Social Authentication: Harder than it Looks

Hyoungshick Kim, John Tang, and Ross Anderson

Computer Laboratory,
University of Cambridge, UK
{hk331, jkt27, rja14}@cam.ac.uk

Abstract. A number of web service firms have started to authenticate users via their social knowledge, such as whether they can identify friends from photos. We investigate attacks on such schemes. First, attackers often know a lot about their targets; most people seek to keep sensitive information private from others in their social circle. Against close enemies, social authentication is much less effective. We formally quantify the potential risk of these threats. Second, when photos are used, there is a growing vulnerability to face-recognition algorithms, which are improving all the time. Network analysis can identify hard challenge questions, or tell a social network operator which users could safely use social authentication; but it could make a big difference if photos weren't shared with friends of friends by default. This poses a dilemma for operators: will they tighten their privacy default settings, or will the improvement in security cost too much revenue?

1 Introduction

Facebook[1] recently launched a new user authentication method called "social authentication" which tests the user's personal social knowledge [14]. This idea is neither unique nor novel [17] but Facebook's implementation is its first large-scale deployment. A user is presented with a series of photos of their friends and asked to select their name of a highlighted face from a multiple-choice list (see Figure 1). The current system is used to authenticate user login attempts from abroad.

Facebook has invited security experts to find flaws in the current system before a wider roll-out. If it were deployed for regular authorization and login systems and attacks were to be found subsequently, this could have wide repercussions for the many online merchants and websites which use Facebook to identify their customers, using the Facebook Connect OAuth 2.0 API[2]. We therefore set out to find the best attacks we could on social authentication, and this paper presents our results.

Social authentication is based on the intuition that the user can recognize her friends while a stranger cannot. At first glance, this seems rather promising. However, we argue here that it is not easy to achieve both security and usability:

[1] `http://www.facebook.com/`
[2] `http://developers.facebook.com/docs/authentication`

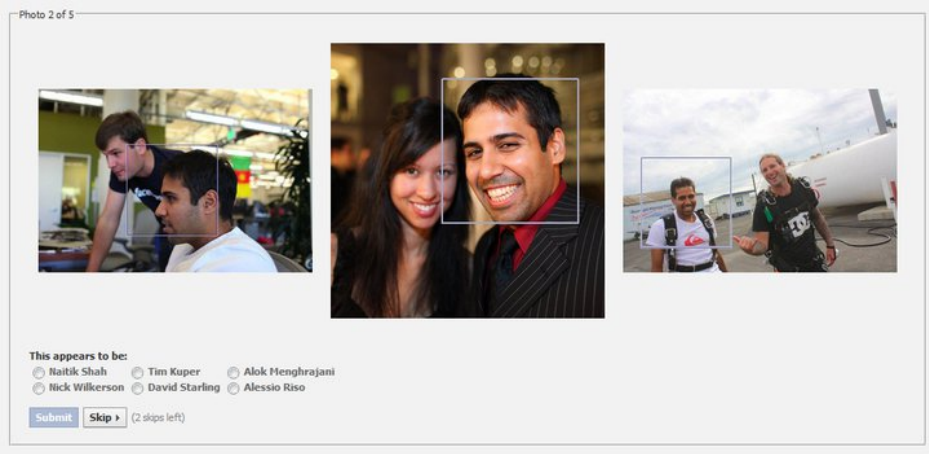


Fig. 1. Social authentication on Facebook. Facebook typically asks the user to name people in three photos.

(1) the user's personal social knowledge is generally shared with people in her social circle; (2) photo-based social authentication methods are increasingly vulnerable to automatic attacks as face recognition and social tagging technologies develop; and (3) we face the same problems as in previous "personal knowledge questions".

In the rest of this article, we will analyse the risk of guessing attacks, then propose several schemes to mitigate them. In community-based challenge selection we use social topology; if a user's friends divide into several disjoint communities, we can select challenge sets that should not be known to any individual friend. We can also reduce the risk of impersonation attacks leveraging the mutual friends between the target user and the adversary; we demonstrate this empirically on realistic data.

2 Why is it difficult to provide secure social authentication?

We analyse three security issues in the photo-based social authentication used in Facebook.

2.1 Friend information is not private enough

Social authentication may be effective against pure strangers. However, the people against whom we frequently require privacy protection are precisely those in our own social circle. For example, if a married man is having an affair, some random person in another country is not likely to be interested; the people who are

interested are his friends and his wife's. In short, users may share a lot of their friends with their adversaries. This is nothing new; 2,400 years ago, Sun-Tzu said 'Keep your friends close, and your enemies closer'. So a proper assessment of the protective power of social authentication in real social networks must be made using real data.

