Towards a Metric for Recognition-Based Graphical Password Security

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*—Recognition-based graphical password (RBGP) schemes are not easily compared in terms of security. Current research uses many different measures which results in confusion as to whether RBGP schemes are secure against guessing and capture attacks. If it were possible to measure all RBGP schemes in a common way it would provide an easy comparison between them, allowing selection of the most secure design. This paper presents a discussion of potential attacks against recognition-based graphical password (RBGP) authentication schemes. As a result of this examination a preliminary measure of the security of a recognition-based scheme is presented. The security measure is a 4-tuple based on distractor selection, shoulder surfing, intersection and replay attacks. It is aimed to be an initial proposal and is designed in a way which is extensible and adjustable as further research in the area develops. Finally, an example is provided by application to the PassFaces scheme.**

I. Introduction

Password authentication is failing as an authentication mechanism due to lack of memorability and security [1], [10] and graphical passwords are presented as an alternative [3], [7]. In particular, with a recognition-based graphical password (RBGP) user authentication scheme the user provides a username and selects (or is provided with) a number of "passimages" (term by [5]). Upon commencing the authentication process, the user provides their username and is presented with a challenge screen which contains one of their passimages along with a selection of alternative images called distractors. To successfully authenticate the user must recognise and select their passimages from the distractor images for a number of challenge screens. Whilst the potential memorability and usability of such schemes have been examined [6], [14] , analysis of their security is often limited to countermeasures for a specific attack. As such, it is unclear whether RBGP schemes provide a secure alternative to passwords. There remains no standardised method of measuring the level of security of a RBGP scheme. The contribution of this work is to combine literature on the security of RBGP schemes in a way which results in an extensible 4-tuple evaluation of the security of RBGP schemes. In the remainder of this paper we present a threat model for RBGP schemes, examine counter measures and propose a heuristic approach to evaluating security levels for RBGP schemes. The result follows that different RBGP schemes can be compared in a common manner.

II. Threat Model

Here we present a threat model for RBGP schemes, where the aim of the attacker is to obtain a users passimages and username. The threat model does not consider possible software exploits since such exploits would be very specific to an implementation and it would not be possible to generalise these for application to all RBGP schemes.

As categorised by DeAngeli *et al.* [6] the main attacks can be split into three distinct areas of concern; guessability, observability and recordability. Literature on RBGP schemes was examined and attacks were identified and categorised into the three areas. In addition to this, one further attack was identified by examination of the dictionary attack threat relevant to password authentication schemes. The following attacks were identified from literature and are considered in our threat model (as shown in Figure 1): shoulder surfing, intersection attacks, eavesdropping resulting in a replay attack, phishing and social engineering tactics for guessing. The remaining attack, ordered guessing, was identified by comparison with dictionary attacks and is also included in our threat model.

III. Security Analysis

In order to construct a heuristic model for evaluating the security of RBGP we take the following approach.

1) Examine the relevant literature and extract the possible counter measures to the specific attacks presented in the threat model.
2) Abstract these counter measures from specific implementations to general approaches.
3) Construct a series of key questions for each attack
4) Construct flow charts which combine the key questions for each attack. Questions regarding the set up of a RBGP scheme will be asked and if counter measures which result in a more secure scheme against that specific attack are implemented, then the security measure will be increased. Where no counter measures or an otherwise insecure setup is used, the level of security for that scheme will be reduced.

Each area of concern and related attacks from our threat model shown in Figure 1 are discussed in the remainder of this section which covers points 1 to 3 for each potential attack (where possible).

