Using Shared Knowledge Questions for Authentication and Access Control

ABSTRACT

We propose a novel security mechanism for controlling access to socially sensitive content, such as online photographs. Rather than require separate authentication and entry of explicit access control lists, we allow users to control access implicitly through shared knowledge questions. We implemented a partial prototype, and conducted studies that enumerate the context of photo sharing security, gauge the difficulty of creating shared knowledge questions, measure their resilience to adversarial attack, and evaluate users' abilities to understand and predict this resilience.

ACM Classification: K6.5 [Management of Computing and Information Systems]: Security and Protection. - Authentication.

Keywords: Privacy, Access Control, Photo Sharing

INTRODUCTION

People are increasingly sharing photos, videos, blogs, location, activity, exercise logs and other personal artifacts online, but might prefer a boss, family member, or a stranger not see some of them. Consequently, sharers must specify *access control*: a set of rules that allow access to some people, and deny it to others. Unfortunately, contemporary access control leads to usability and social difficulties. We will now illustrate these in scenarios before explaining them in more detail.

Scenario 1: Mark has a new girlfriend. He wants to remain friends with his ex, who he talks to periodically on Facebook, but still shield her from photos of his new girlfriend. Without internet-defriending his ex, how can Mark upload these photos to Facebook for his other friends to view?

Scenario 2: Susan wants to share vacation photos of her children and family with relatives. Many photos depict her young children in bathing suits, and she is not comfortable making them available to anonymous web surfers and potential web predators. (We learned this is a common situation from interviews with a major photo-sharing website.) However, if she makes the photos "private", she must expend a great deal of effort finding current email addresses for all 80 of her relatives and friends, adding them to her "friends and family" list on the photo sharing website, sending them invitations, and ensuring that list is enabled to see her vacation photos. Unfortunately, she might have to repeat this process for people whose email addresses have changed, relatives who have forgotten their account logins and passwords, and people she forgot to include the first time. Her aunt Audrey in particular was upset to be omitted from the access list.

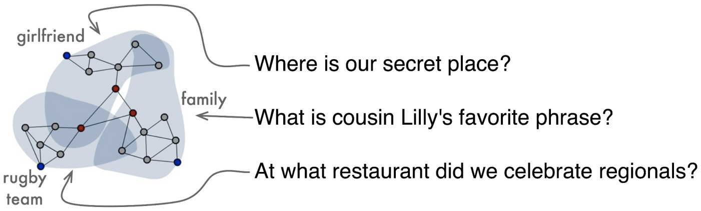
girlfriend
family
rugby team
Where is our secret place?
What is cousin Lilly's favorite phrase?
At what restaurant did we celebrate regionals?

Figure 1: A concise question of shared knowledge can implicitly define a region of friends, family, and colleagues in a social network, without explicitly describing the network or its cliques

Scenario 3: Ryan is a college student looking for a job. She befriended colleagues, potential employers, and college buddies on Facebook, and does not want to explicitly defriend any of them, or create a "Facebook limited profile", which she is worried they could find disrespectful. However, she also thinks it is inappropriate for unfamiliar business acquaintances to view her college party photos.

Scenario 4: Jane is an artist who sometimes blogs intimate details of her life. Complete strangers and some relatives might find her blog vulgar. On the other hand, a password access control might severely limit blog readership within her art community. How can she satisfy these competing constraints?

Current photo sharing websites use variations of *whitelists* and *blacklists*, in which users explicitly list people or groups who should or should not be given access. We argue that white and blacklists are tedious and inflexible, and can be rude.

We propose, instead, that sharers design *guard questions* such as “where did I travel this summer” or “what is my dog’s name” that must be answered to view a photo or album, leveraging the shared knowledge preexisting in social networks (Figure 1). We conducted a study to investigate the design and security of guard questions. Our work is guided by the observation that social security may not need to be "hard" in the cryptographic sense, and might prioritize usability, flexibility, ambiguity, and social nuance instead, thus being useful in a new array of situations.

Traditional access control: whitelists and blacklists

Whitelists and blacklists require users to explicitly translate social relationships into lists of account names and/or email addresses. This is problematic in a few ways:

Tedious

Creating and maintaining lists for many photos or albums, each with many sharees, requires substantial work, particularly for people without existing website accounts, and makes it easy to forget to include people.

Inexpressive or complicated

To alleviate the tedium of large lists, websites let users white or blacklist predefined groups of users, such as "friends and family". However, these do not allow fine-grained exclusions, such as in scenarios 1 and 3.

On the other hand, more expressive grouping mechanisms, such as those in operating systems, become complicated to use in ways similar to programming: they require education, abstract reasoning, advance planning, and debugging.