Formally, we view social connections between users in Facebook as an undirected graph $G = (U, E)$, where the set of nodes U represents the users and the set of edges E represents "friend" relationships. For any user $u \in U$, we use f_u to denote the set of u's friends. If each challenge image is selected by the method $\mathcal{M}$, we define the advantage of an adversary a who tries to impersonate the target user u as:

$$\mathbf{Adv}_{\mathcal{M},a}(u,k,\rho) \geq \prod_{i=1}^{\min\{k,|f_u|\}} \Pr\left[c_i \in f_a^{(i)} : c_i \xleftarrow{\mathcal{M}} f_u^{(i)}\right] \cdot \rho \tag{1}$$

where $f_x^{(i)} = f_x - \{c_1, \cdots, c_{i-1}\}$ and k is the number of challenges (such that all k challenges need to be answered correctly) and ρ is the adversary a's *average* success rate to recognize a person in a challenge image c_i when $c_i \in f_a$. It seems reasonable to introduce ρ less than 1 since it may sometimes be difficult to recognize friends if tricky images are selected. For simplification, however, we use ρ as a system parameter.

For any u, k and ρ, we define the impersonation attack advantage of $\mathcal{M}$ via

$$\mathbf{Adv}_{\mathcal{M}}(u,k,\rho) \geq \max_{a \in A_u} \{\mathbf{Adv}_{\mathcal{M},a}(u,k,\rho)\} \tag{2}$$

where the maximum is over all potential adversaries $a \in A_u$ and A_u is the set of users who share mutual friends with u.

In other words, at least one potential adversary a can impersonate the user u with probability at least $\mathbf{Adv}_{\mathcal{M}}(u,k,\rho)$ when k challenge images are provided by the selection method $\mathcal{M}$. If we assume that k challenge images of different friends are randomly selected, the advantage of the impersonation attack in Equation (2) can be computed as follows:

$$\mathbf{Adv}_{\mathcal{R}}(u,k,\rho) \geq \max_{a \in A_u} \left\{ \prod_{i=1}^{\min\{k,|f_u|\}} \frac{|f_{ua}| - (i-1)}{|f_u| - (i-1)} \cdot \rho \right\} \tag{3}$$

where f_{ua} is the intersection of f_u and $\{f_a \cup a\}$ and $\mathcal{R}$ denotes the random selection method.

For example, in Figure 2, since $|f_u| = 5$ and $|f_{ua}| = 2$, we get the probability that a chooses the answer correctly for a challenge image about u is at least $(2/5) \cdot \rho$ when $k = 1$. The probability decreases to $(1/10) \cdot \rho$ when $k = 2$.

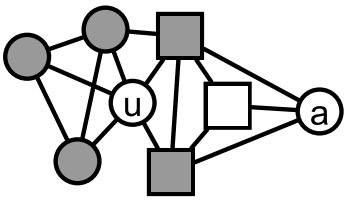


Fig. 2. An example graph with u and a. Nodes represent users and links represent *friend* relationships. The nodes u and a have five (f_u, grey) and three (f_a, square) friends, respectively. They commonly share two friends (f_{ua}, grey-square).

One might think that authentication might be made arbitrarily secure since increasing k will lead to an exponential decrease in the adversary success probability. We decided, however, to use real datasets to explore what value of k might give a good balance between usability and security. With an ideal ρ value ($\rho = 0.99$), we compute the $\mathbf{Adv}_{\mathcal{R}}(u, k, \rho)$ value for each user u by varying k from 1 to 5 on the real Facebook network crawled from both university and regional sub-networks. These sub-networks are summarised in Table 1.

Table 1. Summary of datasets used. $\langle d \rangle$ and n_{cc} represent the "average number of friends" and the "number of connected components", respectively. The sub-networks of universities are highly connected compared to those of regions.

Network	Type	$\lvert U \rvert$	$\lvert E \rvert$	$\langle d \rangle$	n_{cc}
Columbia	University	15,441	620,075	80.32	16
Harvard	University	18,273	1,061,722	116.21	22
Stanford	University	15,043	944,846	125.62	18
Yale	University	10,456	634,529	121.37	4
Monterey Bay	Region	26,701	251,249	18.82	1
Russia	Region	116,987	429,589	7.34	3
Santa Barbara (SB)	Region	43,539	632,158	29.04	1

We display the histograms to show the distributions of the $\mathbf{Adv}_{\mathcal{R}}(u, k, \rho)$ values for all the users in each sub-network. The experimental results are shown in Figure 3.

