Social Access Control for Social Media Using Shared Knowledge Questions

Michael Toomim[1] and **Xianhang Zhang**[2] and **James Fogarty**[1] and **Nathan Morris**[3]

[1]{toomim,fogarty}@cs.washington.edu
Computer Science & Engineering
DUB Group
University of Washington
Seattle, WA 98102

[2]xianhang@u.washington.edu
Human Interface Technology Laboratory
DUB Group
University of Washington
Seattle, WA 98102

[3]nathanms@u.washington.edu
Chemical Engineering
DUB Group
University of Washington
Seattle, WA 98102

Abstract

Managing privacy of online content is difficult. We present a simple social access control where sharers specify test questions of shared knowledge, such as "what is our school mascot," instead of creating authenticated accounts and specifying explicit access control rules for all potential accessors. This demo will let attendees interact with our Facebook prototype. We will also explain prior studies that elucidate 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.

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?

We propose that sharers design guard questions of shared knowledge such as "what is our school mascot" or "where did I travel this summer" that must be answered to view a photo or album, leveraging the shared knowledge preexisting in social networks (Figure 1). We observe 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.

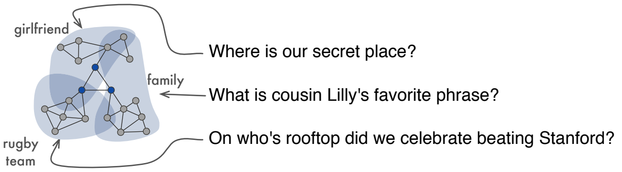
girlfriend
Where is our secret place?
family
What is cousin Lilly's favorite phrase?
rugby team
On who's rooftop did we celebrate beating Stanford?

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

In prior work (Toomim *et al.* 2008) we studied the design and security of guard questions and presented a simple algorithm to verify ambiguous responses. In this demo attendees will interact with our prototype that protects Facebook content.

Problems with traditional access control

Traditional access control uses variants of whitelists and blacklists, requiring users to explicitly translate social relationships into lists or groups of account names and/or email addresses to be allowed or denied. They are 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 (Tajfel *et al.* 1971). 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.

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

Figure 2: Desired and undesired people to see photos, as described by participants. *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.

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 like an ex-boyfriend.

On the other hand, more expressive grouping mechanisms, such as UNIX groups, become, 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.

Preventing guesses from unintended users

It is reasonable to expect some motivated users to guess answers to questions. They might covertly persuade the sharer’s friends to reveal hints, or brute-force guess a question with a finite set of choices such as “what color is my car?” Our application mitigates such attempts with two mechanisms. First, hard *guess limits* hinder brute-force attacks. Second, *access logs* record and display the guessers to the sharer, creating social repercussions *e.g.* for friends that convince relations to leak an answer they were not supposed to know. The access log also displays friends that forget answers, so the sharer can whitelist them. Although we do not require authenticated accounts, the implementation of these features requires some knowledge of the guesser’s identity, and we discuss tradeoffs amongst three levels of identification in (Toomim *et al.* 2008).

Study

We ran a study to learn to whom users want to grant or deny access, the types of questions they use to divide these groups, the basic resilience of the questions to attack from adversaries without access to social knowledge, and user’s abilities to predict this resilience. This section summarizes our results from (Toomim *et al.* 2008).

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%

Figure 3: Categories of participant-designed questions

We first collected data on the groups of people with whom people want to share and not share photos, and how important they are. 31 participants found 179 personal photos and reported who they wanted to and not to see each photo. We clustered the responses in Figure 2. Demonstrating a need for flexible access control, 83% of participants had photos to blacklist from “family” or “friends”, which are commonly assumed to be *whitelist* groups in sharing websites.

We then had participants design questions to protect each of these photos. We clustered their responses in Figure 3. Participants were able to find questions that implement their inclusion/exclusion preferences for 98% of the photos, indicating shared knowledge exists to represent most privacy situations. It took a mean of 15 seconds to design a question, with standard deviation of 28. For comparison, it takes the first author 90 seconds to create a 10-person whitelist of email addresses using the Mac OSX Address Book.

Finally, we uploaded the questions as jobs to Amazon’s Mechanical Turk, and rewarded anonymous Internet users to guess the answers. Guessers had a 6% chance of cracking a question in 3 guesses, and only 7 of the 168 questions (4%) were more than 30% easier to crack than sharers estimated.

Conclusions

Questions of shared knowledge are a lightweight alternative to traditional access control. Their security is enforced socially as well as technically. Our demonstration of this system in a real Facebook application will give attendees a feel for what it is like to be a sharer or accesser in this model.

References

Tajfel, H.; Billig, M. G.; Bundy, R. P.; and Flament, C. 1971. Social Categorization and Intergroup Behaviour. *European Journal of Social Psychology* 1(2):149–177.

Toomim, M.; Zhang, X.; Fogarty, J.; and Landay, J. A. 2008. Access Control by Testing for Shared Knowledge. In *CHI ’08: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, to appear.*