Fig. 1. Threat Model

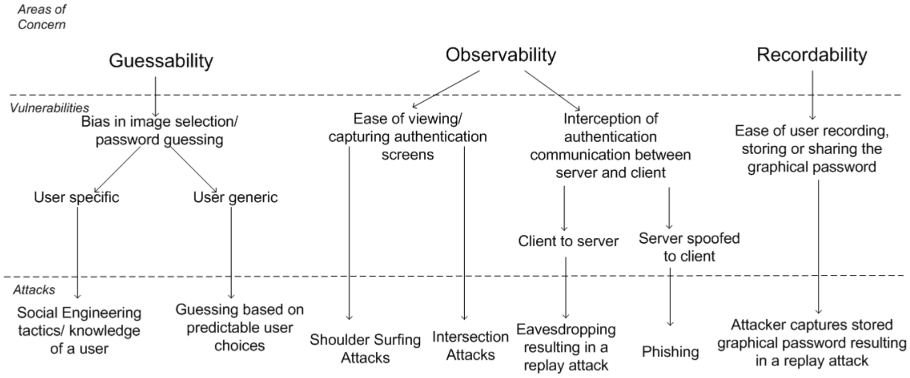
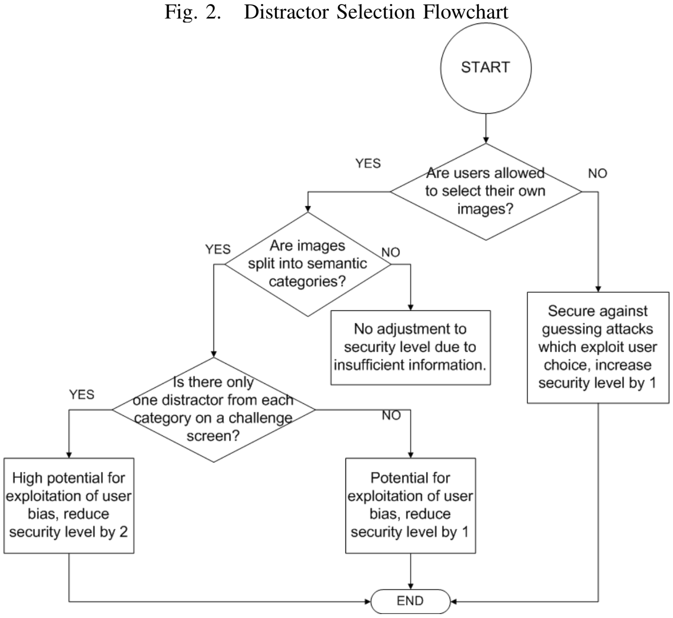

Areas of Concern
Guessability
Observability
Recordability
Vulnerabilities
Bias in image selection/ password guessing
User specific
User generic
Ease of viewing/ capturing authentication screens
Interception of authentication communication between server and client
Client to server
Server spoofed to client
Ease of user recording, storing or sharing the graphical password
Attacks
Social Engineering tactics/ knowledge of a user
Guessing based on predictable user choices
Shoulder Surfing Attacks
Intersection Attacks
Eavesdropping resulting in a replay attack
Phishing
Attacker captures stored graphical password resulting in a replay attack

Fig. 2. Distractor Selection Flowchart


START
Are users allowed to select their own images?
YES
NO
Are images split into semantic categories?
YES
NO
No adjustment to security level due to insufficient information.
Secure against guessing attacks which exploit user choice, increase security level by 1
Is there only one distractor from each category on a challenge screen?
YES
NO
High potential for exploitation of user bias, reduce security level by 2
Potential for exploitation of user bias, reduce security level by 1
END

A. *Guessability: User Specific - Social Engineering Tactics*

Social engineering is the process of exploiting information about or provided by users in order to attack a system [15]. It exploits users in a non-technical approach. An educated guess approach is highlighted in [11] where they use an oil filter applied to images to avoid such attacks. As the authors highlight in their future work section, an experiment to examine the effectiveness of the proposed countermeasure is yet to be reported. Another approach which removes the guessability issue completely is to assign images to a user, this means their choice cannot be exploited for an attack. However, this could interfere with memorability. Overall, this area requires further research. In particular experiments should be conducted to determine how easily images are guessed. This is proposed in the future work section.

B. *Guessability: User Generic - Ordered Guessing Attacks*

When considering alphanumerical passwords, an offline dictionary attack is common and involves constructing a list of words and trying each member of the list until a matching password is found [4]. In the context of an offline dictionary attack for recognition-based graphical password scheme, this type of attack appears infeasible since the passimage is known to be on the screen and trying all possibilities until a matching image is found is not applicable in this instance. It follows that a traditional offline dictionary attack would not be applicable to a recognition-based scheme. However, we propose that there is a possibility for an online guessing attack which prioritises more commonly selected images.

We have conducted studies into exploring the feasibility of these “prioritised guessing attacks”, in which the attacker selects the “most probable” image given the challenge screen presented with the purpose of increasing the probability of selecting the correct image. This was conducted by collecting a graphical passimage set of four images for 64 users. Each image belonged to one of twelve distinct categories. The number of selections in each category and ordering the categories from most to least probable. An attack was then launched by constructing a challenge screen for each possible passimage and choosing the image from most likely category given the ordering noted. It became apparent 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. For more information on the details of this study, please see [9].

