Access Control by Testing for Shared Knowledge

Michael Toomim[1], Xianhang Zhang[2], James Fogarty[1] and James A. Landay[1]

Computer Science & Engineering[1]
DUB Group, University of Washington
Seattle, WA 98195
{toomim,jfogarty,landay}@cs.washington.edu

Human Interface Technology Laboratory[2]
DUB Group, University of Washington
Seattle, WA 98195
xianhang@u.washington.edu

ABSTRACT

Controlling the privacy of online content is difficult and often confusing. We present a social access control where users devise simple questions of shared knowledge instead of constructing authenticated accounts and explicit access control rules. We implemented a prototype and conducted studies to explore the context of photo sharing security, gauge the difficulty of creating shared knowledge questions, measure their resilience to adversarial attack, and evaluate users' ability to understand and predict this resilience.

Author Keywords: Privacy, Access Control, Photo Sharing.

ACM Classification Keywords: H5.m. Information interfaces and presentation: User Interfaces

INTRODUCTION

People are increasingly sharing their lives online in photos, videos, blogs, location and activity status, exercise logs and other personal artifacts. But they often require that a boss, family member, or 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.

Although contemporary access control, based on explicit blacklists and "friend" whitelists, is mathematically precise, it can also be too tedious, inflexible, complicated, or rude in many scenarios. How can a mother share photos of her children with 80 extended family members and family friends, but *not* potential Internet predators, without enumerating all 80 viewers, finding their email addresses, getting them accounts and passwords, and whitelisting them? How can an artist give her local art community access to her personal blog, without requiring a login and password, which could severely limit readership? How can a man prevent an ex-girlfriend from seeing his new girlfriend's Facebook photos, visible to all "friends", without defriending his ex? How can a college student conceal Facebook party photos from employers without blocking them on a potentially offensive blacklist?

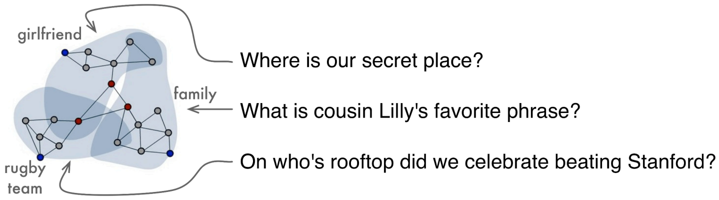
girlfriend
family
rugby team
Where is our secret place?
What is cousin Lilly's favorite phrase?
On who's rooftop did we celebrate beating Stanford?

Figure 1: A concise question of shared knowledge can define a region of friends in a social network without explicitly defining the network or its cliques.

We observe that social cliques overlap with regions of shared knowledge (Figure 1), and propose that sharers design *guard questions of shared knowledge,* such as "what is cousin Lilly's favorite phrase" that must be answered to view a photo or album. We present a discussion of design issues and a study investigating the design and security of shared knowledge questions. Our work is guided by the observation that social security may not need to be "hard" in the strict, cryptographic sense, but may rather prioritize usability, flexibility, ambiguity, and social nuance instead, thus being useful in a new array of situations.

Traditional Access Control: Whitelists and Blacklists

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

Tedious

Authenticating accounts and creating and maintaining lists for many photos or albums, each with many accessors, requires substantial work, and makes it easy to forget people.

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 policies that allow more social nuance.

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 personalized groups, such as "close friends", or special exclusions.

CHI 2008, April 5–10, 2008, Florence, Italy.

On the other hand, more expressive grouping mechanisms, such as UNIX groups, 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. Shared knowledge may be more flexible.

COPING WITH GUESSERS AND FORGETTERS

On the other hand, shared knowledge systems must cope with motivated or clever users guessing answers to questions they do not know, and others forgetting answers they should know. Our approach uses social pressures and technical barricades, directed towards three classes of guessers:

1. Socially disconnected strangers and voyeurs that know little of the sharer or her friends have little information to guess with, so we *limit the number of guesses* that can be made.
2. Guessers with connections in the social graph have the resources to make better guesses, but face the counter-incentive of social disgrace if caught inappropriately guessing answers, which we leverage by *logging and displaying access attempts* to the sharer.
3. Friends who forget or mis-phrase an answer appear in logs with an interface to explicitly provide access. Alternatively, they might ask someone for the answer, since questions such as "where did our club eat" implicitly describe who *should* get access.

