Proceedings of the Eighth International Symposium on
Human Aspects of Information Security & Assurance (HAISA 2014)

Modelling the Security of Recognition-Based Graphical Passwords

R. English

University of Glasgow, United Kingdom
e-mail: Rosanne.English@glasgow.ac.uk

Abstract

Recognition-based graphical passwords have received attention in recent research as an alternative authentication mechanism. The research often presents new schemes, usability studies or proposes countermeasures for specific attacks. Whilst this is beneficial, it does not allow for consistent comparison of the security of recognition-based graphical password schemes. This paper contributes a proposed solution to this problem. Presented here are mathematical models for estimating the number of attacks required before success for four attack types. These models combine to provide an overall metric of the security of recognition-based graphical password schemes. The metric presented provides a consistent, repeatable, and quantitative method for comparing recognition-based graphical password schemes which was previously not possible.

Keywords

Recognition-based graphical passwords, metrics, security

1. Introduction

A recognition-based graphical password (RBGP) scheme is an alternative authentication mechanism where the user selects a number of images called passimages (Charrau *et al.*, 2005) to be used to authenticate. In this paper the collection of the user's passimages will be called their passimage set. When the user attempts authentication they are presented with a number of challenge screens that present at least one of their passimages and a number of alternative images, called distractor images. To successfully authenticate, the user must identify and select their passimage from the distractor images on each screen. Further information on graphical passwords can be found in literature, for example reviews by Biddle *et al.* (2011) and Suo *et. al* (2006).

In addition to RBGPs, there are two further categories of graphical passwords - recall and cued recall. The security of recall and cued-recall graphical passwords have been considered in terms of the ease of guessing in a consistent manner by examining potential password space and bias in user selections. In contrast, analysis of the security of RBGPs has been arguably inconsistent. For example, one approach to calculating the entropy (hence guessability) of a RBGP is proposed by Hlywa *et al.* (2011), whilst a different approach to measuring guessability is reported by DeAngeli *et al.* (2005) and Dhamija and Perrig (2000). In addition, consideration of the security of RBGPs often focuses on countermeasures for a specific attack (e.g.

shoulder surfing is the focus of Wiedenbeck *et al.* (2006) and Sasamoto *et al.* (2008)). Whilst this is useful, it remains unclear how to compare proposed RBGP schemes in terms of their security. There remains no standardised method of measuring the level of security of a RBGP scheme. This paper contributes to the measurement of RBGP security by proposing a metric that allows the security to be assessed and compared in terms of resistance to four identified attacks.

The approach taken is to construct a tuple that consists of individual metrics for each attack type considered. Each metric is presented as a mathematical model, which uses the configuration of the RBGP scheme to estimate the number of attacks required before the attacker is successful. Models for four attacks: random guessing, guessing based on category bias, frequency attacks, and shoulder surfing attacks are presented. These are combined into the overall security metric tuple which is applied to a number of examples and evaluated.

The remainder of this paper is structured as follows. The scope of the work by considering the variables to be used in the metric models (the configuration of the RBGP schemes) and the attacks under consideration are presented in Section 2. The individual mathematical models which combine to give the overall metric are presented in Section 3. A summary of how to apply the metric and examples are presented in Section 4, and the conclusion is presented in Section 5.

2. RBGP Scope

To clearly identify the scope of the work it was necessary to establish the attributes of a RBGP scheme which contribute to their configuration and identify the attacks to be considered in the metric. These are presented in this section.

2.1. RBGP Configurations

A list of variables which contribute to the configuration of RBGP schemes is established here to identify aspects of the configuration which may have an impact on the success rate of different attacks. The different aspects contributing to the configuration of RBGPs were established from current literature as the number of passimages (denoted p for this work), the number of challenge screens (s), the number of distractors per challenge screen (d), the number of constant distractors per passimage (c), and whether the passimages were assigned to the user or not.

A review of RBGP schemes in literature identified 17 RBGP schemes with sufficient information regarding configuration. These schemes can be split into two groups. One group consists of a single challenge screen with multiple passimages presented on this screen. Nine of the 17 RBGP schemes identified presented only one challenge screen. The remaining eight schemes presented represent the group of schemes which present a single passimage on multiple challenge screens, thus the metric considers this configuration.