In order to identify the high-advantage attackers, we calculate the Pearson correlation coefficients between $\mathbf{Adv}_{\mathcal{R}}(u, k, \rho)$ and some representative network centrality that are widely used for measuring the relative importance of nodes in network: degree (**Deg**), closeness (**Clo**), betweenness (**Bet**) and clustering coefficient (**CC**) centrality (see 'Appendix: Network centrality'). The scatter plots in Figure 4 showing the correlation between the adversary's advantage and network centrality visually when $\rho = 0.99$ and $k = 3$. For degree, closeness and betweenness centrality, we can see a negative correlation between the adversary's

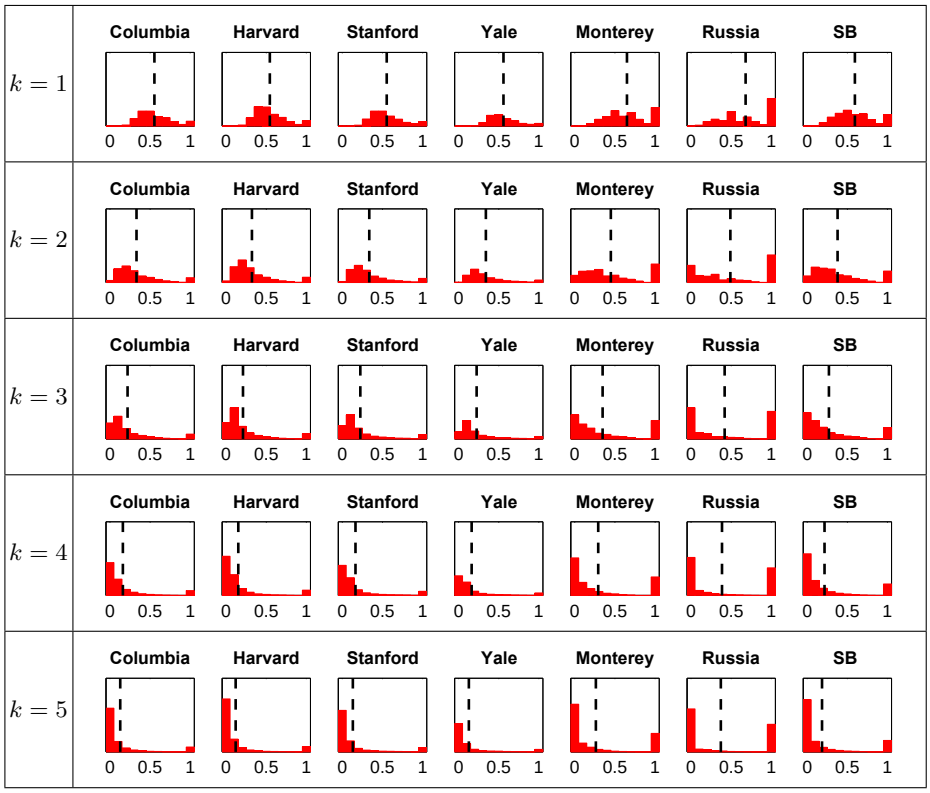


Fig. 3. The histograms of $\mathbf{Adv}_{\mathcal{R}}(u, k, \rho)$ when $\rho = 0.99$ for the users in the seven sub-networks of Facebook in Table 1. The black dotted lines represent the mean of $\mathbf{Adv}_{\mathcal{R}}(u, k, \rho)$ values over the all users in a sub-network.

advantage and nodes' centrality values, although this trend appears to be rather weak for betweenness. In particular, the correlation coefficients for the university datasets are much higher than those for the region datasets. For example, the correlation coefficients between the adversary's advantage and closeness centrality of -0.485 and -0.584 are obtained for each scatter plot graph of the university sub-networks, respectively, while those ranged from -0.00439 to -0.0425 for the region sub-networks. These results indicate that social authentication should not be offered to people with low centrality values.

Another key observation is the correlation between the adversary's advantages and nodes' clustering coefficients. We can see there is a clear correlation (ranged from 0.307 to 0.633) between them although the results are somewhat inconsistent in the cases of 'Monterey Bay' and 'Santa Barbara'. That is, users with high clustering coefficients will become more vulnerable than those with low clustering coefficients. It is natural; the clustering coefficient quantifies how well

