Using Personal Photos as Pictorial Passwords

Thomas S. Tullis and Donna P. Tedesco
Fidelity Investments
82 Devonshire St., V4A, Boston, MA 02109 USA
[tom.tullis, donna.tedesco]@fidelity.com

ABSTRACT

Pictorial passwords, where the user recognizes “target” images among “distractors”, appear to have potential for improving the usability of authentication systems. We conducted three exploratory studies on the use of personal photos for authentication over a three-month period. Participants provided 8-20 photos of personal significance to them but which they believed others would not recognize. They also chose four photos to remember from a set of stock photos. Recognition accuracy for the personal photos was significantly higher than the stock photos. We also manipulated the number of target and distractor photos as well as their similarity, and we tested how well others who know the users could guess their photos. Larger numbers of distractors and greater similarity to the targets made it harder for others to guess the correct photos, while having no impact on the user’s own recognition accuracy.

Author Keywords

User authentication, pictorial passwords, photos, photographs.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Password memory and security in user authentication has long been a concern in the computing industry. When choosing passwords, users tend to choose very easy, memorable passwords that can often be guessed [5,6]. When given meaningless strings of passwords by the system, users are often unable to recall them, resulting in help-desk calls and the costs of resetting passwords. With all of the passwords we use daily, users even forget the easy passwords that they’ve chosen for themselves—unless they use the same password for every system or write their passwords down, both of which are a security risk.

A variety of studies have investigated the viability of alternative methods for user authentication, including biometrics [1], doodling [4], inkblots [7], and images [2,3,7,8,11]. Our efforts specifically focus on the use of images for authentication, or “pictorial passwords”. This has been a popular topic for investigation, as psychological literature shows that images are much more easily learned and recognized than words [9,10]. Dhamija [3] studied the use of actual photos as well as “visual hashes,” or unstructured, random images derived from data strings. She found that users recognized photos better than the visual hashes and preferred to choose photos that had more personal meaning. Davis, Monrose, and Reiter [2] studied the use of pictures of faces as passwords, similar to the method used in a commercial system called Passfaces™[8]. They found that faces that users chose were highly correlated with their race and gender, almost to the point of being predictable and easily guessed. Weinshall and Kirkpatrick [11] used pictures of random items for authentication, where the system assigned a number of pictures to the user for memorization. They found a high level of accuracy when pictures had clear themes and individual distinctions, and when training and testing were more frequent.

Requiring users to remember a set of assigned pictures and training with them can be more laborious than learning or creating a textual password. As an alternative, we decided to study a method where users would provide their own personal photos as “targets” for authentication. Personal photos are highly meaningful and unforgettable to users, as well as fun to view. However, a concern with this method is “break-ins” by imposters who are somewhat familiar with the user. The overall goals of this series of studies were to assess recognition accuracy for personal photos over a period of months, compare that to the recognition of stock photos, and identify ways of minimizing the likelihood that others could guess a user’s photos.

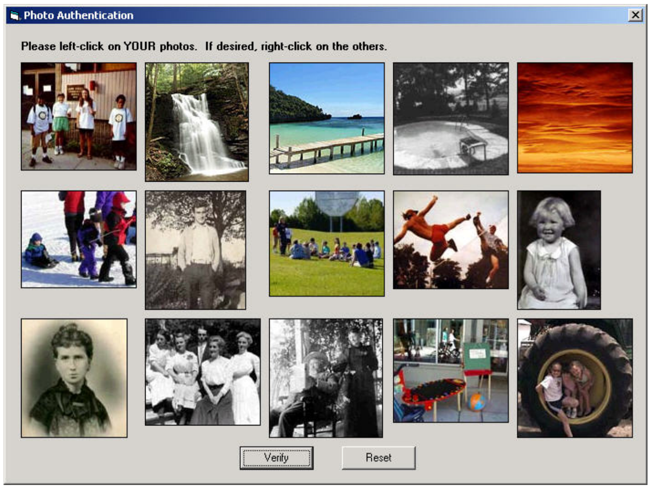
STUDY 1

Fourteen employees of our company participated in the first study. Prior to the study, each was asked to provide 8-20 photos of personal significance to them but which they believed others would not be able to readily associate with them (e.g., baby pictures, old family photos, favorite vacation photos, pets). After receiving all of the photos from the participants, we selected about 300 additional photos found through a variety of sources that seemed to be of reasonably similar subject matter for use as the pool of distractors. All photos were then reduced to thumbnail size (125 pixels maximum in either dimension).