Using a single image on multiple screens can be further refined by the passimage selection being restricted to a specific order, or order being irrelevant. No schemes which had multiple challenge screens with one passimage per screen where order was important. Thus, the metric assumes unordered selection. Images can also be assigned to the user, selected by the user from a provided set of images, or uploaded by the user. These options were approximately equally distributed in the schemes identified, however it was felt that allowing users to upload their own image was potentially too guessable (see Tullis and Tedesco, 2005) and conversely using an assigned set of images may impact memorability. Thus it was decided to examine user selected images from a pre-defined set.

2.2. Attacks Considered

Once these configurations were established, it was necessary to identify areas of potential threats to identify the attacks to be considered. DeAngeli *et al.* (2005) propose that security of authentication mechanisms can be judged in terms of three aspects; guessability (the probability an attacker can guess the user's password), observability (the probability of an attacker being able to observe the authentication process), and recordability (the ease with which a user can record the user's password). Recordability was defined as outside the scope as it relates to how easily the password can be recorded. It is unclear to what extent users may record their passimages and how easily an attacker may gain access to this information. This is not an aspect which could feasibly be modelled.

A total of four attacks were considered - two guessing attacks and two observation attacks. These were random guessing, semantic ordered guessing, shoulder surfing, and frequency attacks. Random guessing, shoulder surfing, and intersection/frequency attacks were identified as attacks which are often identified in literature and hence deserved consideration for the proposed metric.

A semantic ordered guessing attack is an attack where guessing is prioritised based on the semantic category of the images (assuming the common approach of using semantically themed images such as faces, objects etc.). Studies exploring the feasibility of these semantic ordered guessing attacks in which the attacker selects the "most probable" image given the challenge screen presented are reported in English and Poet (2011). Results showed that bias in user choice could decrease the estimated guessability by varying degrees dependent on how distractors are selected for a given challenge screen. On average, guessing using a prioritised attack was 13 times more likely to succeed than random guessing for a passimages scheme. The work by Davis *et al.* (2004) and the prior related work both indicate the feasibility of prioritised guessing attacks. There may be a number of different approaches to prioritising images for guessing attacks. A SOGA was included to represent a prioritised guessing attack similar to that proposed by Davis *et al.* (2004) where the information required to construct the model was readily available.

At this stage the variables relating to the configuration of a RBGP have been identified, and the attacks under consideration have been selected. Thus, the next

step is to construct the metric. The approach taken is to establish a tuple that consists of an estimated number of attacks required before successful authentication for each type of attack considered. A tuple approach was considered appropriate instead of combining values (e.g. by summing the scores) or using a Euclidean metric since the interpretation of security is context sensitive. For example, in the context of authentication in a home environment where no other individual is present, a negative shoulder surfing value would not be a concern. Thus, it would not be appropriate to reduce the overall security score due to this. Another approach could have been to weight the individual values before combining them. However, the weighting could be different depending on context. The resulting tuple represents the security of a RBGP scheme in terms of the attacks identified. Presented here is a 4-tuple metric consisting of four estimated values of the number of attacks required before successful authentication. There is one estimate for each of the attacks; random guessing, semantic ordered guessing, shoulder surfing and frequency attacks. The calculation of each of the component parts is summarised in the following four subsections.

3. Establishing the Metric

3.1. Models for Guessing

3.1.1. Random Guessing

The estimate of the number of random guessing attacks required before success is obtained from the calculation of the probability of success. This is commonly reported as $\frac{1}{x^s}$ where x is the number of images shown on a challenge screen (the number of distractors plus one passimage, d+1) and s is the number of challenge screens. The denominator of this calculation is used to provide an estimate of the number of random guessing attacks required before success, thus the RG value is calculated as shown in Equation 1.

$$RG=(d+1)^s \tag{1}$$

3.1.2. Semantic Ordered Guessing

The calculation of the number of semantic ordered guessing attacks (SOGA) required before success relies on an estimate of the number of attacks which are successful for a given potential passimage set. This is calculated by performing simulations of SOGAs based on the category distribution of real user choices. Further details of such simulations are presented in English and Poet (2011) where the following percentages of success were achieved: 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 success rates can be used as estimates for user selected passimage schemes where the images can be split into semantic categories.