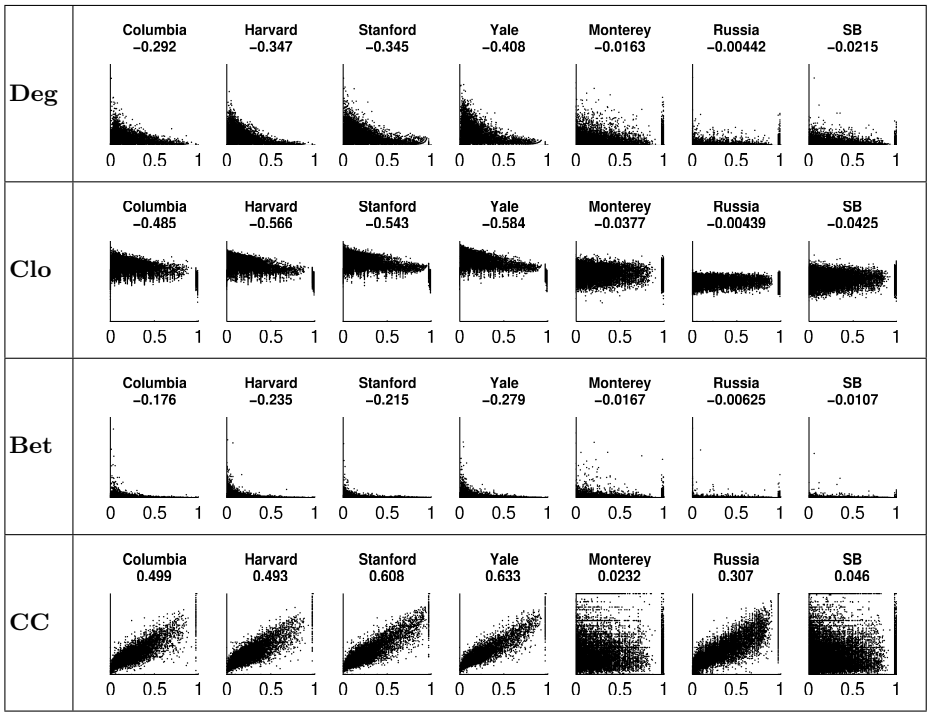


Fig. 4. Scatter plot graphs showing the correlation between the adversary's advantage (X-axis) and network centrality (Y-axis) over nodes when $\rho = 0.99$ and $k = 3$. We also calculate the Pearson correlation coefficient for each scatter plot. These graphs indicate that there exists a negative correlation between the adversary's advantage and network centrality while there exists a positive correlation between the adversary's advantage and clustering coefficient.

a node's friends are connected to each other — we should conclude that social authentication is not recommended for users with high clustering coefficients.

2.2 Automatic face recognition

Social authentication is an extension of image-recognition CAPTCHAs. So we should consider its vulnerability to machine learning attacks; Golle [8] showed that Microsoft's image-recognition CAPTCHA (Asirra) can be broken using machine learning by an adversary who can collect and label a reasonable sample set. So automatic image recognition will be a significant threat to photo-based social authentication. Although face recognition is not a completely solved problem, face recognition algorithms do well under certain conditions. For example, current algorithms are about as good as human judgements about facial identity for "mug shot" images with frontal pose, no facial expression, and fixed illumination [7].

Recent evaluation of face recognition techniques with the real photo images in Facebook [3] showed that the best performing algorithms can achieve about 65% accuracy using 60,000 facial images of 500 users. This shows that the gap between the legitimate user and a mechanised attack may not be as large as one might think.

As with CAPTCHAs, if adversaries use ever-better face recognition programs, the designers could use various tricks to make image recognition – e.g. by noise or distortion – but such images are also hard for legitimate users to identify. The usability costs could be nontrivial. For example, if we reduce ρ to 0.9, then even for $k = 3$ we get an unacceptable user success rate of $(0.9)^3 \approx 0.73$.

To make matters worse, face recognition attacks could be easily extended to large-scale automated attacks by combining the photo collection and recognition processes. As Facebook provides APIs to get images with Facebook ID easily from photo albums, an adversary might automatically collect a lot of high-quality images from the target's friends since many casual users expose their photos in public [1,12]. Although some users do have privacy concerns about sharing their photos, many casual users often struggle with privacy management [13]. Social networks make it difficult for users to manage privacy; it is in their commercial interests for most users to stick with the (rather open) default settings. Therefore an adversary attempting to circumvent social authentication could simply login to Facebook with her own account, access the photos of the victim's friends via the openly available *public search listings* [5,10].

2.3 Statistical guessing attacks

Finally, we revisit statistical guessing attacks which have been studied in the context of personal knowledge questions [9,6]. In particular, Bonneau et al. [6] showed that many personal knowledge questions related to names are highly vulnerable to trawling attacks. The same issues arise in social authentication when the names of a user's friends are sought. The probability distribution of names is not uniform but follows Zipf's law, and the target's language and culture can give broad hints. Even a subject's racial appearance can increase the guessing probability. Since there is a significant correlation between name and race (or gender), the subject's appearance may help an attacker guess his or her name.

3 Toward more secure social authentication

Having identified security problems of photo-based social authentication in Section 2, we now consider what can be done to improve matters.

3.1 Community-based friend selection

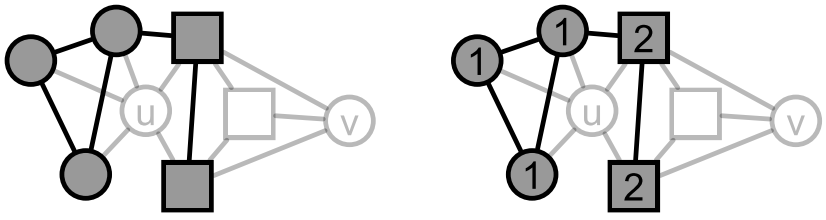
In Section 2.1, we observe that there exists a potential adversary a who can impersonate the target user u with a high probability if the number of challenges

k is small. This is because a shares many mutual friends with the user. In this case, random selection of challenge images may be ineffective.