Thus, white and blacklists exist in a bounded sea of zero-sum tradeoffs: without groups they are tedious, with arbitrary groups they are complicated, and with predefined groups they are inexpressive. Guard questions may be more flexible.

Rude and lacking social nuance

Social relations are inherently soft and ambiguous, yet white/blacklists are hard and binary. The mere act of categorizing individuals into groups is known to produce prejudice and discrimination. [4] It can be insulting to learn you are on a friend's blacklist; it is less offensive to be unable to answer a question about her summer travels. As a medium, the internet already polarizes social relationships, and it is worth pursuing authentication policies that allow more social nuance.

STUDY

Our study probes guard questions for photo sharing via three progressive levels of inquiry. First, with whom do sharers want to show or hide their photos? Second, what types of questions do sharers devise, and how difficult are they to design? Finally, how easy is it for a stranger to crack these questions, and are users able to predict crackability? To answer the first two questions, we had participants devise questions for their own photos. To answer the third, we then uploaded these questions as challenges to Amazon Mechanical Turk, and rewarded anonymous internet workers to guess the answers.

Designing guard questions

We first recruited 31 people to find a total of 179 photos that they wanted to share with some people, but not with others. Subjects reported who they would want and not want to see each photo, as well as the importance of seeing or not seeing it on a 4 point ordinal scale, ranging from (1) "I barely care" to (4) "I care a whole lot". Finally, they designed guard questions that they felt would effectively control access to each photo. For each question, they reported how long the design took and how many of 10 random strangers they thought could guess the answer in 10 guesses. Our participants were a fairly diverse group: 47/53% male/female, mean age 27 (stdev 8), recruited through flyers in 2 websites and 3 urban neighborhoods.

Results: desired and undesired recipients

We clustered 315 responses of desired photo recipients and 401 of undesired recipients into 9 emergent categories:

Category of person or group of people	Desired		Undesired	
	Freq.	Imp.	Freq	Imp.
Friends	90%	2.2	41%	3.0
Family	76%	2.4	79%	3.0
Strangers	0%	--	72%	2.8
Specific people by name	46%	2.8	24%	2.4
Common interest group	38%	1.7	41%	3.0
Friends of photographed	34%	2.5	0%	--
Authority figures	21%	3.2	42%	3.0
Ex-friends and romances	0%	--	14%	2.7
Potential romances and employers	10%	3.5	7%	3.6

Table 1: Desired and undesired people to see photos. *Freq* is percentage of responses that include a category. *Imp.* is mean rated importance of the responses in a category, on our 1-4 ordinal scale.

Demonstrating a need for flexible access control policies, 58% of participants had a category that both *should* see one photo but *should not* see another, which most simple white/blacklists do not support. Additionally, 83% of participants had photos to hide from friends, family or co-workers: people who are likely to be on most social networking access lists. On average, people cared more about preventing access (2.6) than providing it (2.2) (p<.001).

Results: guard questions

Subjects easily understood the concept of guard questions, and could readily create them after reading a one-paragraph description. They designed 168 unique guard questions (and 11 duplicates), which we clustered into 6 categories:

Question Type	Example Question	Freq.
About themselves	What's my favorite spirit for mixed drinks?	48%
Knowledge of a mutual friend	What was the name of Susan's hairy dog?	13%
About a specific place or event	In what country did I work in Europe?	12%
About the guesser	What river did we float down for Keith's B-Day?	10%
Inside joke or reference	Spiky red hair on the dance floor drink	8%
General Knowledge	The "AP" in AP Stats stands for?	6%

Table 3: Categories of questions generated

Subjects successfully created questions for all but 3 of the 179 photos, a 98% success rate, indicating that there exists appropriate shared knowledge to separate most inclusion/exclusion groups. Subjects spent a median of 8 seconds designing each guard question, according to self report. This compares favorably to the design time of many whitelists—it takes the first author an average of 9 seconds per person to search for, select, copy and paste a list of friends' email addresses from his OSX address book. However, some guard questions in the tail of the distribution took much longer. The mean and standard deviation were 15 and 28 seconds, respectively. We also observed strong individual differences. One subject averaged a report of 155 seconds over her 8 questions, with the longest being 600 seconds. Future work should investigate the cause. We found no significant effect of design time on end security, to which we now turn.