Once the percentage of success has been estimated, one can calculate the estimated number of attacks as shown in Equation 2 where *s* denotes the number of challenge screens. If the passimages are assigned to the user, then this attack is not applicable and this is denoted by *.

$$\left(\frac{100}{successPercentage}\right)^{s} \tag{2}$$

3.2. Models for Observability

To establish the models for observability simulation software was built, the purpose of which was to represent a RBGP scheme with a given configuration, construct a user's passimage set and allow frequency and shoulder surfing attacks to be emulated against that set. The RBGP scheme can have a varied configuration, in addition if a shoulder surfing attack is being simulated an attacker has a percentage of recall, which reflects their ability to recall the passimages observed. After construction of the simulation software it was possible to simulate each attack type to establish which variables of the configuration of a RBGP scheme had a significant impact on the success rate of the attacks. These variables were then used to run 500 simulations at a variety of configurations for each variable. This resulted in a collection of data which could be used as the basis for mathematical modelling. The modelling process was repeated multiple times to obtain more accurate (better fitted) models. Due to space restrictions further information cannot be included, but is available in detail in Chapters 6 and 7 in http://www.dcs.gla.ac.uk/~rose/2012EnglishPhd.pdf. The following sections present the final models for shoulder surfing attacks and for frequency attacks.

3.2.1. Shoulder Surfing Value

As for the semantic ordered guessing value, one must estimate the percentage of recall rate or success rate of an attacker given a specific shoulder surfing countermeasure. This can be done by performing an experiment to establish how successful shoulder surfing attacks are for the countermeasure implemented. Alternatively an estimated value of successful recall between 1 and 100% can be chosen. Once the recall value has been established, the shoulder surfing value can be calculated as shown in Equation 3 where *p* denotes the number of passimages in a user's passimage set, *s* is the number of challenge screens in a session, and *r* is the percentage of recall. The modelling was based on log_2 of the median number of attacks and so the final equation includes a power of 2.

$$SS = 2^{1.3852p - 0.0824p^2 - 0.2143s - 0.0472r + 0.0002r^2} \tag{3}$$

3.2.2. Frequency Value

An intersection attack, as defined by Dhamija and Perrig (2000) is an attack in which the attacker records multiple challenge screens and notes the images which are constant between two screens. Assuming all distractor images change this would result in the passimage being identified. Takada *et al.* (2006) identify a similar attack which they call a frequency attack. In a frequency attack, the attacker notes multiple challenge screens and notes the frequency with which each image appears then selects the image which occurs most frequently for any given screen. For this work a frequency attack will considered primarily since an intersection attack can be thought of as a special case of a frequency attack.

Unlike the previous two calculations, the frequency model relies primarily on the configuration of the RBGP scheme (and not user choice distribution or attacker recall). This includes the number of distractors kept constant per passimage (denoted by c) in addition to the number of screens (s), the number of distractors per screen (d), and the passimage set size (p). The frequency value can be calculated as shown in Equation 4. The modelling was based on log_2 of the median number of attacks and so the final equation includes a power of 2.

$$FREQ=2^{0.0156p+1.6655s+0.9497c-0.5575d+0.018p^2+0.0132s^2-0.0344c^2+0.0309d^2} \quad (4)$$

This equation should only be used if the number of distractors kept constant per passimage is less than the number of distractors per challenge screen. If the challenge screens are constant then a frequency attack will be reduced to a random guessing attack. In this case, * denotes the attack is not applicable.

3.3. Overall Metric

Now each individual model has been determined, it is possible to combine these into the final metric. The metric is denoted as shown in Equation 5 where RG denotes the random guessing value, SOGA denotes the semantic ordered guessing attack value, SS denotes shoulder surfing value and FREQ denotes frequency attacks value. If for any of the attacks a countermeasure is implemented which means the attack is not possible, then a * is used to denote this.

$$(RG,SOGA,SS,FREQ) \quad (5)$$

There are a number of limitations of the metric which should be considered. The final metric models are based primarily on simulations, and so the reality of attacks may be different. This approach provided a flexible and controlled alternative to a large scale user study which was attempted but was unsuccessful in recruitment of sufficient participants. Also, the work primarily considers RBGP schemes with a predetermined set of images (which was constant for the duration of the work) and does not consider user provided images. It is necessary to be careful not to use the models for prediction, i.e. applying configuration values outside the values used in the simulations. This is because the model was based on the configurations used in

the simulations, and values outside this could deviate substantially from the models. The models could be used outside the ranges, but care must be taken in interpretation of the prediction. Note that a prediction arises from the metric where configurations outside those upon which the models are based are used. An estimation is provided where configurations used were incorporated into the model. To minimise the need to apply values outside the configurations used, the simulations used configurations from literature to date and values either side. For example, 4 challenge screens are common, and simulations were run with 1 through to 10 screens.