We propose instead "community-based challenge selection"; our intuition is that a user's friends often fall into several social groups (e.g. family, high school friends, college classmates, and work colleagues) with few, if any, common members. So if we select challenges from different groups, this may cut the attack success probability significantly. We describe this process in detail. For a user u, the k challenges are selected as follows:

1. Extract the subgraph H induced on the user u's friends' nodes f_u from the social graph G.
2. Find the set of community structures $S = \{\eta_1, \cdots \eta_l\}$ in H where η_i represents the ith community structure in H and $l = |S|$.
3. For ith challenge generation for $1 \leq i \leq k$, choose randomly c and remove it from $\eta_{(i \texttt{ MOD } l)}$ where $\eta_l = \eta_0$. After removing c from $\eta_{(i \texttt{ MOD } l)}$, if $\eta_{(i \texttt{ MOD } l)}$ is empty, remove it from S and decrease the indices of the following community structures $\{\eta_m : (i \texttt{ MOD } l) < m \leq l\}$ and the total number of community structures l by 1.

For example, we extract the subgraph H induced on f_u in Figure 5(a) and then find two community structures of H by applying a community detection algorithm. Although a specific heuristic method [4] is used here for community detection, we expect that any community detection algorithm can be used for this purpose. In this example, unlike the results of the random selection in Section 2.1, v cannot impersonate u since we choose a challenge from the community structure η_1 in Figure 5(b) when $k = 1$.



(a) Subgraph H (b) Communities of H

Fig. 5. An example of how the community structures are detected. **(a)** The subgraph H is induced on the user u's friends' nodes f_u (f_u, grey). **(b)** Two community structures $S = \{\eta_1, \eta_2\}$ are detected in H.

Formally, if we select k challenge images using community-based challenge selection, the advantage of the impersonation attack A, $\mathbf{Adv}_A(u, k, \rho)$, can be computed as follows:

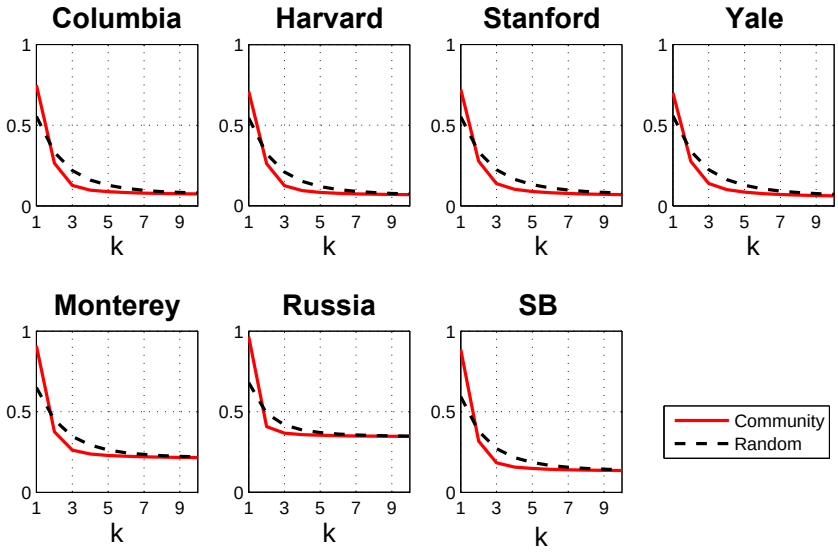


Fig. 6. Comparison of the mean values of adversary advantage between community-based challenge selection (red solid line) and the random challenge selection (black dashed line).

$$\mathbf{Adv}_{\mathcal{C}}(u,k,\rho) \geq \max_{\substack{v \in U \\ u \neq v}} \left\{ \prod_{i=1}^{\min\{k,|f_u|\}} \frac{|\eta_{(i \bmod l)}(v)|}{|\eta_{(i \bmod l)}|} \cdot \rho \right\} \tag{4}$$

where $\eta_i(v)$ is the intersection of η_i and $\{f_v \cup v\}$ and $\mathcal{C}$ denotes the community-based challenge selection method.

To validate the effectiveness of this selection method, we compute the mean values of $\mathbf{Adv}_{\mathcal{C}}(u,k,\rho)$ on the preceding datasets in Section 2.1 and compare those of $\mathbf{Adv}_{\mathcal{R}}(u,k,\rho)$ with random selection. The experimental results (for $\rho = 0.99$) are shown in Figure 6 which shows almost the same slope patterns for all the datasets. Community-based selection (**C**, red solid line) performed significantly better than random selection (**R**, black dashed line) from $k = 2$ to 5. But if we use a single challenge image (i.e. $k = 1$), it does worse! Since the first challenge is selected from the first community η_i only, the attack success probability of anyone in that community η_i is increased. The gap between community-based and random challenge selection is largest for $k = 3$ or 4, and the mean values of the adversary's advantage tend to converge slowly. In fact, community-based challenge selection is comparable at $k = 3$ to random selection at $k = 10$.