Although we do not require authenticated accounts, *guess limits* and *access logging* do need to know the guesser's identity. Depending on the incentives of users and attackers, a system might require identification from friend-confirmed accounts, regular accounts, or just IP addresses, providing varying levels of resilience to savvy, motivated users that create fake accounts. For instance, a Nike+iPod exercise log might need no more than IP addressess. As a failsafe, such a system can enforce a per-question global guess limit, blocking access until the sharer checks or changes the question. IP addresses can also be used to infer geographic locations for access logs, from which identity might be guessable, *e.g.* for an inquisitive ex who lives in San Diego.

Implementations must choose amongst these designs to fit their circumstances, striking a balance in the incentive structure between security and overhead of guard questions.

STUDY: DESIGN OF QUESTIONS FOR PHOTO SHARING

Since the summative effectiveness of shared knowledge security depends on its social context of use and these implementation decisions, our formative study instead probes the underlying issues. First, with whom do sharers want to show or hide their photos, and does shared knowledge *exist* to divide these groups? Second, what types of questions do sharers devise, and how difficult are they to design? Finally, how vulnerable are the questions to guessing, and do sharers anticipate the vulnerability? To answer the first two questions, we had participants devise questions for their own photos. To answer the third, we uploaded these questions as challenges to Amazon Mechanical Turk, and rewarded anonymous Internet workers to guess the answers.

Designing 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 within 10 guesses. Our participants were fairly diverse: 47/53% male/female, mean age 27 (stdev 8), recruited through flyers on two websites and in three urban neighborhoods. They completed the survey online and received $15 USD.

Results: Desired and Undesired Recipients

We clustered 315 responses of desired recipients and 401 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
Authority figures	21%	3.2	42%	3.0
Friends of photographed	34%	2.5	0%	--
Potential romances and employers	10%	3.5	7%	3.6
Ex-friends and romances	0%	--	14%	2.7

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

Demonstrating a need for flexible access control policies, 83% of participants had photos to blacklist from friends or family, which are commonly assumed to be *whitelist* groups on sharing websites. On average, people cared more about preventing access (2.6) than providing it (2.2) (p<.001).

Results: Questions Designed

Subjects easily understood the concept of guard questions, and could readily create them after reading a one-paragraph description. They designed 168 unique questions (and 11 duplicates), which we clustered into 6 categories in Table 2. Subjects successfully designed questions for all but 3 of the 179 photos, a 98% success rate, suggesting that there exists shared knowledge to separate most inclusion/exclusion groups (though we did not evaluate inclusion effectiveness). The median subject spent 8 seconds designing a guard question, according to self report. For comparison, it takes the first author 90 seconds to create a 10-person whitelist of

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 2: Categories of questions generated

email addresses using the Mac OS X address book. However, 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 reported 155 seconds on average over her 8 questions; her longest was 600 seconds. Future work should investigate the cause. We found no significant effect of design time on vulnerability to guessing.

Cracking the Questions

To learn how vulnerable questions are to guessing, we uploaded the questions 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. They were motivated with a bounty of $.75 for a correct guess within 3 guesses, and $.25 for one 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 designed the incentives to emulate those of unknown voyeurs (group 1; "strangers" in Table 1), with no connection to the sharer or their social network of shared knowledge. We plan to evaluate social relation (group 2) guessing ability in future work, using a field study to account for access logs and social pressures. We manually verified the quality of Turk guesses; a few poor responses were rejected, but the vast majority were of very high quality, *e.g.* showing clear evidence of clever thought and Web searching for answers.

As can be seen in Figure 2, Turk workers guessed correctly 6% of the time given 3 attempts, and 11% given 10. Thus, a guess limit of 3 could cut guesses roughly in half. It is unclear if guesses beyond 7 make much difference, but the data suggests their added value may taper off.

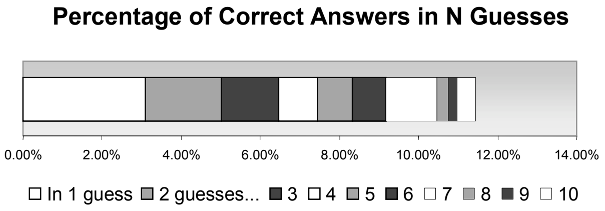

Figure 2: If we allow 3 guess attempts, strangers have a 6% chance of guessing correctly. With 10 attempts: 11%.