Another potential issue is with the interpretation of the values resulting from these models. One must not consider the values reported as a concrete value of the number of attacks required in any given case. The values reported are estimates based on simulations, in reality other factors such as a combination of shoulder surfing, frequency and guessing attacks could be used which cannot be represented by these models. However, the purpose of this work was not 100% accuracy, but to provide an estimate which could be used to achieve a comparison of the security of different RBGP configurations.

4. Using the Metric and Example Application and Comparison

Now the metric has been established it is appropriate to discuss how to apply the metric and use it in decision making. This section aims to discuss these aspects.

To calculate the component values for the tuple the following approach is taken. First examine the RBGP scheme to establish values for the configuration (as previously indicated, *p,s,d,c,r* and *successPercentage*). Next, establish if any of the attacks are not feasible for the scheme being examined. For any such attack, use a * in the appropriate place in the metric tuple to denote the attack is not applicable to the scheme. For each of the attack types remaining use the appropriate configuration values (identified in step one) in the appropriate mathematical model described in Section 3. Next, round each of the model values to the nearest whole integer. Finally, combine the values in the order (RG, SOGA, SS, FREQ) to obtain the final metric as applied to the scheme under consideration.

The metric as applied to the scheme under consideration can now be examined in terms of the security either individually or against other schemes. For simplicity, this paper considers the comparison between two schemes as an example. This could be easily extrapolated to examine more than two schemes. To compare two schemes, scheme 1 and scheme 2, one should consider the values for each of the models within the tuple. Let us call the constituent tuple values of each scheme RG1, SOGA1, SS1, and FREQ1 for scheme 1 and RG2, SOGA2, SS2, FREQ2 for scheme 2. It is then possible to compare the values for each of the attacks e.g. RG1 can be compared to RG2 and so forth. Thus, if for example RG1 is larger than RG2 then we can deduce that scheme 1 is more resistant to random guessing attacks. Follow a similar approach for the remaining attacks.

In using this for decision making, for example to select an appropriate scheme, one should consider the context in which the scheme will be deployed. For example if observability is a key concern, but guessability less so then particular attention should be paid to the observability values. This may result in a situation where one scheme has a higher resistance for one attack and a lower resistance for the other whilst the scheme it is being compared to has the opposite resistance. In this situation it is down to judgement of the decision maker to consider what is most important in the context. Having now discussed how to use the metric, the next section aims to provide some example applications. Due to space restrictions only two schemes are included, but more applications of the metric are available in Chapter 8 of http://www.dcs.gla.ac.uk/~rose/2012EnglishPhd.pdf together with further discussion of benefits and limitations of the approach.

4.1. Application to PassFaces

The application of the final metric to the PassFaces scheme is presented here. From reviewing the PassFaces white paper (available from PassFaces (2005)) the following information on the configuration of the scheme was extracted s=p=4, d=8, c=8. Images are assigned and so a SOGA is not applicable, represented by *. Images appear highlighted upon selection potentially making shoulder surfing more successful as shown by Tari *et al.* (2006) where approximately 60% of attacks were successful, thus this value is used for the recall rate of PassFaces. The resulting metric for PassFaces is then calculated as (6561,*,2,*) where * represents that a frequency attack will be no better than random guessing since the number of distractors kept constant is equal to the number of distractors per screen.

From this result the weakest aspect of the security is shoulder surfing. If one were authenticating where the process could be viewed, then this could be an issue. The number of attacks required could be increased by doubling the number of passimages to 8, which results in a SS value of 7. It could be further increased by allowing keyboard entry, which results in a success rate of approximately 11% (again, shown by Tari *et al.* (2006)) which results in a shoulder surfing value of 22.

4.2. Application to Adapted VIP

Whilst the VIP scheme proposed by DeAngeli *et al.* (2002) has only one screen, it is adapted here to multiple challenge screens. This allows the metric to be applied to the scheme and provides an additional example. The metric is now applied to the adapted VIP1 scheme. Since there are four passimages in a session s=4 is used. From the defining paper, the configurations were as follows; with four passimages in a challenge session, p=10, d=9, c=0. The shoulder surfing recall was estimated at 60% (as assumed for the PassFaces scheme) since there was no details on highlighting the images upon selection, but the images were selected on a touchscreen. A SOGA was not applicable to the adapted VIP1 since the images were randomly assigned to the users. There was no mention of maintaining constant distractors for passimages and so this was assumed to be 0. It should be noted that the random guessability value may underestimate the resistance as the calculations do not account for sequence,

which is incorporated into the adapted VIP1 scheme. Also, location was maintained and thus there is potential for the shoulder surfing value to be overestimated as could be arguably easier to shoulder surf a passimage which stays in one position. The resulting metric is (10000,*, 6, 80).