We hypothesised that setting k to the "number of community structures" would enable community-based selection to get a good tradeoff between security and usability. In order to test this, we analysed the average number of community structures for each user's friends. The results are shown in Table 2 where friends

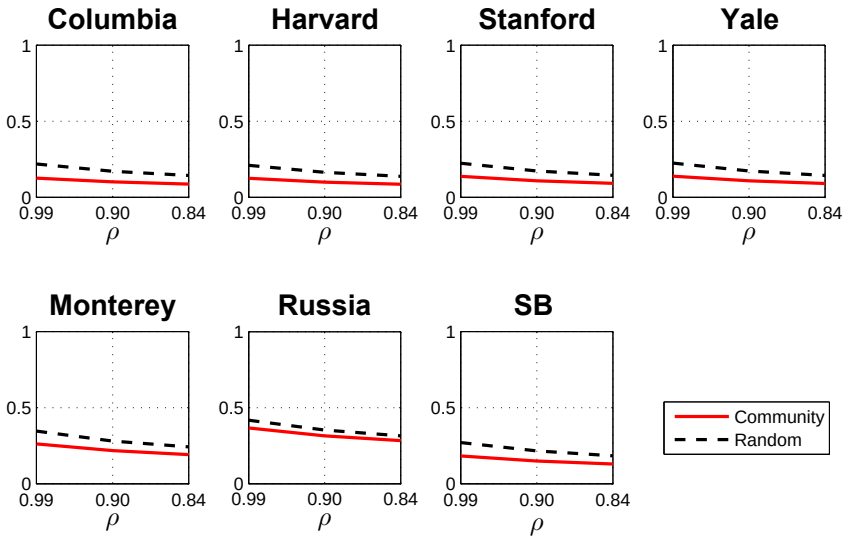


Fig. 7. the adversary advantage between the *community-based challenge selection* (red solid line) and the *random challenge selection* (black dashed line) by varying ρ from 0.99 to 0.84 when $k = 3$.

can be divided into about three or four communities on average except in the Santa Barbara sub-network.

Table 2. The average number of communities for each user's friends.

Columbia	**Harvard**	**Stanford**	**Yale**	**Monterey**	**Russia**	**Santa**
3.779	3.371	3.227	2.812	3.690	3.099	4.980

We verified this hypothesis by calculating the average number of community structures for each user's friends and found that indeed friends can be divided into about three or four communities on average; the exception being Santa Barbara sub-network which had 5 communities.

We now discuss how adversary advantage may change with the friend recognition success rate ρ (see Figure 7). To demonstrate this we fix $k = 3$. As ρ decreases from 0.99 to 0.84, the advantage values of both selection methods also slightly decrease. However, the change of ρ does not significantly affect the advantage values compared to the change of k or the challenge selection methods. These values were derived from user success rates for existing image-based CAPTCHAs [11].

In all our experiments, the average number of communities is always a small number (less than 5). Since we use campus or region networks, the number of communities might be small compared to real friendship patterns in Facebook,

which could include structures of high school friends, college classmates, work colleges and so on. Recently, some social networking services such as Google+[3] and Facebook have started to encourage users to divide their friends into explicit community groups; community-based challenge selection should be even more useful in such situations.

3.2 Exclusion of well-known or easily-recognizable friends

In order to mitigate the threat via automatic face recognition program discussed in Section 2.2, some might suggest that we should educate users about these attacks, but that has been found in many applications to not work very well; "blame and train" is not the way to fix usability problems.

One approach may be to exclude users who make all their photos visible to everyone or "friends of friends" – an option in Facebook. This will prevent collection of the training data needed for automatic face recognition tools. There may be technical options too. As face-recognition software tools improve, they can be incorporated into the challenge generation system – by rejecting candidate challenge images whose subjects they can identify.

However, we should be cautious in using a long blacklist of photos; such a policy may shrink the space of challenge photos to the point that an adversary can just guess the answer to a given challenge.

3.3 Weighted random sampling

In order to reduce the risk of the statistical guessing attacks discussed in Section 2.3, and which leverage the probability distribution of people's names, we suggest using weighted random sampling instead of uniform random sampling.

Under uniform sampling, a name n is selected with the probability $f(n)$ where f is the probability density function for a set of names of people $\mathcal{P}$. Alternatively, in weighted random sampling, n is selected with the following probability:

$$w(n) = \frac{f(n)^{-1}}{\sum_{p \in \mathcal{P}} f(p)^{-1}} \tag{5}$$

Intuitively, in this case, friends with infrequent names will be selected with higher probability compared to friends with popular names when a challenge image is chosen. In a global view, the estimated probability density function of the users' names in challenge images might tend to be the uniform distribution if the number of users with popular names is much greater than that with unpopular name. So selecting popular names as challenge answers won't help the attacker any.