CHI 2005, April 2–7, 2005, Portland, Oregon, USA.
ACM 1-59593-002-7/05/0004.

The Prototype

The prototype authentication system first prompted the user for a User ID, since the system needs to know who this person *claims* to be. Users were then shown a screen (Figure 1) containing a total of 15 photos in random order. The number of targets in each case was randomly selected between 2 and 5. The user's task was to click on their photos, in any order. Whenever the user clicked on *any* photo, it disappeared. After clicking on all the desired photos, the user clicked on the "Verify" button. Users got up to three incorrect attempts before being locked out.



Please left-click on YOUR photos. If desired, right-click on the others.

Figure 1. Sample screen from Study 1.

In building the first set of photos to show a user, a "distractor pool" of the same size as the "target pool" was created. So, for example, if the user had 10 targets in their pool, a pool of 10 distractors was built. If the user was unsuccessful on their first attempt, subsequent attempts built the array of 15 photos by choosing from these pools created on the first attempt. In this way, repeated photos were just as likely to be a target as a distractor.

Method

The study was conducted in our Usability Lab. The procedure for each participant was as follows:

1. Each participant performed about eight trials where they "logged in" by recognizing their own photos.
2. Each participant was given a list of all the participants in the study (who were mostly from the same department) and asked to pick two others that they know from the list and to try "breaking in" as them.
3. Each participant was then shown a categorized collection of about 2,000 stock photos (e.g., famous people, animals, famous paintings) and asked to pick at least 4 photos from that collection to remember.
4. Each participant then performed about eight trials where they "logged in" using the stock photos they had just chosen. The pool of distractors was stock photos from the same categories as the ones they had just seen.

Results

Personal Photos

When logging in using the personal photos they had provided, all participants were successful. Out of 112 individual trials, users made errors on only 5 (4%). These errors were all on first attempts; they then succeeded on the second attempt. Average time per trial was 11.2 seconds. No users succeeded in "breaking in", after three attempts, as one of the other two users they selected. The following formula, expressed as a percentage, was used for calculating the accuracy of each attempt:

Accuracy = (# of Targets Clicked) /
(Total # of Targets + # of Distractors clicked)

For example, if there were 4 targets, and the user clicked on 3 of them plus 1 distractor, that would be an accuracy of 3/(4+1) = 3/5 or 60%. This break-in accuracy averaged 29%, ranging from 8% to 50% per participant. The highest accuracy on any one attempt was 80%.

Stock Photos

All participants successfully logged in using the stock photos they had just chosen. Out of 112 individual trials, users made errors on only 7 (6%). These errors were all on first attempts; they then succeeded on the second attempt. Average time per trial was 10.3 seconds.

Subjective Ratings

User comments on the recognition of personal photos as a way of authenticating themselves were quite positive. Many users obviously enjoyed the experience. On post-test subjective ratings, participants rated both photo techniques as easier, more enjoyable, and more secure (in comparison to the traditional login technique), and one they would be very likely to use. Personal photos were rated significantly higher than Stock photos on most scales.

Discussion

The high accuracy rates when users logged in using their Personal photos were very encouraging. The high accuracy rates when users logged in using the Stock photos were also encouraging, but not surprising since they had just selected the photos a few minutes earlier. On the attempted "break-in" trials, however, some users came uncomfortably close to guessing another person's photos. Some of the participants in the study have worked together for over 10 years. In these cases, they were sometimes able to correctly guess a few of the other person's photos (e.g., "I know she likes to make quilts, so I'm going to guess that quilt is hers.").

STUDY 2

The primary goals of Study 2 were to assess recognition accuracy for both Personal and Stock photos after the passage of time, and to determine if the guessing accuracy from "break-in" trials found in Study 1 could be reduced.

