Privacy in ubiquitous computing

Lots of questions, a couple of answers

Mika Raento

mika.raento@cs.helsinki.fi

Helsinki Institute for Information Technology – Basic Research Unit

Introduction

What is privacy?

- Bodily (body searches)
- Territorial (home)
- Communication
- **Information**

Themes – applications

- Presence services
- User modelling
- So wanting to collect either for distribution to others or analysis

Themes – constraints

- User attitudes/needs
- User behaviour
- Legislation

Privacy principles

Marc Langheinrich, Privacy by Design Principles of Privacy-Aware Ubiquitous Systems, Ubicomp 2001 Proceedings

- Openness and transparency: subject aware
- Individual participation: subject can see and modify records
- Collection limitation: not excessive for purpose
- Data quality: relevant, correct and up-to-date

Privacy principles

continued...

- Use limitation: only for stated purposes, access controls
- Reasonable security: relative to data collected
- Accountability: subject able to verify compliance

Reflected in EU and US legislation.

Issues in ubicomp

Openness and transparency

- Systems are supposed to be invisible
- How can the user be aware of when data is being collected, and what data
- Legal issues: getting consent for all collection
- User issues: how can we enable the user to have an accurate mental model of the systems' working

Individual participation

- System cannot function as a black box
- What about inferred data: models, predictions
- How can the user correct a model built by the system?

Collection limitation

- We want to build systems that (maybe) use as much data as possible, without necessarily knowing how relevant attributes are
- Good motivation for finding out relevance of e.g. presence information attributes!
- What about length of history stored?

Data quality

- How do we show that inferred data or models are accurate?

Use limitation

- How do individuals give permission to distribute data to others?
- Legally and practically

Reasonable security

- Maybe ’reasonable’ doesn’t have to be very much in a research setting
- **If a presence service distributes data to other general-purpose computers there is no way of limiting where that data ends up**

Users

User preferences vs. behaviour

Spiekermann, Grossklags, Berendt (2001) Stated Privacy Preferences versus Actual Behaviour in EC environments: a Reality Check, Proc 5th Int Conf Wirtschaftsinformatik

- 75% of people studied were concerned about their privacy or commercial profiling (30% ’privacy fundamentalists’)
- 87% of participants disclosed large amounts of private information in exchange for uncertain, smallish financial gains
- The exact numbers aren’t necessarily interesting, but the study shows that people do not act according to their stated preferences

User preferences vs. behaviour

- Affects suitability of research methods
- True/well simulated situations essential to measurements
- Even if users aren't necessarily really interested in their privacy:
- real risks of damage exist, and systems that do not protect from this adequately are not useful
- does not free us from legislative constraints

Getting those preferences

- Much research in specification of privacy preferences in e.g. data collection or presence services
- Not extremely intresting, we may well assume that arbitrarily complex systems can be generated that allow any kinds of rules necessary
- The interesting problem is: how can we get the users to set these preferences so that they maximize (benefit-damage)

Getting those preferences

Leysia Palen (1999), Social, individual and technological issues for groupware calendar systems

- Well-established in HCI that users don’t change default settings
- Holds even for (at least some) private information (calendars)
- Users can find preference settings too difficult or not rewarding enough
- How to ’fix’ both? Can we? How to study this?

Getting those preferences

- Can we build a framework wherein we can reason about the power of the preference system in relation to complexity of configuration? (BRU?)
- How much effort are users willing to expend? Initially? Per recipient of presence information? Per situation? (ARU?)
- How much information do we need to guard from damage?

How much privacy do we need?

Feasibility

Following four slides based on Langheinrich (2001)

- Scott McNealy: ’You already have zero privacy anyway, get over it.’
- Can we build systems that can *enforce* privacy? (security, use restrictions, accountability)

Convenience

- Or cost vs benefit
- Free flow of information can enable us to build better personalized, proactive systems
- Protect only highly sensitive data?
- Research issues: how much is there to gain? how much is there to lose?

Communitarian

- Society as a whole can benefit from less privacy (e.g. lessen criminality)
- Can be smaller social groups (families, workplaces) as well: more honesty?
- Huge risks? Big-brother/Nazi -like societies
- Large differences in attitudes between Europe/US

Egalitarian

- No watchers and watched, you know as much about anybody else as they know about you
- New forms of social interaction based on egalitarian knowledge
- What about legitimate power structures? (e.g. families) Do such exist :-) ?
- Maybe privacy controls can be based on reciprocality (and have been based on)

User modelling

Hypothesis

- Proactive systems anticipate users’ needs
- Need personalized/learning/predictive models
- Not necessarily true?
- But assume it for now

Per-user modelling

- Maybe we can have the user store and analyze the data on a device controlled by them, so no issues
- But if the model is to be used *ubiquitously* it has to be transmitted to other devices/systems
- How sensitive is the model?
- Can the model be applied to data (so don't distribute, answer queries instead) without giving the secrets away?
- Probably not

Central modelling

- Learning from groups of people can lead to much better results
- Recommender systems good example
- Cryptographic protocols exist that allow secure multi-party computing of any reasonable functions
- Assume e.g. that 50% of users are available when analysing and that 2/3 are honest
- But only the global model is known afterwards, not individual data
- Research area: privacy-preserving data mining (BRU)

Some conclusions

From ideal/user perspective

- Build systems and data collection that the users can understand and give permission for
- Distribute data only to entities the users trust/are willing to give the information to (can be situation-specific)
- Make the setting up of trust relations easy enough for users
- Make the system compelling enough so that the users are willing to configure it
- Plenty of interesting and hard problems

Note

- The following statements are deliberately harsh
- maybe we can come up with solutions?

From regulative viewpoint

- We cannot guarantee access control or security for presence data in a contractual sense (trust is not contractual) (at least without trusted computing)
- We cannot describe the contents of user models or let users correct them (?)
- We cannot get user consent explicitly for each observation
- So we are not allowed to collect/distribute data that can be connected to individuals $\Longrightarrow$ pseudonymity or anonymity needed

From regulative viewpoint

- If anonymity or pseudonymity can be guaranteed, we are allowed to collect data
- Presence services are not possible with anonymity, pseudonymity needed: server doesn't know users' real identities, users can tell their pseudonyms to others themselves
- For ubiquitous/proactive services the user is often physically identifiable when using the system and so pseudonymity can be compromised
- Also if we collect much everyday data, the user may be identifiable from the data (e.g. locations)

Finally

Our ethics

Personally, I:

- Would probably be willing to tell quite a lot about myself to friends and family
- Don’t like the idea of trusted computing, even if it would allow us to distribute sensitive data
- Think that current legislation should in no case be relaxed
- Wouldn’t necessarily consider it harmful if people would have to be more honest about their activities

What about you?