Not That Kind of Friend: Misleading Divergences Between Online Social Networks and Real-World Social Protocols (Extended Abstract)

Jonathan Anderson and Frank Stajano

University of Cambridge
Computer Laboratory

Abstract. Current social networking technology provides users with little of the privacy that they expect and are promised. While some individuals may truly wish to be "open books," violating most users' privacy expectations leads to serious consequences for both the users and their network. We propose that, in order for users to truly be in control their personal information, we need to respect and learn from the social protocols that have been built up over the millennia.

1 Introduction

Social networking technology is used by hundreds of millions worldwide to help them connect to and share information with others. Sometimes, however, the information they share reaches a wider audience than they intend [1,2]. We believe that problem this stems from two primary sources: the implicit trust that must be placed in network operators and the exceptionally poor quality of users' access control.

In this paper, we propose a methodology which addresses these problems and seeks to truly put users in control of their personal information. We describe how a decentralized architecture, informed by the semantics of real-world social interaction, could provide the technical means for users to share information in a manner which they already understand how to manage. Rather than teaching users about the ins and outs of computer security, we should rely on implied social contracts; they are the foundation of social interaction and no technical details should obscure this. We will not make false promises about revocation to lure users into a false sense of security, nor will we use arbitrary constructs like friendship lists to provide incentives that do not exist in the real world.

2 What's Wrong With Social Networks?

Social network technology exists to assist people in connecting with others and managing their interpersonal relationships. The former, taken in isolation, is very easy. The latter, however, requires delicacy, confidence and privacy. Put another way[1], Facebook has the happy talent of helping you find friends; whether it is equally capable of helping you keep them is less certain.

Aside from obvious implementation details, we believe that there are two essential problems with today's social networking technologies:

1. network operators require implicit trust, and
2. users don't manage access control well.

[1] with apologies to Jane Austen

2.1 Operator Trust

Most users don't think very hard about it, but using a social network ought to be a sobering activity. You tell a third party information that they can use to impersonate you [3], then tell them who all of your friends are – what keywords would be effective in a phishing attack against you [4], then post information which, should it reach an unintended audience, could get you blackmailed or fired [5,6,7].

Social network operators have databases of such information for millions of users. The value of this information is obvious, and the only thing keeping it from being stolen *en masse* by identity thieves is the competence and good will of the network operators and their employees.

The first of these qualities – competence – can be assumed to vary across the many social network operators in the market. The second quality – good will – is different, though. Many networks assert the right to share whatever information they like with whatever parties they deem trustworthy [8], and the only paying customers are their advertisers. In this light, consider a choice before a hypothetical social network:

1. spend time and money securing personal information against unauthorized access by corrupt insiders, or
2. spend time and money exposing personal information to advertisers to increase the value of their ads

Which is more likely to be chosen?

I am not claiming that major social networks have been overrun by identity thieves or that they don't care about their reputation. I *do* assert, however, that the requirement for absolute trust in a large network operator is worrisome and there may be a chilling effect among some users: I'd share a lot less information with friends if I knew that strangers were recording a transcript of all my conversations.

2.2 Access Control

It is immediately apparent that a social network must allow information sharing in order to be useful. It is slightly less apparent that this sharing often depends on the *assumption* of effective access control – I share photos with my friends and implicitly assume that my boss won't be looking at them. The system falls down when the assumption is not justified – as has been shown time after time on Facebook – but the system also fails to work effectively when its users allow the fear of poor access control to prevent them from sharing information.

Why isn't access control effective in social networking? A large part of the problem is that social networks are fun and easy to use, but their access control schemes are tedious and incomprehensible.

Tedious People use social networking technology for the sake of their social relationships, not for the purpose of spending time in front of a computer. Faced with a choice between a) finding friends and chatting or b) entering meta-data and managing ACLs, most users choose option (a) [9].

Incomprehensible Many of the implications of security policy decisions are not immediately apparent to users. These implications include problems with both fundamentals (e.g. a trust model with false in-/out-group distinctions or which assumes trust is transitive) and details (e.g. policy interactions that mystify PhD students in computer security).

3 What Can We Do About It?

How can we build a social network that doesn't succumb to these problems? Many data protection policies pay lip service to the concept of "putting people in charge of their own information," but we should actually *do it*: rather than "user control" consisting of a policy that users trust the system to follow, personal data ought to be protected such that if something isn't *authorized*, it isn't *possible*. Such a policy could be achieved by addressing the above problems of operator trust and poor access control, and we believe that privacy in online social networks could be greatly improved by observing real-world social protocols.