However, if an adversary can crawl all names of a victim's friends successfully, weighted random sampling is worse than uniform random sample unlike our

[3] `https://plus.google.com/`

expectation; an attacker can choose a name from the crawled names in proportion to the above probability since the challenge image is chosen with the probability. Thus in practice a more complicated weighted random sampling technique should be considered based on real statistics of privacy settings. As part of the future work, we plan to design more advanced weighted sampling methods.

4 Related work

Our work focuses on the security and usability of photo-based social authentication methods. Social authentication was introduced under the belief that adversaries halfway across the world might know a user's password, but they don't know who the user's friends are.

Yardi et al. [17] proposed a photo-based authentication framework and discussed some security issues including Denial of Service (DoS) attacks: an adversary can spam the system with photos with wrong tagging information so legitimate users cannot pass the authentication test. They also mentioned attacks by a network outlier belonging to the same group as the target. We extended this attack formally and experimentally measured the level of threat.

In social networks, photo privacy may become even more problematic as social networking websites such as Facebook have become the primary method of sharing photos between people [16]. Ahern et al. [2] examined users' decisions when posting photos to Flickr[4] with mobile camera phones, finding that many users were concerned with protecting their personal images and keeping them out of public view. Most social networking websites already provide mechanisms for fine-grained photo sharing control, but user surveys [1,12] have shown that over 80% of social network users do not change their privacy settings at all from the default. This implies that photo-based social authentication is very vulnerable in practice to face recognition tools.

5 Conclusion

Facebook recently launched an interesting authentication method [14], and is currently waiting for feedback from the security community before pushing it out to a wider range of authentication and login services including, potentially, third-party merchants who utilise the Facebook Connect API.

This article provides that feedback. We found that the current social authentication scheme is susceptible to impersonation both by insiders and by face-recognition tools, and a naive approach to selecting friends isn't effective against either attack. It is hard to identify the social knowledge that a user holds privately since social knowledge is inherently shared with others. A critical observation is that many likely attackers are 'insiders' in that the people who most want to intrude on your privacy are likely to be in your circle of friends.

[4] `http://www.flickr.com/`

We set out to formally quantify the difficulty of guessing the social information of your friends (and your friends' friends) through the analysis of real social network structures and analysed how this can interact with technical attacks such as automatic face recognition and statistical guessing.

We proposed several ways to mitigate the threats we found. Community-based challenge selection can significantly reduce the insider threat; when a user's friends are divided into well-separated communities, we can select one or more recognition subjects from each. We can also avoid subjects with common names or who are known in multiple communities. But perhaps the most powerful way to improve social authentication will be to exclude subjects who make their photos visible to friends of friends. At present, that's most users, as 80% of users never change the privacy defaults – presumably there was some marketing advantage to Facebook in having relaxed privacy defaults, in that making the photos of friends' friends visible helped draw in new users, increasing the network effects; so a change to a default of sharing photos only with friends could give a real security improvement.

In analysing the adversary's advantage, we assumed some fixed constants (e.g. the adversary's average success rate to recognize a person in a challenge image) rather than actual testing results through user studies on Facebook. So our analysis is still rather limited. To verify this point in a practical environment, we plan to conduct a user study to evaluate the effectiveness of the attack and mitigation techniques.

Acknowledgements

We thank Ben Y. Zhao and Joseph Bonneau for their Facebook datasets.

Appendix: Network centrality

Formally, we use the standard definition [15] of the *degree*, *closeness* and *betweenness* centrality values of a node u.

Degree centrality simply measures the number of direct connections to other nodes. This is calculated for a node u as the ratio of the number of edges of node u to the total number of all other nodes in the network. Degree centrality can be simply computed but does not take into account the topological positions of nodes and the weights of edges.

Closeness centrality expands the definition of degree centrality by measuring how close a node is to all the other nodes. That is, this metric can be used to quantify in practical terms how quickly a node can communicate with all other nodes in a network. This is calculated for a node u as the average shortest path length to all other nodes in the network:

$$\mathbf{Clo}(u) = \frac{1}{|V|-1} \sum_{v \neq u \in V} \mathrm{dist}(u, v) \tag{6}$$

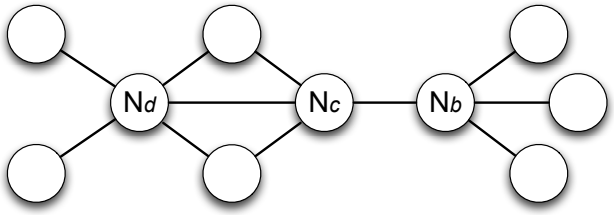


Fig. 8. The characteristics of network centrality. In this network, N_d has higher *degree* centrality than N_c since N_d has five neighbours while N_c has higher *closeness* centrality than N_d. We note that N_d is located at the periphery of the network compared to N_c. Interestingly, N_b has the highest *betweenness* centrality. We can see that N_b plays a 'bridge' role for the rightmost nodes.

where dist(u, v) is the length of the shortest path from node u to node v. In an undirected graph, $dist(u, v)$ is the number of hops in the shortest path from node u to node v.

