Educated Guess on Graphical Authentication Schemes: Vulnerabilities and Countermeasures

Eiji Hayashi
Human-Computer
Interaction Institute
Carnegie Mellon University

Jason I. Hong
Human-Computer
Interaction Institute
Carnegie Mellon University

Nicolas Christin
Information Network Institute
Carnegie Mellon University

ABSTRACT

This paper investigate security of graphical authentication tokens against educated guess attacks. Results of two user studies indicate that, if we use original photos as authentication tokens, the authentication tokens are vulnerable to educated guess attacks. The results also demonstrate that we can mitigate the vulnerability using distorted pictures.

1. INTRODUCTION

Educated guess attacks are attacks against authentication tokens (e.g., password), where an attacker try to guess the authentication tokens based on informaiton about users. In particular, educated guess attacks can be categorized into two types. In the first type of educated guess, an attacker exploits knowledge common to many users. For instance, if the attacker knows that users frequently choose the word "password" as their actual password, trying "password" has a higher chance of success than trying a random sequence of letters in guessing a user's password. In the context of password-based authentication systems, this type of educated guess is generally called a *dictionary attack*.

A different type of the educated guessing attacks relies on knowledge of information specific to a given user. For instance, if an attacker knows that a user has a wife named Pam, the attacker can guess that "pam," or slight variations on these strings, have a considerably higher probability of being the user's password than other possible sequences. By analogy with the terminology used to characterize variants of phishing attacks, we call this latter form of educated guessing the *spear educated guess attack*.

One possible solution for the educated guess attackes is to improve memorability of self-chosen passwords by using image-based authentication schemes. However, educated guesses are still one of the foreseeable attack vectors.

Because graphical authentication tokens preserve more information – such as semantic meanings of an image – than text-based authentication tokens do, graphical authentication could actually be *less* resilient against spear educated guess attacks than their text-based counterparts. The first research question we address in this paper is to quantify the extent to which graphical authentication tokens are resilient against (or vulnerable to) spear educated guess attacks.

We further consider a graphical variant of the dictionary attack, where an attacker tries to guess a user's authentication tokens based on general knowledge about users' preferences. Specifically, if people have a bias toward a specific types of images, such as photos of faces, they choose as authentication tokens, then an attacker may use this bias to improve his chances of guessing the right authentication tokens. The second research question we address in this paper is to prove or disprove the existence of biases toward certain types of images in graphical authentication tokens chosen by users; and to characterize such biases if they do exist.

One countermeasure for the the educated guess attacks against graphical authentication schemes is the use of distortion [2, 3]. Hayashi et al. showed high memorability of the distorted pictures [3] and claim that their system is therefore more secure; however no study thus far quantifies how much added security distortion provides. Thus, the third research question is to what extent the distortion technique improve security of image-based authentication schemes.

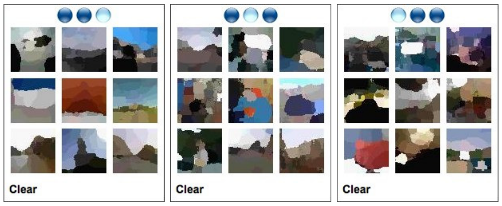

Figure 1: Screenshots of an authentication system used in the first user study. A user is asked to choose pre-determined three images from 27 images to be authenticated. In the first user study, we ask a participants to guess his friend's graphical authentication tokens using this interface to quantify vulnerability of graphical authentication tokens against educated guess attacks and impact of distortion on the attacks.

2. USER STUDY

In order to investigate the three research questions, we conduct two user studies.

In the first user study, we recruit "friend pairs" as participants. We define "friend pair" as two persons who meet three times a week and register each other as friends in Facebook. Then, we ask friend pairs to guess their friends' graphical authentication tokens. In other words, we asked participants to be mock attackers and to guess authentication tokens of targets whom they know well, i.e., spear educated guess attack. The target can be either his friend or a stranger. Additionally, the authentication tokens can be original photos or distorted photos. Table 1 shows the success rates of the mock attackers in four conditions.

As Table 1 shows, in the original conditions, the success rates are higher than that of the random guess. These results demonstrate that if we use original photos as authentication tokens, attackers can guess the authentication tokens at a higher success rate than random guesses. Especially, the high success rate in friend-original condition indicates that graphical authentication tokens are vulnerable to the spear educated guess attack. On the other hand, in the distorted conditions, there is no statistical significance between the success rates and the success rate of random guesses ($p > 0.01$).

	Friend	Stranger
Original	0.53* (8)	0.20* (3)
Distorted	0.067 (1)	0.00 (0)

Table 1: Success Rates of Educated Guesses. Asterisks "*" denote that the success rates are higher (statistically significant) than success rate of random guesses. Numbers in parentheses stands for numbers of mock attackers out of 15 who made correct guesses.