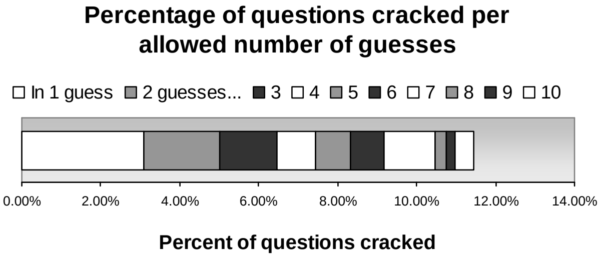

Figure 2: If we allow 3 guesses, 6% of questions are cracked. With 10 guesses, 11% are cracked.

Cracking the questions

Understanding the breadth of disclosure is critical for privacy-sensitive systems [2]. To learn both how hard questions are to crack, and gauge subjects' ability to predict such hardness, we uploaded the questions to be cracked as jobs on Amazon's Mechanical Turk, a web marketplace that pays people to complete small tasks. We recruited 10 workers per question to take 10 guesses each. Workers were motivated with a bounty of $.75 for a successful crack within 3 guesses, and $.25 for a crack within the remaining 7. For reference, many Turk jobs pay pennies for a similar time commitment. All Turk workers received $.05 just for guessing. We manually verified the quality of Turk responses; a few poor responses were rejected, but the vast majority were of high quality.

As can be seen in Figure 2, Turk workers cracked an average of 6% of the questions in 3 guesses, and 11% in 10 guesses. Thus, allowing only 3 guesses reduces the crack rate by almost half. The crack rate is less than subjects' average predicted rate of 14%, thus the average subject has slightly better security than she expects.

We compare subjects' predicted with actual crack rates in Figure 3. The mispredictions are in the lower-right and upper-left. Of the 168 questions, only 11 have predicted crack rates off by more than 30%, and of these, just 7 (=4%) are less rather than more secure than expected. A strong majority (143=85%) have both crack rates under 20% and predictions under 40%. A linear regression gives R^2=44% between coordinates. These numbers are all for 10 guesses.

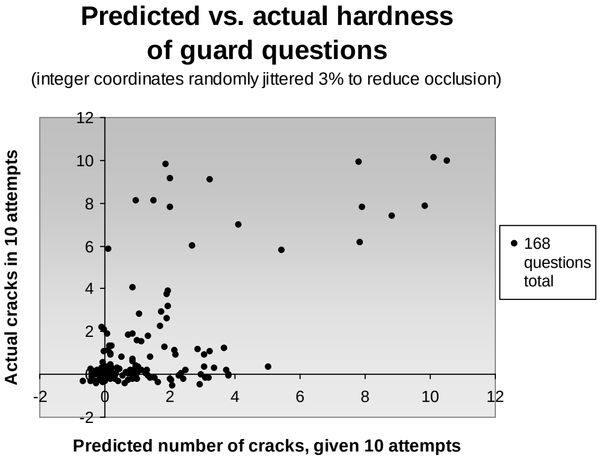

Figure 3: Most questions were difficult to crack, as subjects predicted. The questions in the upper-left (4-6% of the data) were cracked more frequently than subjects predicted.

We examined the 7 cases in the upper-left with crack rates more than 30% higher than predicted rates. We found two common flaws: 5 questions asked for an easily-enumerable class of answers, such as a small number, color, or day of the week (e.g. "What night of the week do I usually stay out late?"); and 2 questions could be answered by searching Google for the question and browsing the first page of results (e.g. try searching "Who lives in Chris's closet on FG?"). One could imagine a system that counts ontologies and does web searches to discover such weak passwords automatically and suggest alternatives.

We suspect these rates are acceptable. If we introduce discretionary use of guard questions vs. whitelists, shorter guess limits, and a user interface that informs sharers of wrong guessers and their guesses (giving the guesser social liability, and the sharer a chance to fix weak questions), then cracking might be a manageable problem.

HANDLING AMBIGUOUS ANSWERS

Some correct guesses do not exactly match their answer's text. Here we describe the issues in, and our approach to, implementing automated ambiguous answer verification.

To derive ground truth, we manually labeled all guesses as *correct*, or *incorrect*. In this process, we discovered the following textual deviations, for which we created a rubric.

Intra-word deviations: We allowed spelling errors and stemming differences, such as "Teriers" for "Terrier".

Alternative words: Abbreviations, acronyms, and synonyms were treated as different, incorrect words.

Extra or missing words: We ignored *stop words*, such as "and", "or", and "to". If a guess had a few *extra* words, such as "seattle downtown" instead of "seattle", we considered it over-specified and correct. If a guess was *missing* words, such as "grandparents" instead of "gabe's grandparents", it was considered under-specified and incorrect.

This rubric was problematic in two cases: the university "case western" was judged correct for the university "western", even though "case western" is *not* a specialization of "western". Similarly, the answer "2005 and 2007" incorrectly accepted a guess of "2003 2004 2005 2006 2007". As a solution to the latter, the question designer could specify whether a guess must *be* or *contain* the answer.