Method

The same fourteen people from Study 1 participated in Study 2. Study 2 was conducted approximately 30 days after Study 1. Participants did not have access to the prototype during the intervening period. The following changes to the prototype were made for Study 2:

- The number of target photos shown on any given attempt was increased from 2-5 to 3-6.
- The pool of distractor photos was modified by adding all of the participants' target photos as potential distractors and by adding more distractors with a "snapshot" flavor. This increased the distractor pool to about 500 photos. Obviously, participants' own target photos were not displayed as distractors to them. This change was primarily a result of feedback in Study 1 that some of the distractor photos did not look enough like the targets (i.e., they looked "too professional").

The procedure for each participant was as follows:

1. Attempted about 8 trials logging in using their Personal photos.
2. Attempted to "break in" as two other participants in the study using Personal photos.
3. Attempted about 8 trials logging in using the Stock photos they had chosen in Study 1. (The pool of distractors for the Stock photos was the same as before.)
4. Attempted to "break in" as two other participants in the study using Stock photos.

Results

Personal Photos

When logging in using their Personal photos, all participants were successful. Out of 116 individual trials, they made errors on only 7 (6%). These errors were all on first attempts; they then succeeded on the second attempt. Average time per trial was 11.4 seconds. Much to our surprise, on the attempted "break in" trials, there were four instances where participants succeeded in guessing another participant's target photos. Across the 14 participants, the guessing accuracy, calculated using the same technique as in Study 1, averaged 38%, ranging from 18% to 50%. The highest accuracy on any one attempt was, of course, 100%.

Stock Photos

There were seven instances where participants failed to log in using their chosen Stock photos, even after three attempts. This happened to four participants at least once. One participant never succeeded. Out of 116 individual trials, they made errors on 33 (28%). Average time per trial was 13.3 seconds. No users succeeded in "breaking in" as one of the other users they selected using the Stock photos. The guessing accuracy averaged 23%, ranging from 11% to 43% per participant. The highest accuracy on any one attempt was 60%.

Discussion

Even after one month had elapsed since the previous study, recognition accuracy for participants' own Personal photos remained high. However, recognition accuracy for Stock photos after one month was disappointingly low, with several instances of participants failing to recognize their chosen photos even after three attempts. The other disappointment was that the guessing accuracy on the "break-in" trials using Personal photos actually increased relative to Study 1, from 29% to 38%. Obviously, the changes intended to decrease the guessing accuracy did not work.

STUDY 3

For Study 3, we decided to drop the Stock photos due to their disappointingly low recognition accuracy in Study 2. Consequently, the primary goals of Study 3 were to determine if the high levels of recognition accuracy for Personal photos would hold after another month elapsed, and to see if the guessing accuracy could be reduced.

Method

Twelve of the fourteen participants from Studies 1 and 2 participated in Study 3. Study 3 was conducted approximately 30 days after Study 2. The following changes were made to the prototype for Study 3:

- The total number of distractor photos was significantly increased by switching to a multi-screen approach. The user had to select one photo on each of 5 screens containing 16 photos each. As soon as the user clicked on a photo, the next screen was shown automatically. Feedback was given only after all five screens.
- The total number of distractors was increased again, up to 1,367. As in Study 2, we continued to add photos that had a "snapshot" or "amateur" look to them.
- Two conditions were introduced manipulating the similarity of the targets and distractors. The Random Distractors condition was the same as in Studies 1 and 2: the distractors were chosen at random for each target. In the Tailored Distractors condition, the distractors for a given target were selected to be similar to the target (Figure 3). For each target photo, we manually identified 16-25 photos from the distractor pool that were "similar" to the target photo in some way.

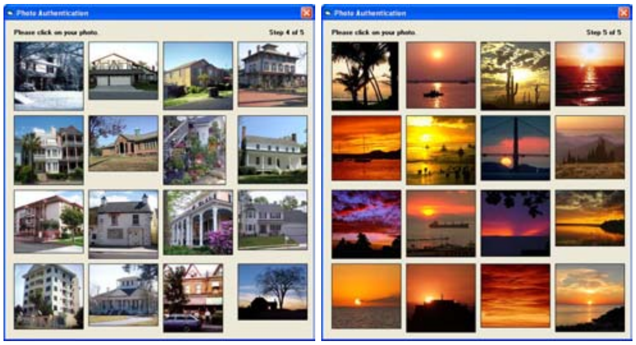


Figure 3. Examples of Tailored Distractors Condition

The procedure for each participant was as follows:

1. Attempted at least 5 trials logging in using their Personal photos shown among Random Distractors.
2. Attempted "breaking in" as two other participants in the study using Random Distractors.
3. Attempted at least 5 trials logging in using their Personal photos shown among Tailored Distractors.
4. Attempted "breaking in" as two other participants in the study using Tailored Distractors.