This demonstrates that distorted pictures are more resilient againt the educated guess attacks than origianl pictures.

In the second user study, we test whether there are biases toward certain categories among pictures chosen by users as their authentication tokens.

We ask participants to classify two sets of pictures into 12 categories according to semantic meanings of the images. The first set consists of 120 photos randomly chosen from Flickr.com. We used a Perl script to select the 120 photos from all photos posted on Flickr under Creative Commons license [1]. The other set consists of 180 photos chosen as graphical authentication tokens by the participants in the first user study. By comparing how the photos in these two sets are categorized, we can evaluate whether photos chosen by the participants are biased toward certain categories.

We use Amazon Mechanical Turk to categorize images according to their semantic meanings into pre-defined 12 categories. Table 2 shows ratios photos categorized into the 12 categories. The ratio $p_i, (i = 1, 2, 3..., 12)$ is calculated as $p_i = c_i/N$ where c_i and N denote the number of photos categorized to ith category and the total number of photos.

Category	Random choice		Users' choice	
Human	0.26	(153)	0.26	(251)
Transportation	0.06	(38)	0.02*	(21)
Animal	0.07	(39)	0.05	(38)
Insect	0.00	(0)	0.00	(1)
Interior	0.10	(60)	0.04	(35)
Landscape	0.10	(58)	0.15*	(127)
Plant	0.01	(3)	0.03*	(24)
Building	0.11	(68)	0.25*	(218)
Food	0.09	(51)	0.04*	(34)
Clothing	0.00	(2)	0.00	(1)
Object	0.17	(101)	0.13	(122)
None of them	0.04	(26)	0.03	(28)
Total	1.00	(600)	1.00	(900)

Table 2: Distribution of categories given to original images. The numbers in parentheses stand for actual number of category tags given. Asterisks "*" denotes that the categories have statistically significant differences in the ratios between random choice and users' choice.

In Table 2, the numbers in random choice column shows categorization of photos people *take*. On the other hand, the numbers in normal users' choice column shows categorization of photos participants *choose* as their authentication tokens. The differences between numbers in these two columns indicates participants' tendency to choose or not to choose categories of images as their authentication tokens.

Table 2 shows that the participants are less likely to choose photos of transportation and food as their authentication tokens. On the other hand, they are more likely to choose photos of landscapes, plants and buildings as their authentication tokens. This results indicates that there are biases in user-chosen authentication tokens.

On the contrary, when we ask Mechanical Turk users to categorize distorted pictures, there is no statistically significant difference in the ratios between randomly chosen images and user chosen images.

This observation implies that an attacker can launch dictionary attacks against graphical authentication tokens without ditortion, and that distortion mitigate the vulnerability.

3. DISCUSSION

The results obtained in this paper yield a number of interesting implications.

First, there is an inevitable trade-off between the security and the memorability of graphical authentication tokens. Our results imply that semantic meanings, which authentication tokens have, help an attacker to make better guesses about the authentication tokens, while the semantic meanings also help a user to memorize them more easily.

Second, the results show that the semantic meanings of distorted photos are difficult to guess without knowing their original photos. Thus, by using distortion, we can make information asymmetry between a user and an attacker. This asymmetry gives a user a clear advantage over an attacker.

Finally, the results can be applied to other types of graphical authentications. For instance, secret pictures are being used in online banking sites for users to verify that the web sites are legitimate. Although we need further investigation, we would be able to use a distorted photo as a secret picture to make it difficult to guess for an attacker, while keeping the secret picture easy to memorize for a user.

4. CONCLUSION

In this paper, we evaluate the security of graphical authentication tokens against the educated guess attacks. Findings discussed in this paper provide evidence that graphical authentication tokens have the same challenges as character-based authentication tokens, e.g., password. When we allow users to choose their "password", they tend to choose weak ones. However, at the same time, the findings suggest that we can address this problem using distortion. By using distorted photos, we can make graphical authentication tokens more secure against the educated guess attack, while keeping their memorability as high as original photos [3].

More generally, we hope that this paper suggests a new direction to improve the security of graphical authentications tokens by relying on the advantages of the human cognitive process rather than addressing its memorability limitation by using graphics.

5. REFERENCES

[1] Creative Commons. http://creativecommons.org/.

[2] A. Harada, T. Isarida, T. Mizuno, and M. Nishigaki. A user authentication system using schema of visual memory. In *Proceedings of BioADIT'06*, pages 338–345, Osaka, Japan, January 2006.

[3] E. Hayashi, R. Dhamija, N. Christin, and A. Perrig. Use your illusion: secure authentication usable anywhere. In *Proceedings of Usable privacy and security*, August 2008.