentication. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pages 316–323. ACM, 2002.

[6] Antonella De Angeli, Lynne Coventry, Graham Johnson, and Karen Renaud. Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies*, 63(1-2):128–152, 2005.

[7] PG Inglesant and Martina Angela Sasse. The true cost of unusable password policies: password use in the wild. In *of the 28th international conference on Human factors in computing systems*, pages 383–392, 2010.

[8] James L. Massey. Guessing and Entropy. *Proceedings of 1994 IEEE International Symposium on Information Theory*, page 204.

[9] Fabian Monrose, M.K. Reiter, and Darren Davis. *Graphical Passwords*, chapter 9, pages 163–174. O'Reilly, 2005.

[10] Ron Poet and Karen Renaud. A Mechanism For Filtering Distractors for Doodle Passwords. *International Journal of Pattern Recognition and Artificial Intelligence*, 23(5):1005–1029, 2009.

[11] Ron Poet and Karen Renaud. An Algorithm for Automatically Choosing Distractors for Recognition Based Authentication using Minimal Image Types. *The Ergonomics Open Journal*, 2(3):178–184, January 2010.

[12] Amirali Salehi-Abari, Julie Thorpe, and PC Van Oorschot. On Purely Automated Attacks and Click-Based Graphical Passwords. *2008 Annual Computer Security Applications Conference (ACSAC)*, pages 111–120, December 2008.

[13] Julie Thorpe and PC Van Oorschot. Human-seeded attacks and exploiting hot-spots in graphical passwords. *Proceedings of 16th USENIX Security*, pages 103–118, 2007.

[14] Julie Thorpe and P.C. van Oorschot. Graphical Dictionaries and the Memorable Space of Graphical Passwords. *usenix.org*, 2004.

[15] P.C. van Oorschot and Julie Thorpe. On the Security of Graphical Password Schemes, 2005.