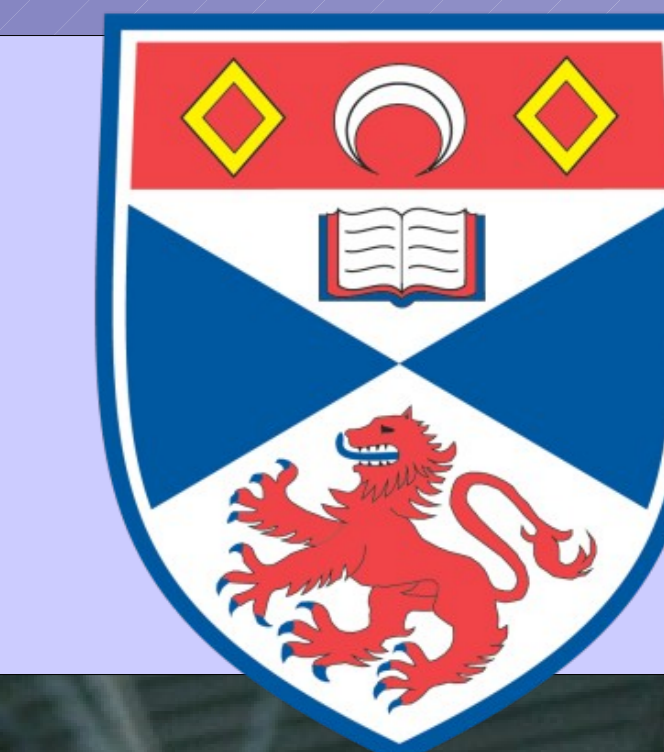


Users' Privacy Concerns in Online Social Networks

Iain Parris <ip@cs.st-andrews.ac.uk>

Supervisor: **Tristan Henderson**

Funded by EPSRC Research Grant EP/G002606/1



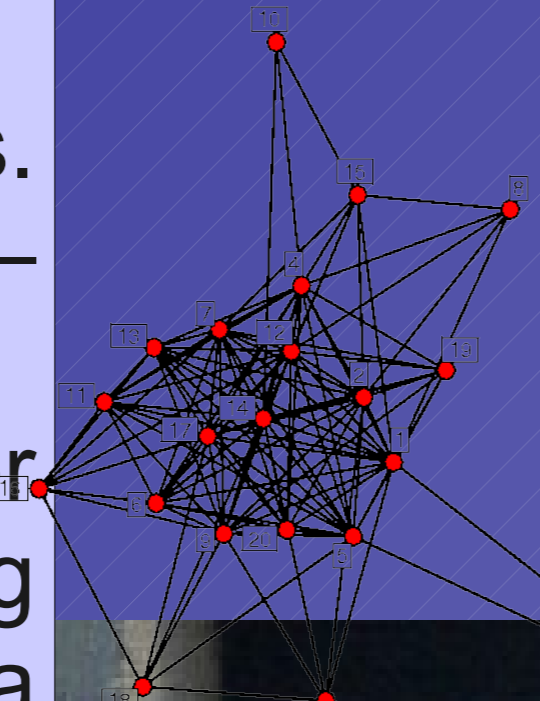
Introduction

In recent years, **online social networks** – such as Facebook, MySpace and Bebo – have increased dramatically in popularity.

Numerous privacy threats are associated with privacy networks. However, relatively little work has been done to assess these threats – and in particular, users' concerns.

A lack of understanding of these threats can be costly: consider Facebook's *Beacon* system, which displayed potentially-embarrassing data of purchases on other websites on people's Facebook profiles. As a result of a large, and unexpected, user backlash, this system was made *opt-in*, instead of *opt-out* as before. Had it not, the financial impact on Facebook (due to users leaving) may have been dramatic.