Results

Random Distractors

All participants were successful in logging in with their Personal photos shown among the Random Distractors. In fact, participants made *no* errors in the 71 trials. Average time per trial was 20.3 seconds. In the attempted "break-in" trials, there was one instance of a participant successfully guessing another participant's photos. Due to the different structure of the prototype, the "guessing accuracy" had to be calculated differently from Studies 1 and 2. In this case, since there were always 5 target photos, the accuracy was simply what percentage of those 5 photos were identified. The average guessing accuracy was 30%, ranging from 10% to 60% for each participant.

Tailored Distractors

All participants were successful in logging in with their Personal photos shown among the Tailored Distractors. As with the Random Distractors, there were no errors in the 72 trials. Average time per trial was 21.3 seconds. In the attempted "break-in" trials, there were two instances of participants successfully guessing another participant's photos. The average guessing accuracy was 20%, ranging from 7% to 47% for each participant. Guessing accuracy using the Tailored Distractors was significantly less than that using the Random Distractors (t-test, p=.03).

CONCLUSIONS AND FUTURE DIRECTIONS

The following conclusions appear to be warranted by these three studies:

1. Users can easily and accurately recognize their personal photos even after several months have elapsed and even among very similar distractors.
2. Users are significantly more accurate at recognizing their personal photos than stock photos that they selected. In addition, most users actually seem to enjoy viewing and choosing their personal photos from distractors.
3. Although we have reduced the levels of "guessing accuracy" through various techniques, the likelihood of someone else who is familiar with the user being able to accurately guess their personal photos is still too high to be acceptable. Ways of reducing that guessing accuracy further need to be explored.

The most promising technique appears to be presentation of the target photos among very similar distractors. Variations on this technique will be investigated in future studies. Some options to be considered include restricting the types of personal photos users can select (e.g., no pictures of people), manipulating the total number of targets and distractors, introducing the possibility of a target not being shown on a given screen, and investigating ways of having the user categorize their personal photos for matching with similar distractors.

REFERENCES

1. Coventry, L., Angeli, A., and Johnson, G. Usability and Biometric Verification at the ATM Interface. *Proc CHI 2003*, ACM Press (2003), 153-160.
2. Davis, D., Monrose, F., and Reitner, M. On User Choice in Graphical Password Schemes. *Proc. 13th USENIX Security Symposium*, USENIX Assoc. (2004), 1-14.
3. Dhamija, R. Hash Visualization in User Authentication. *Proc. CHI 2000*, ACM Press (2000), 279-280.
4. Goldberg, J., Hagman, J., and Sazawal, V. Doodling our Way to Better Authentication. *Proc. CHI 2002*, ACM Press (2002), 868-869.
5. Gong, L., Lomas, M.A., Needham, R.M., and Saltzer, J.H. Protecting Poorly Chosen Secrets from Guessing Attacks. *IEEE Journal on Selected Areas in Communications*, 11, 5 (1993), 648-656.
6. Klein, D. "Foiling the Cracker": A Survey of, and Improvements to, Password Security. *Proc. 2nd USENIX Workshop on Security* (1990), 1-11.
7. Pictures as Passwords. *The Economist*, *Sept. 16, 2004.* http://www.economist.com/science/tq/displayStory.cfm?story_id=3171359
8. Real User Technology and Products (2004). Passfaces™System, http://www.realuser.com/published/RealUserTechnologyAndProducts.pdf
9. Shepard, R.N. Recognition Memory for Words, Sentences, and Pictures. *Journal Verb Learn Verb Behav* 6 (1967), 156-163.
10. Standing, L., Conezio, J., and Haber, R.N. Perception and Memory for Pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, 19, 2 (1970), 73-74.
11. Weinshall, D. and Kirkpatrick, S. Passwords You'll Never Forget, but Can't Recall. *Proc. CHI 2004*, ACM Press (2004), 1399-1402.