The evaluation of distractor selection is presented in Figure 2, a flow chart which combines a number of questions relating to the set up of a RBGP scheme which coincide with the results of the research determining how these answers affect security against a prioritised guessing attack. If users are not allowed to select their own images, no bias can be exploited in a guessing attack (assuming random assignment of passimages) and in this case the security measure is increased by one. If users can select their own images from a collection which cannot be categorised into semantic categories, then there is insufficient information to analyse potential security impact. In this case no adjustment is made to the security measure. If users select their own images from a collection which can be categorised into distinct semantic categories and the distractors are selected entirely randomly, then guessability can be decreased by construction of a semantic ordered guessing attack and the measure is reduced by one to reflect this. If distractors are selected from distinct categories other than the category of the passimage, guessability can be decreased by construction of a semantic ordered guessing attack and the measure is reduced by two to reflect this. If however the distractors are not selected from distinct categories, then the reduction is less and the security level is reduced by only one.

C. Observability:Shoulder Surfing

In [17], the authors propose that counter measures for a shoulder surfing attack can be placed under two categories; using no indicators of passimage selection or disguising indicators of passimage selection . We extend this categorisation by inclusion of an additional approach taken by DeAngeli *et al.* [6] and Dunphy *et al.* [8]. In this approach, a "key image portfolio" concept for security against observation attacks is used where the user selects a number of passimages which exceeds the number of challenge screens presented. This means that in any authentication session, a subset of the users passimages are presented.

There are a number of specific counter measures for shoulder surfing. An example of a scheme where no indicators are shown include the work in [18] where a convex-hull approach is taken. In this approach the user is required to click any point in a convex-hull which is created from mentally joining the users passimages shown on screen. An example of a scheme were image selection is disguised is presented in [11] where passimages have an oil painting filter applied to disguise the image .

We now highlight several key issues when considering construction of a RBGP scheme which is resistant to shoulder surfing. If the scheme does not provide any indication of a passimage being selected (e.g. by border, or highlighting) then the scheme is very secure against shoulder surfing and the security metric can be increased to reflect this. If there is an indication of which image has been selected, but this has been disguised then the scheme is similarly secure. If however there is no attempt to disguise image selection but the scheme implements a "key image portfolio" as in [6] and [8] then the scheme reduces potential for shoulder-surfing, but does not eliminate it completely and hence the metric can be incremented, but not by the same value as being completely secure against shoulder surfing.

If there is an indication of the image selected, with no attempt to disguise this selection and no implementation of a key image portfolio, then the scheme is almost definitely insecure against shoulder surfing and the measure is decremented to reflect this. These results are summarised and combined into the flowchart shown in Figure 3.

Fig. 3. Shoulder Surfing Flowchart

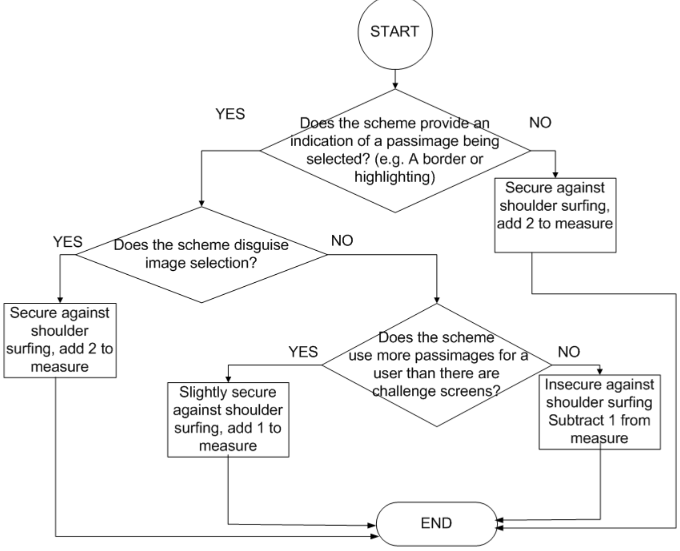
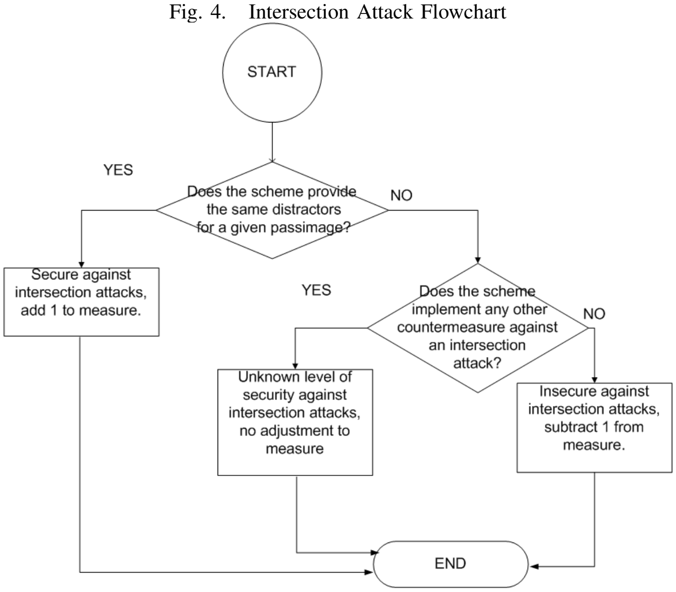