The human brain seems to be hard-wired to handle a fixed number of social relationships [10], and we have spend a very long time learning how to manage our social interactions with them. The great hubris of social network operators seems to be believing that they've fundamentally changed the way that humans beings interact, and we can see the results of this hubris played out all around – some people make bad access control decisions, but many people are afraid to use social networking sites and lose out on their potential benefit.

If we want to make a social network that behave as people expect, let's pay attention to the social understanding that people already have. Instead of trying (and failing) to teach users to think like a computer, let's train the system to figure out what the user wants to do [11] by following the semantics of real-world social relationships.

3.1 Centralization

One of the most obvious difference between online and offline social networking is the degree to which each is centralized. In the real world, no third party is required for you and I to be friends[2], but this model hasn't been picked up by social network operators. Instead, we have the all-trusted network operator. We can deal with the "trusted operator" problem in one of two ways:

1. stop trusting the operator or
2. stop using an operator.

The former solution has been proposed (e.g. "keep using Facebook, but encrypt everything in your profile"), but there is an economic problem with this model: you take away the central party's business model, namely the ability to serve highly targeted advertisements. This makes your system a *parasitoid*: a parasite which incubates within its host during its larval stage and eventually kills the host [12]. If your system is successful, it will kill the centralised network; if the network is smart, it will kill your system first.

This leaves us with the decentralized option, which brings us back to something which looks a lot like the offline protocol.

3.2 Social contracts

Social networks need to make us feel safe enough to share our personal information (exactly *how* safe depends on the individual – some people will admit criminal guilt via public broadcast, but many would not). In order to accomplish this goal, network operators often provide "locks and bars" that make security feel like a technical problem, but security is fundamentally a *human*

[2] You could argue that a communications infrastructure is required in many cases, but it's not always true, and people can choose a wide range of communication options (phone, letter, "when you see Jon, tell him..." etc).

problem, and we would do well to consider the real underpinning of human relationships: social contracts.

In real-world social networks, there are no technical barriers to my close friends running around and telling others my most personal secrets, so it shouldn't be surprising that the same behaviour is possible in an online setting. The only thing that really prevents this type of behaviour is an implied social contract: if I tell a close friend a secret, I expect them to keep it. If I tell a stranger the same secret, I have less expectation of them keeping it. If I allow a computer to share the secret with anybody in a group whose membership we don't really understand, however, then somehow we expect the secret to be kept.

We need to remove this layer of obscurity and clearly admit that social network security is based on an implied social contract, and that social contracts are not externally enforced [13]. Anything more is a promise we can't keep.

3.3 Revocation

One component of this obfuscation layer is the concept of permission revocation. If I tell you a personal secret, there is no way for me to make you forget (that doesn't involve a heavy stick). I can ask you to keep it secret, I can hope that you forget it and stop doing things that might remind you, but I cannot cut into your brain and remove the information. Current social networks, however, provide an interface which gives the appearance of perfect, retroactive revocation without actually backing it up.

This appearance of perfect revocation leads users to post things which they "can always delete later", but which in reality may never disappear entirely from the network. I can check a photo's ACL and see that only my University friends are allowed to view it, but if family members were ever allowed to see it, then who knows which browser cache or physical album the photo is tucked away in?

Better would be a system where the only revocation that is offered is *future* revocation. Such a system would serve the twin purposes of stripping the glossy veneer off of inherently risky activities and allowing users to view not just an optimistic guess at who we *hope* can view information, but an accurate picture of who *actually* can see it.

3.4 Friendship lists

One of the main features of social networks – friend lists – may be their undoing in the long term. Such a concept simply does not exist in the real world, and there are several reasons why I think that pretending it does is a bad idea. These reasons include:

False Dichotomies Social relationships are complicated affairs, but social networks would have us believe that they are binary in nature, and usually bidirectional. In real life, before sharing a secret with anyone, I think about my relationship with them, how reliable they are, etc. In a social network, however, the only easy-to-use access control is usually the question "is this person in my friend list or not?" Such a false dichotomy – either you're in my "inner sanctum" or the public at large – can lead to poor access control decisions.

Dilution A friendship list can be viewed as a list of people with whom a greater degree of intimacy is warranted than with the general population. Often, "all friends" is the most restrictive access control that can be applied without significant effort (e.g. ACLs), yet a common complaint about social networks is that "your mother is on Facebook," so the friendship list doesn't mean what it once did. The larger and more inclusive a list gets, the less it differentiates people from the general population; conversely to how we usually think about network effects, increasing the size of a friendship list decreases its overall value.