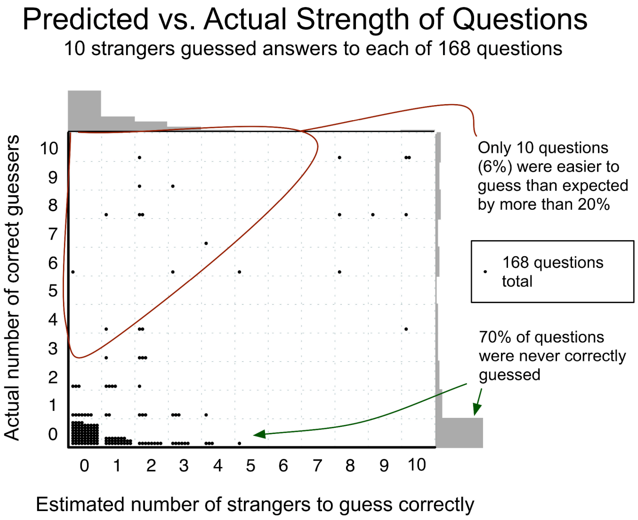

Figure 3: Scatterplot showing sharer ability to predict their questions' resilience to guessing. Histograms at top and right.

However, some questions may be intentionally easy to guess, since users might just want to reduce, not necessarily eliminate, access to a photo. In this case, users should at least be able to *predict* the ease of guessing their questions: understanding the breadth of disclosure is critical for privacy-sensitive systems [2].

We found the average subject has slightly better security (11%) than she expects (14%). We analyze this in more detail in Figure 3. The mispredictions are in the lower-right and upper-left. Of the 168 questions, only 10 (=6%) were less secure than expected by a margin of more than 20%. More common was to predict a few correct guesses for questions that could not be guessed. A linear regression gives R^2=.44 between coordinates. These assume 10 guesses.

We examined the 7 cases in the upper-left with the most unexpected guesses. 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 uses ontologies and Web searches to discover such weak questions automatically and suggest alternatives.

HANDLING AMBIGUOUS ANSWERS

Even users that know the answer may phrase it differently than the sharer. We designed a set of rules to verify ambiguous responses, implemented in a 37-line algorithm:

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

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

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

This algorithm 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, the question designer could specify whether a guess must *be* or *contain* the answer.

RELATED WORK

Many *personal* authentication systems require answers to tests of personal knowledge. For instance, Zviran studied personal authentication questions like "mother's maiden name" [5]. By using *shared* knowledge, these personal tests become group access control. For instance, personal photo knowledge authentication [3] can become an access control by incorporating a group, instead of personal, photo pool.

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 people, at anytime without redistributing passwords.

People naturally gauge one another with shared knowledge tests in real life. We have also found ad-hoc uses on the Web, where a traditional login & password page is accompanied with instructions such as "username perry and the password is our school mascot (in lower case with an s at the end)". Our work formalizes this idea and presents a design and study to broaden its viability.

Recent research has worked on the usability of operating systems access control lists. See Cao [1] for an example.

FUTURE WORK

As mentioned earlier, our formative study does not attempt to evaluate the real-life access rates of social relations and friends (groups 2 and 3), since so many real-life and implementation variables influence their behavior. Informed by the results of the present work, we are currently building a real system to evaluate access and user acceptance in field trials.

There are many potential avenues to reduce error rates, both through interaction and analysis, such as better visualizations of guesses and guessers; interfaces 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 free-form text to avoid paradox of choice and weak question types; cognitive analysis of systematic guess rate underestimates; 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, café wifi access, realtime location data streams, automatically moderating mailing list subscriptions, subgroup CAPTCHAs, and group project Wiki access control. Guard questions could also be combined with traditional access controls in interesting ways. For instance, one might use a guard question over a hidden blacklist to add plausible deniability.

CONCLUSION

We present a type of access control where concise tests of shared knowledge replace accounts and access control lists. Users readily learn the concept, and design most questions with little effort. Most questions are secure to guesses from strangers. Users can generally predict the security of their questions, but sometimes underestimate the ability of attackers to use Web searching or enumeration to discover answers. By lowering the threshold to access control, shared knowledge tests could enable more types of information to acquire collaborative value on the Internet.

ACKNOWLEDGEMENTS

We thank Tadayoshi Kohno for his advice, and Eytan Adar, Jiun-Hung Chen, Jon Froehlich, Susumu Harada, Tapan Parikh and our helpful reviewers for their detailed feedback on the paper. The first author was supported by a National Science Foundation fellowship, and the second by a grant from the Department of Homeland Security. This work was inspired by a feature of Jonathan M. Hsieh's personal webpage, and by the tribulations of the first author's mother named Susan, who had a hairy dog named Bernie.

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.
5. Zviran, M., Haga, W.J. (1990). *User Authentication by Cognitive Passwords: An Empirical Assessment. Jerusalem Conference on Information Technology*, 137-144.