START
Does the scheme provide an indication of a passimage being selected? (e.g. A border or highlighting)
YES
NO
Secure against shoulder surfing, add 2 to measure
Does the scheme disguise image selection?
YES
NO
Secure against shoulder surfing, add 2 to measure
Does the scheme use more passimages for a user than there are challenge screens?
YES
NO
Slightly secure against shoulder surfing, add 1 to measure
Insecure against shoulder surfing Subtract 1 from measure
END

Fig. 4. Intersection Attack Flowchart


START
Does the scheme provide the same distractors for a given passimage?
YES
NO
Secure against intersection attacks, add 1 to measure.
Does the scheme implement any other countermeasure against an intersection attack?
YES
NO
Unknown level of security against intersection attacks, no adjustment to measure
Insecure against intersection attacks, subtract 1 from measure.
END

D. Observability: Intersection Attack

An intersection attack, as defined by Dhamija *et al.* [7] (and discussed in [8], [13]) is an attack in which the attacker records multiple challenge screens and notes the images which occur with high frequency. The authors successfully summarise general approaches to counter measures for this issue as:

- Use the same decoy images and pass images for each session.
- Use a small subset of decoy images for each session.
- Present a number of challenge screens which, if a user fails at one stage, subsequent screens only display distractor images. This means an attacker cannot impersonate a user in order to discover their complete set of pass pictures.
- Implement a limit on the number of incorrect authentications a user can perform, this stops an impersonator

attempting to discover all of the images. (A "three strikes and you're out" approach)

Maintaining the same distractors for a given passimage does ensure that an intersection attack is not possible. The remaining three options only serve to minimise the potential for an intersection attack.

Figure 4 shows the flow chart for this attack. Where the same distractor images are used for a given passimage the security is increased as this stops a possible intersection attack. Where one of the remaining mitigation techniques are implemented, no adjustment to the level of security is made since the possibility of an intersection attack is still present, even though this risk is reduced. If no attempt is made to mitigate or stop an intersection attack, the level of security is reduced.

E. Recordability: Eavesdropping resulting in a Replay Attack

Application of a replay attack to a recognition-based graphical authentication scheme is similar to considering a replay attack on an alphanumerical password scheme. A man-in-the-middle attack is constructed and the data copied when a user performs the authentication process and sends the authentication data to the server. The data is then "replayed" to the server at another time, potentially resulting in a false positive authentication. Discussion of this type of attack in terms of RBGP schemes is limited. Ku *et al.* [12] implement a remote user authentication protocol which sends a hash based on the time stamp (essentially performing as a salt value), code from the passimage (in this case a drawing) and other components. However, Ku's work is based on the Draw a Secret scheme and not on a RBGP scheme.

