Saveface: Save George's faces in Social Network where Contexts Collapse

By Fuming Shih, Sharon Paradesi
fuming@mit.edu, paradesi@csail.mit.edu
MIT

1. Introduction

Social networks have repeatedly perplexed Web architects with the dilemma of privacy versus publicity. Is it ever possible to protect privacy on a platform that exists and is designed to share user's information publicly? Despite the efforts of increasing privacy controls for its users, Facebook has continuously made the headline news concerning privacy issues. Seemingly, more controls to a user's content does not afford a satisfactory solution to diminish privacy risks [1]. Studies have shown that it is unreasonable for a user to foretell all the possible consequences of information disclosure and set an ex ante rule about privacy. [3]. The real challenges about privacy protection in social networks, as some researchers argue, are the dynamics of social contexts [2], and the deficiency of current architecture in utilizing contextual information [4]. Through this paper, we demonstrate a need to create a sustainable architecture to create and implement viable privacy awareness measures and practices for different actors within the social network 'eco-system'. Also, by monitoring transfer of information and social actions within the system, the system would be able to suggest the policy decisions being made to the user.

The rest of the paper is structured as follows. We explain the meaning of context within social networks, the limitations in the current social networks and finally highlight research directions that will help in implementing a sound privacy infrastructure. Our paper focuses solely on Facebook, but we envision that these principles are universal in nature.

2. Contexts in Social Networks:

In Social Networks, contexts imply a content's original intent, the original audience addressed, content's object references, the original activity or practice; and context tied to social or public in which the content is produced [5]. In general, when a user reads a piece of information, she usually places it within an implicit context in order to interpret it. Multiple contexts together act as a "container" of the disclosed information to differentiate it from other similar kinds. Without defining contexts, the consumers of the information have no reference to appropriately evaluate the message conveyed. In the famous Seinfeld episode *The Pool Guy,* George Costanza [6] was upset when confronted with the situation of his new girl friend infiltrating his group of friends. He has successfully kept his relationships in independent silos, and he does not expect they will mix together. Same with the Facebook users today, when their moms and bosses "friend" their ways to access a user's content, privacy concerns arise due to the lack of context supported in the Facebook architecture.

3. Current limitation - lack of context in a social network:

Neglect of contexts when processing information could cause systems to disregard user's expectations and thereby raise serious privacy issues [8]. In the real-world practice of privacy management, companies have gradually shifted their focus from content to the context within which the user's data is present. This change aims to make appropriate use of users' data to meet their expectation [9]. To prevent users from experiencing *George's panic*, the social network should incorporate user's contexts explicitly so that application that plug into it would treat information per user's specifications. On the other hand, only a user him/herself could best evaluate the privacy risk based on the effects brought by the change to his/her context.

For example, George could specify a policy stating that all comments attached to pictures from or tagged by certain friends (those he parties with, for instance) should not be used for employment decisions or dating recommendations. Once George states this context within which his data should be interpreted, smart Facebook applications that implement "employee-finder" systems or other external programs that mine Facebook data [10] should bypass certain pieces of information when aggregating data about George that they would otherwise utilize. By setting such restrictions on contexts, George will still be able to receive personalized services like the targeting ads but is now able to state when certain information cannot be used against him.

Such a mechanism is currently lacking in Facebook and we think that the directions described in the next section would lead to the development of such a more robust infrastructure.

4. Directions toward a more efficient privacy setting in social networks:

a. Creating an awareness about contexts in users

A formal description of "contextual integrity" developed by privacy law researcher Helen Nissenbum, introduced inextricable relation between privacy and context [4]. However, in practice of social network, it is unreasonable to ask the user to manually identify the relevant contexts for their privacy concerns. To alleviate the user's burden, the author in [7] proposed an privacy mechanism to learn a user's privacy policy based on context inference. In this paper, we propose a different tool with similar ideas to help a user explore relevant contexts to raise the awareness of possible privacy risks. In the project Saveface, we are developing a game-like interface for people to retrospectively recall privacy settings that are in conflicts with their expectations. Say, George updated his friend list to include Susan. We can give him a warning about the content that will be displayed to Susan. Give George a preview about how his worlds will collide. In other words, the content she can see before becoming his friend and after. This reflects the current access control policy of a typical Facebook user. Now George can choose to create a new group for Susan. We are experimenting how much content is necessary to expose to the user to raise awareness in the mindset of a user.

b. Need for more flexible access controls

Most of the current social networks provide some level of access control. For example, Facebook and Flickr provide default settings of control where a user can restrict certain people from viewing his/her content. However, in reality, having well-defined groups of friends is a very naive assumption from the perspective of the dynamism of human relationships. A more reasonable way would be to enable users to create dynamic groups based on attributes that matter to the user. These attributes would then form the context within which the friends, or even

strangers, would be placed.