Two key questions, thus, are: **what are the privacy problems in social networks**; and: **which of these problems are of concern to users?** Specifically, **can we model users' concerns?**



Expected results

- A new understanding of the **contexts in which people want to share private data**, and those in which they don't.
- Based on this new understanding, **enhancing privacy-aware social networking applications**. Example potential enhancements may include:
 - Developing a new, simple, intuitive user interface for managing privacy settings
 - Building an underlying privacy-preserving infrastructure for social networks – thus allowing the development of higher-level technologies that leverage social networks, without having to worry about privacy
 - Utilising smart phone sensors to intelligently determine what the user's expectations of privacy are in the current context

Current status

- Working on a **threat analysis** of privacy problems. This will be used during user studies to determine which problems concern users.
- Developing methods for anonymising a social network graph for use in “social routing”. The efficiency of using the graphs, anonymised in different ways, may then be studied.

Background: Social networks and sensors

Online social networks are hugely popular, and attracting an ever-growing user base. At the same time, there has been a rise of **smart phones** – phones carrying sensors, such as *Global Positioning System (GPS)* sensors which can pinpoint location.

Might the two be linked?

Social networks and sensors may be used together, to enable new applications. For example, if a person wanted to find which of their *friends* (people in their social network) were within walking range of their current position, this may be facilitated by social networks with sensor input.

This has begun (with services like *dodgeball.com*, *brightkite.com*), but isn't yet mainstream. With such sensitive data, and with phones likely to carry even more sensors in future, **privacy** now becomes an even higher priority.[1]



Expected methodology for measuring users' concerns

When asked broad and imprecise questions, such as *Are you concerned about privacy?*, little useful information is obtained. Even worse, there is a disconnect between what people say and what people do: despite being “worried” about privacy, a person may pragmatically choose to use an online social networking site because the benefits outweigh the privacy concerns.[3]

How can we obtain useful information about privacy, therefore?

One core problem is that people will tend to respond differently to the same question in the lab to **in situ**: out “in the real world”. Another is that privacy concerns are **context-specific**. For example, a person may not be worried about their location at work being shared with others, but may wish to keep their home address private.

A possible approach is to use the **Experience Sampling Method** – which is a way to get responses from people *in situ*[4]. We ask the person a question requiring an immediate answer, while they are performing the task. Using **mobile phones**, this becomes possible and fairly straightforward.

Rich, relevant data may thus be collected. Even better, the mobile phones' **sensors** can be exploited to ask **relevant questions**.

The problems: Potential privacy problems

What sort of privacy problems may arise?

Examples of data that a person may wish to keep private are:

- **who** they are friends with
- **how many** people they are friends with
- **groups** of friends (such as not wanting to admit that they socialise with computer scientists)

Safeguarding such data may be non-trivial. **Anonymity** is hard to ensure when releasing data about social networks; just removing names is insufficient, because the **network structure itself reveals clues** as to the identity of the participants.[2]

References

- [1] D. Anthony, et al. (2007). 'Privacy in Location Aware Computing Environments (PLACE): How users' social context influences willingness to share location information'. *IEEE Pervasive Computing* 6(4):64-72.
- [2] L. Backstrom, et al. (2007). 'Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography'. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pp. 181-190, New York, NY, USA. ACM Press.
- [3] C. Jensen, et al. (2005). 'Privacy practices of Internet users: Self-reports versus observed behavior'. *International Journal of Human-Computer Studies* 63(1-2):203-227.
- [4] S. Consolvo & M. Walker (2003). 'Using the experience sampling method to evaluate ubicomp applications'. *Pervasive Computing, IEEE* 2(2):24-31.

Photos used under the *Creative Commons* license:

- "CCTV Heads – d*base" [by Joffley: <http://flickr.com/photos/joffley/>]
- "Cyber-shot cellphone "W61S" (2008)" [by mujitra: <http://flickr.com/photos/mujitra/>]
- "Satellite, Ariel I, Reconstructed Satellite" [by cliff1066: <http://flickr.com/photos/nostri-imago/>]