Un-friending In real life, "un-friending" somebody is easy: you stop inviting them to parties, revealing personal information, etc. until you've drifted apart. Such pruning should be easy, because it is necessary – the human mind can only accommodate so many close relationships. On a social networking site with hard binary relationships, however, letting a relationship go requires users to state, in effect, "I no longer value my relationship with Alice." Increasing the internal complexity of pruning friendships leads to a stale picture of current social relationships and access control decisions that no longer reflect the user's intentions [3].

Metric Gaming If software engineering has taught us anything, we should know that introducing arbitrary success metrics (e.g. lines of code, bugs fixed) leads to actors maximizing their own "stats," often to the detriment of the larger group. Introducing the concept of a friend count incents users to make "friendship" links which they otherwise would not have made, as well as keeping links which they otherwise would not bother to keep. There are Facebook users with thousands of "friends" and MySpace users with millions; how many of these are actual friends, and how many of these "relationships" exist solely as a pair of mouse clicks between strangers? Keep in mind: this is a realm where "Only Friends" is as inclusive as many ACLs get.

3.5 Sharing Defaults

In the real world, ideas and secrets start out hidden away in one's grey matter, then are shared with friends via an act of will. Photographs follow a similar progression, with images being captured by a camera, developed (or downloaded) and then shared by the photographer's effort via paper and/or e-mail.

Since maintaining ACL metadata is tedious at best, current social networking technologies prefer to share everything by default. This can lead to unwanted information being shared, and is only mitigated by the [partial and ineffective] revocation promised by the network operators. We need not follow this poor practice, however, as long as we are not slaves to its precondition (poor access control). We can make a system whose default setting is "private" as long as sharing information is easy and understandable.

4 Conclusion

It we want to build a social network that people can use effectively while maintaining privacy, we should pay heed to the lessons learned over the last several thousand years of recorded civilization. Online social networking is not a fundamentally different thing from offline networking, so we should build virtual systems that work something like the real world.

An appropriate level of abstraction for social networks deals with concepts like friends, photos and events; we ought to hide details like transport protocols and encryption keys from users. It is *not*, however, appropriate to try to protect users from themselves by hiding social facts like "there is no such thing as true revocation." On the contrary, social networks should, to the user, look like the real world: decentralised and based on implied social contracts, with no concepts of perfect revocation, friendship lists or permissive defaults. If we make the computer network behave like real social networks, we just might provide a platform that people can safely use with confidence.

[3] In fact, such staleness is even bad for advertisers, who rely on the high accuracy of information in social networks in order to target their ads accurately!

References

1. R. Bennett, "Plea to ban employers trawling Facebook." http://technology.timesonline.co.uk/tol/news/tech_and_web/article3613896.ece, 25 Mar 2008. The Times.
2. J. Shepherd and D. Shariatmadari, "Would-be students checked on Facebook." http://www.guardian.co.uk/uk/2008/jan/11/accesstouniversity.highereducation, 11 Jan 2008. The Guardian.
3. A. Rabkin, "Personal knowledge questions for fallback authentication," in *Proceedings of the 4th Symposium on Usable Privacy and Security - SOUPS 08*, pp. 13 – 23, ACM, 2008.
4. T. N. Jagatic, "Social phishing," *Communications of the ACM*, vol. 50, no. 10, p. 94, 2007.
5. E. Pilkington, "Blackmail claim stirs fears over Facebook." http://www.guardian.co.uk/business/2007/jul/16/usnews.news, 16 July 2007. The Guardian.
6. D. Randall and V. Richards, "Facebook can ruin your life. And so can MySpace, Bebo...." http://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-can-ruin-your-life-and-so-can-myspace-bebo-780521.html, 10 Feb 2008. The Independent.
7. CBC News, "Student recruits unfit for service, say former border guards." http://www.cbc.ca/canada/british-columbia/story/2007/10/01/bc-borderguards.html, 1 Oct 2007. Canadian Broadcasting Corporation.
8. "Facebook Privacy Policy." http://www.facebook.com/policy.php, Dec 2007.
9. A. Whitten, *Making Security Usable*. PhD thesis, Carnegie Mellon University, 2004.
10. L. C. Aiello and R. Dunbar, "Neocortex Size, Group Size, and the Evolution of Language," *Current Anthropology*, vol. 34, no. 2, pp. 184 – 193, 1993.
11. K.-P. Yee, "Aligning security and usability," *IEEE Security and Privacy Magazine*, vol. 2, no. 5, p. 48, 2004.
12. H. C. J. Godfray, "Parasitoids," *Current Biology*, vol. 14, no. 12, p. R456, 2004.
13. K. Binmore, *Game Theory and the Social Contract, Volume 2: Just Playing*. MIT Press, 1998.