For instance, say that George uploads content on Facebook hoping to reach a large audience to convey some message. This audience could be larger than any group he currently has but he does not want to manually create a new list of friends just for this update. In such a scenario, a flexible way would be to obtain the attributes that George wants his audience to have and automatically create such a list for him.

c. Augmenting access control with usage restrictions

Relying solely on access control can only get you so far. In extreme cases, it only makes users paranoid about who is viewing and using their information. Currently, one can only restrict access his or her content, but has no say over how that content will be used. We would like to see the infrastructure progress to an ideal world where everyone can exercise their freedom of speech. To reach such a goal, users should be able to attach usage restrictions of how they want the data to be interpreted and used by others.

In short, if Elaine uploads content about George on her Facebook account and tags George, all his usage restrictions will be applicable on this content, as they would carry over.

d. Standardizing usage restrictions across multiple networks

Implementing usage restrictions in a single network seems straightforward. However, the real challenge comes when we need to reconcile these restrictions across multiple networks and media. The crucial aspect is that the restrictions should be tied to and travel along with the piece of information.

Example:
Continuing with our example of George, say that he has a friend on Flickr called Elaine. Elaine also has a Flickr account. Suppose that George created a usage policy on his Facebook account saying that no content from his account can be used for targeted advertising. Similarly, Elaine created a usage policy on her Flickr account saying that certain albums cannot be used to make employment decisions for her.

Mr. Steinbrenner, George's boss, wants to make a decision about George's performance review. Mr. Steinbrenner is friends with George on Facebook and Elaine on Flickr and is browsing their respective content. A typical accountable system on both networks would behave as follows.

Rewrite the example from this standpoint
==========
When Aintno queries George's social graph, a typical accountable system would infer that a health insurance decision is not targeted advertising and thus grants access to George's content. Aintno sees that George has a friend named Jo and looks up Jo on other social networks. In due time Aintno comes across the Flickr account of Jo and queries her profile. An accountable system at this end would reason that this insurance decision is not being made for Jo and therefore grant Aintno access to all of Jo's content. While browsing through Jo's Flickr photos, Aintno comes across one that highlights some poor choices made by George. This leads to Aintno rejecting George's health insurance application.

In the above scenario (plausible in the near future), no usage policy has been explicitly violated. However, Aintno was shown data that they shouldn't have seen/used in a particular context. This is something we need to guard against.

Issues:
- Discrepancies between policies and their meanings across different social network and media.
- Ability to carry a usage restriction over from one network to another.

5. Discussion:

With the proliferation of data been created specifically from social network platforms, the user faces a plethora of mixed messages with different quality. When a piece of information is created under certain context, some is just gossip with friends and some is a fact stated for specific purpose. During the process of information, both data consumer and creator need a environment to guide privacy practices. For privacy controls, both defaults by the system and finer grained settings from users have proved to be problematic that fail to meet user's expectations. Apparently, most access control mechanisms are not sufficient for an dynamic and evolving environment like the social network. To improve, we argue that "context" is the missing piece in the architecture that technology should exploit to raise the awareness in the user. The system should treat privacy as a function of user's expectations in particular attributable contexts [9]. Also, privacy protection through usage restriction....[TBD]

6. References:

[1] L. Kagal and H. Abelson. Access control is an inadequate framework for privacy protection. In W3C Privacy Workshop, 2010.
[2] D. Boyd 2008. Facebook's Privacy Trainwreck: Exposure, Invasion, and Social Convergence. *Convergence*, 14 (1), 2008.
[3] A.M. McDonald, R.W. Reeder, P.G. Kelley, and L.F. Cranor. A comparative study of online privacy policies and formats. Privacy Enhancing Techonologies Symposium 2009.
[4] A. Barth, A. Datta, J. Mitchell, and H. Nissenbaum, Privacy and Contextual Integrity: Framework and Applications, Proceedings of the IEEE Symposium on Security and Privacy, 2006.
[5] A. Chan, Social context, Facebook Likes, activity and action streams, entry posted April 27, 2010, http://www.gravity7.com/blog/media/2010/04/social-context-facebook-likes-activity.html
[6] Independent George - Worlds collide, http://www.youtube.com/watch?v=uPG3YMcSvzo
[7] G. Danezis. Inferring privacy policies for social networking services. In AISec, 2009.
[8] The Washington Post, Google Buzz Privacy Issues Have Real Life Implications . http://www.washingtonpost.com/wp-dyn/content/article/2010/02/12/AR2010021201490.html
[9] K. A. Bamberger and D. K. Mulligan, Privacy on the books and on the ground. Stanford Law Review 63 (2010). Accessed 14 May 2010.
[10] Datamation, 'Pre-crime' Comes to the HR Dept. http://www.socialintelligencehr.com/home
[11] Z. Wu, L. Wang: Enforcement of Privacy Policies over Multiple Online Social Networks for Collaborative Activities.SCSS 2009: 583-588

[12] M. Helft, Facebook Acknowledges Privacy Issue With Applications, entry posted October 18, 2010, http://bits.blogs.nytimes.com/2010/10/18/facebook-admits-to-privacy-issue-and-makes-fixes/?hp

7. Acknowledgements:

The authors would like to thank Joe Pato (from HP labs and currently at MIT) and Michael Speciner (from MIT) for their insightful comments and discussion.