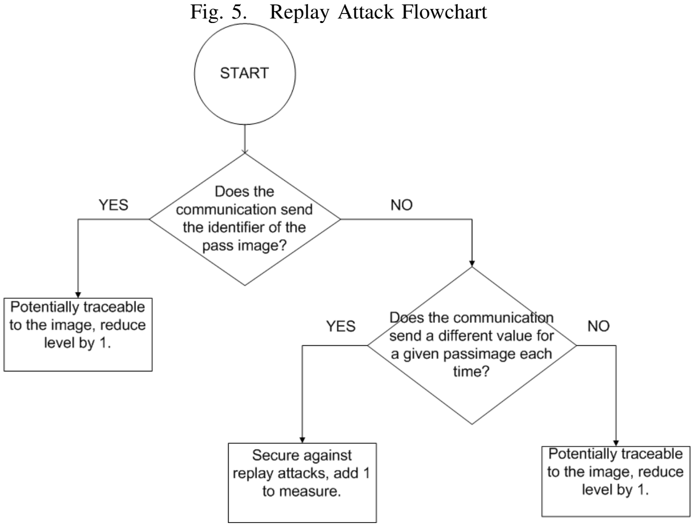
In addition to this approach of mitigating a replay attack, we propose a random assignment of location of the passimage on the challenge screen presented to the user, it should be this position which is sent back to the server to authenticate.

The key aspect of both these approaches is sending a different value each time authentication occurs, which results in avoiding a replay attack. This is incorporated into a flowchart presented in Figure 5 . If the communication between client and server sends the pass image identifier then security is reduced, if this is not the case but a different value is sent each time the security level is increased. In the final option, the passimage identifier is not sent, but it is the same value each time results in a value which may be traceable to the image and so the security level is reduced. With careful design, a replay attack can be completely avoided. However, not all aspects of the design are mentioned when new schemes are proposed and thus the assumption has been made that a replay attack may not have been considered when designing a scheme.

F. Recordability: Phishing

As noted by Biddle *et. al* in [2], a phishing attack is relatively easily applied to a RBGP scheme. In this situation, the attack would be carried out in real time. A man-in-the-middle attack would be performed, and the user's username is sent to the attacker who then uses this in the legitimate site to obtain a valid challenge screen, which is relayed back to the subject of the attack. The process is repeated until authentication has completed. At the moment, there is insufficient research to produce a method which takes onus off the user (since users appear not to take notice of security indicators such as the use of SSL [16]) and thus the best recommendation currently is to implement SSL and hope the user takes notice.

Fig. 5. Replay Attack Flowchart


START
Does the communication send the identifier of the pass image?
YES
NO
Potentially traceable to the image, reduce level by 1.
Does the communication send a different value for a given passimage each time?
YES
NO
Secure against replay attacks, add 1 to measure.
Potentially traceable to the image, reduce level by 1.

IV. Heuristic Model for Security Evaluation

The main result of this paper is the heuristic model for security evaluation, this model arose from the flowcharts shown in Figures 2, 3, 4, and 5 which covered ordered guessing, shoulder surfing, intersection and replay attacks. Each factor (distractor selection, shoulder surfing, intersection and replay) provides one score of a 4-tuple, where the score starts at 0 and increases or decreases depending on the route taken through the corresponding flow chart. The resulting tuple represents the security of a RBGP scheme in terms of these aspects, that is {distractor score, shoulder surfing score, intersection score, replay score}. For example a scheme which is the most insecure in terms of the model presented would have a 4-tuple of {-2,-1,-1,-1} whilst the most secure setup would result in a 4-tuple of {1,2,1,1}.

In general, where secure set up is implemented scores are increased by one, where no countermeasures were implemented scores were decremented by one. In shoulder surfing (Figure 3), there is the possibility to increase the level of security by two scales. This is due to the perceived significance (identified by the quantity of research in this area) of the threat caused by shoulder surfing and also due to the graded levels of counter measures available as described in Section III-C. In addition, there are two instances where no adjustment to the security level is possible. The first case is in distractor selection for an ordered guessing attack (shown in Figure 2), where a non-random distractor selection algorithm is implemented.

In this case it is not possible to determine how successful an ordered guessing attack might be without further analysis. The second instance is when considering intersection attacks, where counter measures which mitigate (but do not remove the possibility of) intersection attacks. The model is designed in such a way that it is extensible. This means that once further research is conducted, the results can be incorporated to the initial framework providing a more complete measure of security levels for RBGP schemes.