4.3. Comparison

The purpose of this metric is to allow consistent comparison of the security of RBGP schemes. Using the metric to demonstrate this it is now possible to compare the security of the PassFaces scheme with the security of the adapted VIP1 scheme. It can be seen from the metrics reported in the previous section that the PassFaces scheme is more secure in terms of frequency attacks, but the adapted VIP1 scheme is more secure against random guessing and marginally more secure against shoulder surfing attacks due to the increased passimage set size. Both schemes are equally secure against SOGAs since passimages are assigned to users. In selecting an appropriate scheme, one would need to consider the context under which the mechanism would be used. For example, if shoulder surfing is not a concern then the PassFaces scheme may be a better choice.

5. Conclusions

This work aimed to present a model for the security of recognition-based graphical passwords. The overall model consisted of four smaller models which allow an estimation of the number of attacks required for the following attack types; random guessing, semantic ordered guessing, shoulder surfing, and frequency attacks. This was an important topic to research since alternatives to alphanumerical authentication are arising more but analysis of security can be limited where recognition-based mechanisms are considered. In particular, it was difficult to compare two schemes in terms of their respective levels of resistance to attack. This work has contributed to a resolution of this issue by proposing a metric which can be used for RBGP schemes where multiple challenge screens are presented with one passimage per screen. The consistent, objective, and quantitative approach now allows schemes to be readily compared in terms of resistance to the guessing and observation attacks discussed here. Previously this was not possible. The work provides an estimated number of attacks before success for each of the following attack types: random guessing, semantic ordered guessing, frequency, and shoulder surfing attacks. This metric can be used to establish the more appropriate scheme given a selection, or as a method of deciding which configuration is most appropriate for a particular context.

6. References

PassFaces Corp., 2005, Two Factor Authentication, http://www.realuser.com/ Last accessed 05/30/2014.

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

Charrau, D., Furnell, S., and Dowland, P., 2005. PassImages: An alternative method of user authentication, *Proceedings of 4th Annual ISOneWorld Conference and Convention, Las Vegas, USA*

Davis, D., Monrose, F., and Reiter, M., 2004. On User Choice in Graphical Password Schemes, *Proceedings of the 13th conference on USENIX Security Symposium-Volume 13*, page 11

De Angeli, A., Coutts, M., Coventry, L., Johnson, G., Cameron, D. and Fischer, M. 2002. VIP: a visual approach to user authentication. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pages 316–323

De Angeli, A., Coventry, C., Johnson, G. and Renaud, K., 2005. Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies*, 63(1-2), pp128–152

Dhamija, R., and Perrig, A., 2000. Deja vu: A User Study Using Images for Authentication. In *Proceedings of the 9th conference on USENIX Security Symposium-Volume 9*, pp 45–48

English, R. and Poet, R. 2011. Measuring the Revised Guessability of Graphical Passwords, *5th International Conference on Network and System Security*, pp.364-368

Hlywa, M., Biddle, R., and Patrick, A. 2011. Facing the facts about image type in recognition-based graphical passwords. *Proceedings of the 27th Annual Computer Security Applications Conference*, volume 36, pages 149–158

Sasamoto, H., Christin, N., Hayashi, E. 2008. Undercover: authentication usable in front of prying eyes. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 183–19

Suo, X., Zhu, Y., and Owen, G. 2006. Analysis and Design of Graphical Password Techniques. *Advances in Visual Computing*, pages 741–749

Takada, T., Onuki, T., and Koike, H. 2006. Awase-e: Recognition-based image authentication scheme using users' personal photographs. In *Innovations in Information Technology,* pages 1–5

Tari, F., Ozok, A., and Holden, S. 2006. A Comparison of Perceived and Real Shoulder-surfing Risks between Alphanumeric and Graphical Passwords. In *Proceedings of the second symposium on Usable privacy and security - SOUPS '06*, page 56

Tullis, T., and Tedesco, D. 2005. Using personal photos as pictorial passwords. *CHI '05 extended abstracts on Human factors in computing systems*, page 1841

Wiedenbeck, S., Waters, J., Sobrado, L., and Birget, J. 2006. Design and Evaluation of a Shoulder-surfing Resistant Graphical Password Scheme, *Proceedings of the working conference on Advanced visual interfaces*, page 177