Betweenness centrality measures the paths that pass through a node and can be considered as the proportional flow of data through each node. Nodes that are often on the shortest-path between other nodes are deemed highly central because they control the flow of information in the network. This centrality is calculated for a node u as the proportional number of shortest paths between all node pairs in the network that pass through u:

$$\mathbf{Bet}(u) = \frac{1}{(|V|-1)\cdot(|V|-2)} \sum_{s\neq u, t\neq u\in V} \frac{\sigma_{s,t}(u)}{\sigma_{s,t}} \tag{7}$$

where $\sigma_{s,t}$ is the total number of shortest paths from source node s to destination node t, and $\sigma_{s,t}(u)$ is the number of shortest paths from source node s to destination node t which actually pass through node u. For normalization, it is divided by the number of all pairs of s and t.

In Figure 8, for example, the nodes N_d, N_c, and N_b illustrate the characteristics of these network centrality metrics. These nodes have the highest *degree*, *closeness* and *betweenness* centrality, respectively.

Clustering coefficients measures the probability of neighbours of a node to be neighbours to each other as well. This is calculated for a node u as the fraction of permitted edges between the neighbours of u to the number of edges that could possibly exist between these neighbours:

$$\mathbf{CC}(u) = \frac{2\cdot\Delta}{(\kappa_u)(\kappa_u - 1)} \tag{8}$$

where Δ is the number of the edges between the neighbours of node u and κ_u is the number of the neighbours of node u (i.e. the degree of node u).

References

1. Acquisti, R., Gross, R.: Imagined communities: Awareness, information sharing, and privacy on the Facebook. In: Proceedings of the 6th Workshop on Privacy Enhancing Technologies. pp. 36–58 (2006)
2. Ahern, S., Eckles, D., Good, N.S., King, S., Naaman, M., Nair, R.: Over-exposed?: privacy patterns and considerations in online and mobile photo sharing. In: CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 357–366. ACM, New York, NY, USA (2007)
3. Becker, B.C., Ortiz, E.G.: Evaluation of face recognition techniques for application to facebook. In: IEEE International Conference on Automatic Face and Gesture Recognition. pp. 1–6 (2008)
4. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Unfolding communities in large complex networks: Combining defensive and offensive label propagation for core extraction. Physical Review E 83(3), 036103 (2011)
5. Bonneau, J., Anderson, J., Anderson, R., Stajano, F.: Eight friends are enough: social graph approximation via public listings. In: Proceedings of the Second ACM EuroSys Workshop on Social Network Systems. pp. 13–18. SNS '09, ACM, New York, NY, USA (2009)
6. Bonneau, J., Just, M., Matthews, G.: What's in a Name? Evaluating Statistical Attacks on Personal Knowledge Questions. In: Financial Cryptography and Data Security '10 (2010)
7. Daugman, J.: The importance of being random: statistical principles of iris recognition. Pattern Recognition 36(2), 279–291 (2003)
8. Golle, P.: Machine learning attacks against the Asirra CAPTCHA. In: CCS '08: Proceedings of the 15th ACM conference on Computer and communications security. pp. 535–542. ACM, New York, NY, USA (2008)
9. Just, M.: On the design of challenge question systems. IEEE Security and Privacy 2, 32–39 (September 2004)
10. Kim, H., Bonneau, J.: Privacy-enhanced public view for social graphs. In: SWSM '09: Proceeding of the 2nd ACM workshop on Social web search and mining. pp. 41–48. ACM, New York, NY, USA (2009)
11. Kluever, K.A., Zanibbi, R.: Balancing usability and security in a video CAPTCHA. In: SOUPS '09: Proceedings of the 5th Symposium on Usable Privacy and Security. pp. 1–11. ACM, New York, NY, USA (2009)
12. Krishnamurthy, B., Wills, C.E.: Characterizing privacy in online social networks. In: WOSP '08: Proceedings of the first Workshop on Online Social Networks. pp. 37–42. ACM, New York, NY, USA (2008)
13. Lipford, H.R., Besmer, A., Watson, J.: Understanding privacy settings in facebook with an audience view. In: Proceedings of the 1st Conference on Usability, Psychology, and Security. pp. 2:1–2:8. USENIX, Berkeley, CA, USA (2008)
14. Rice, A.: A Continued Commitment to Security. `http://blog.facebook.com/blog.php?post=486790652130` (January 2011)
15. Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications. Cambridge University Press (1994)
16. Willinger, W., Rejaie, R., Torkjazi, M., Valafar, M., Maggioni, M.: Research on online social networks: time to face the real challenges. SIGMETRICS Performance Evaluation Review 37, 49–54 (January 2010)
17. Yardi, S., Feamster, N., Bruckman, A.: Photo-based authentication using social networks. In: WOSP '08: Proceedings of the first Workshop on Online Social Networks. pp. 55–60. ACM, New York, NY, USA (2008)