V. Example - Application to PassFaces Scheme

In order to provide a concrete example, the PassFaces scheme (www.realuser.com) was examined. From reviewing the white paper available at http://www.realuser.com/published/TheScienceBehindPassfaces.pdf the following information on the setup of the scheme was extracted:

- in general, four passfaces are assigned to a user and to authenticate users must identify their passfaces from four challenge screens each showing a passface and eight decoy faces
- order of faces on the screen is random
- no grid contains faces from the other grids and these faces are similar in appearance to the passface so that no one face stands out
- the same decoys are used each time for a given passface
- the option is provided to use keypad entry of the passface
- a "mask" is applied to the faces after selection

From this information it was possible to rule out guessing attacks completely due to the random assignment of passfaces, which gives a secure guessability factor of 1. It was not possible to consider eavesdropping or phishing attacks, nor was it possible to consider replay attacks. However, there was sufficient information to examine shoulder surfing and intersection attacks. In terms of shoulder surfing the mask application after selection meant image selection was disguised, resulting in a shoulder surfing security factor of 2. If there was no masking, the alternate keypad entry would result in a shoulder surfing factor of 1, but the highest of the two possibilities was considered appropriate. Since the scheme provides the same decoys (or distractors) for a passface for a given user, the scheme has an intersection security factor of 1. This results in a security 4-tuple of $\{1,2,1,X\}$ where X represents the unknown resistance to replay attacks.

VI. Conclusions and Future Work

The work presented here represents a step towards providing a unified measure of the security of RBGP schemes. The approach taken analysed potential attacks and counter measures and as a result constructed a series of key questions in order to determine resistance to these attacks. The attacks included were ordered/prioritised guessing,shoulder surfing, intersection attacks, and replay attacks. Analysis of each aspect resulted in a flow chart which provided one score of a 4-tuple. The score for each factor starts at 0 and increases or decreases depending on the route taken through the corresponding flow chart.The resulting tuple represents the security of a RBGP scheme in terms of these factors. The area requires further research, in particular research is proposed in the areas of social engineering for a known user where users select personal images for their passimages. The problem of recording a passimage (using a camera or screenshot) also needs to be addressed. The model presented is extensible to include further work as this would add extra flowcharts for any additional attacks analysis which can be added to the evaluation and also additional information can be added for any given attack already summarised.

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 Paul C. van Oorschot. Graphical Passwords: Learning from The First Generation. In *Technical Report TR-09-09, School of Computer Science, Carleton University*, pages 1–20. Carleton University, 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] Mark Burnett and Dave Kleiman. *Perfect Passwords*, chapter 2, page 17. Syngress, 2006.

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

[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] Rachna Dhamija and Adrian Perrig. Deja vu: A User Study Using Images for Authentication. In *Proceedings of the 9th conference on USENIX Security Symposium-Volume 9*, page 4. USENIX Association, 2000.

[8] Paul Dunphy, Andreas P. Heiner, and N. Asokan. A closer look at recognition-based graphical passwords on mobile devices. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, pages 1–12. ACM, 2010.

[9] Rosanne English and Ron Poet. Measuring the Revised Guessability of Graphical Passwords. In *NSS 2011- To appear*, 2011.

[10] Dinei Florencio and Cormac Herley. A large-scale study of web password habits. *Proceedings of the 16th international conference on World Wide Web - WWW '07*, page 657, 2007.

[11] Eiji Hayashi, Nicolas Christin, and Adrian Perrig. Use Your Illusion : Secure Authentication Usable Anywhere Categories and Subject Descriptors. *ACM International Conference Proceeding Series*, 337(Proceedings of the 4th symposium on Usable privacy and security):35–45, 2008.

[12] Wei-chi Ku and Maw-jinn Tsaur. A Remote User Authentication Scheme Using Strong Graphical Passwords. In *The IEEE Conference on Local Computer Networks 30th Anniversary (LCN'05)*, pages 351–357. Ieee, 2005.

[13] 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.

[14] Karen Renaud. Guidelines for designing graphical authentication mechanism interfaces. *Int. J. Information and Computer Security*, 3(1):60–85, 2009.

[15] David Salomon. *Foundations of Computer Security*, chapter 8, page 205. Springer, 2006.

[16] Stuart E. Schechter, Rachna Dhamija, Andy Ozment, and Ian Fischer. Emperors new security indicators: An evaluation of website authentication and the effect of role playing on usability studies. In *In Proceedings of the 2007 IEEE Symposium on Security and Privacy*, 2007.

[17] Hai Tao and Carlisle Adams. Pass-Go: A proposal to improve the usability of graphical passwords. *International Journal of Network Security*, 7(2):273–292, 2008.

[18] Susan Wiedenbeck, Jim Waters, Leonardo Sobrado, and Jean-Camille Birget. Design and Evaluation of a Shoulder-surfing Resistant Graphical Password Scheme. *Proceedings of the working conference on Advanced visual interfaces - AVI '06*, page 177, 2006.