Implementing automated answer verification

We implemented this rubric as an algorithm. In it, we ignore case and punctuation, and remove all stop words from both the guess and answer. Next, we check that each word in the answer appears in the guess, allowing 1 character difference for non-numeric words 2-8 characters long, and 2 differences for larger words. The code is 34 lines. This simple algorithm performed identically to human labeled ground truth, and could be a starting point for a practical system. Future work could also leverage NLP research to handle synonyms, acronyms, and abbreviations.

RELATED WORK

Many *personal* authentication systems require answers to questions on personal knowledge. Cognitive Passwords [5], for instance, probe questions like "mother's maiden name". Many of these ideas could be extended to use *shared* knowledge. For instance, Pering et. al. [3] describe an authentication method in which users identify their personal photos from a long set of distracters, which avoids replay attacks in single user authentication. We suggest that members of a group could also identify photos from *shared knowledge* of a group photo pool, and enable group network authentication, avoiding replay attacks and the need for as large a personal photo library.

Shared passwords and keys are an alternative to allowing access without account creation. However, unlike guard questions, these passwords or keys must be distributed to a whitelist of users, rather than letting them stumble onto content. Furthermore, users must remember or store and manage these foreign passwords (one for each whitelist they are on), whereas shared knowledge answers are by nature easy to remember, since they are aspects of a user's real life. This makes shared knowledge a useful guard for long lived family photo albums, for instance. Finally, guard questions can be changed, allowing different groups of people, at anytime without redistributing passwords.

Recent research has improved usability of systems oriented white/blacklist access controls. See Cao [1] for an example.

We have also found existing ad-hoc implementations of shared knowledge questions on the web, guarding blogs and photo albums. We think an improved understanding and robust implementation may increase their use.

FUTURE WORK

Our study only examines crack attempts from complete strangers. This paints an incomplete picture. It is critical to evaluate access from intended sharees, as well as unintended users with partial knowledge of the subject, or investigative tools (*e.g.* Google stalking or asking friends). Such study may require long term use of a real system. Real system use could also illuminate the social side-effects of guard questions, such as differences in feelings of social exclusion or ostracization vs. white/blacklists.

There are many potential avenues to reduce the rate of unexpected cracks, both through interaction and analysis, some of which we have already mentioned. At a high level, we would like to see real-time visualizations of guesses and guessers; languages for sharers to specify alternative answers and ambiguity bounds; empirical investigations into weak question/answer types; providing a set of predefined questions to choose from rather than freeform text to avoid paradox of choice and weak question types; cognitive analysis of systematic crack rate underestimates; automatic detection of weak questions; and natural language analyses for answer verification and weak question detection.

We would also like to apply shared knowledge challenges to domains beyond photo sharing, such as blogs, real-time location data streams, automatically moderating mailing list subscriptions, subculture-specific Captchas, and group project Wiki access control. Guard questions could also be combined with traditional access controls in interesting ways. For instance, one might put a guard question on top of a hidden blacklist to add plausible deniability.

CONCLUSION

We present a new class of mechanisms for implicit access control using shared knowledge, in which the design of a concise question replaces multi-user authentication and access control lists. Users readily learn the concept, and can design questions with a moderate but variable amount of effort. Most questions are hard to crack. Although users are sometimes unaware of their question's security level, there are possibilities for future work to mitigate this. Shared knowledge questions also have a wide variety of potential cross-domain applications in ubiquitous computing and electronically-mediated social communication.

REFERENCES

1. Cao, X. and Iverson, L. (2006). *Intentional Access Management: Making Access Control Usable for End-Users.* Proceedings of the Symposium on Usable Privacy and Security, (SOUPS 2006) 20-31.
2. Lederer, S., Jason Hong., Dey, A.K., and Landay, J. (2004). *Personal Privacy through Understanding and Action: Five Pitfalls for Designers. Personal and Ubiquitous Computing*. **8**(6), 440-454.
3. Pering, T., Sundar, M., Light, J. and Want, R. (2003). *Photographic Authentication through Untrusted Terminals*. IEEE Pervasive Computing, **2**(1), 30-36.
4. Tajfel H, Billig M G, Bundy R P & Flament C. (1971). *Social Categorization and Intergroup Behaviour*. European Journal of Social Psychology **1**(2), 149-177.s
5. Zviran, M., Haga, W.J. (1990). *User Authentication by Cognitive Passwords: An Empirical Assessment. Jerusalem Conference on Information Technology*, 137-144.