



Institute for Software Research

University of California, Irvine

The Challenges in Preserving Privacy in Awareness Systems



Sameer Patil
Univ. of California, Irvine
patil@ics.uci.edu



Alfred Kobsa
Univ. of California, Irvine
kobsa@ics.uci.edu

April 2003

ISR Technical Report # UCI-ISR-03-3

Institute for Software Research
ICS2 210
University of California, Irvine
Irvine, CA 92697-3425
www.isr.uci.edu

www.isr.uci.edu/tech-reports.html

The Challenges in Preserving Privacy in Awareness Systems

Sameer Patil, Alfred Kobsa

Institute for Software Research,
University of California, Irvine
Irvine, CA 92697-3425, USA
{patil, kobsa}@uci.edu

ISR Technical Report # UCI-ISR-03-3

April 2003

Abstract: Awareness of the activities of one's co-workers is valuable for effective collaboration. The need for awareness is however frequently in conflict with privacy concerns of the people involved. This paper discusses various factors and principles that influence and inform a privacy-preserving design of awareness systems.

The Challenges in Preserving Privacy in Awareness Systems

Sameer Patil, Alfred Kobsa
School of Information and Computer Science
University of California, Irvine
Irvine, CA 92697-3425, USA
{patil, kobsa}@uci.edu

ISR Technical Report # UCI-ISR-03-3

April 2003

Abstract

Awareness of the activities of one's co-workers is valuable for effective collaboration. The need for awareness is however frequently in conflict with privacy concerns of the people involved. This paper discusses various factors and principles that influence and inform a privacy-preserving design of awareness systems.

Keywords: *Privacy, Awareness, Distributed software development, CSCW, Instant messaging*

1. Awareness

Awareness of the activities of collaborators helps individuals plan, orient and coordinate their own work to fit in with the larger scheme of things, with respect to the team, department or organization, thereby increasing efficiency and effectiveness of individual work as well as the work that is carried out collaboratively (Dourish and Bellotti 1992). It is no surprise then that the more tightly-coupled the collaborative activity, the higher the amount of effort and time individuals spend in seeking information about the availability and activities of others and in providing information to others of their own availability and activities (Herbsleb, Mockus et al. 2001).

Awareness information is multi-faceted. It includes information about people's presence, activities (past, present or future), schedules, routines, deadlines, availability and so on. Moreover, such information may be provided and received through

a variety of channels – from physical to social to digital. For instance, by peeking through a partially open office door one may find out whether a colleague is busy. One may also use the knowledge of a colleague's typical routine to infer her availability, or one can consult the colleague's online calendar to check for her availability.

Over the years a variety of (digital) systems have been built with the explicit goal of supporting the collection and dissemination of awareness information. Examples of such systems include Shared Media Spaces (RAVE (Bellotti and Dourish 1997), Portholes (Dourish and Bly 1992), Thunderwire (Hindus, Ackerman et al. 1996)), Shared Calendars, Mailing lists, Shared Workspaces (Polyteam (Mark, Fuchs et al. 1997), BSCW (<http://bscw.gmd.de/>), Docushare (<http://docushare.xerox.com>), CVS (<http://www.cvshome.org/>), Newsgroups), Instant Messaging (e.g. MSN Messenger (<http://messenger.msn.com>), Yahoo! Messenger (<http://messenger.yahoo.com>), ICQ (<http://www.icq.com>) and AOL Instant Messenger (<http://www.aol.com>) etc.), Sensors (Active Badges (Want, Hopper et al. 1992), Motion sensors etc.), Shared Displays (Notification Collage (Greenberg and Rounding 2001), Video monitors etc.). Even systems that are generally regarded as single-user such as email and telephone may be employed for awareness purposes. For example, caller ID may be used to screen calls; automatic email replies may be used to indicate extended unavailability and so on.

We find people typically using a combination of diverse systems and mechanisms in their efforts to

generate, disseminate and receive awareness information. The manner in which various mechanisms are combined and used depends on the people involved, the task(s), the granularity of the awareness information, the frequency of changes in awareness information, the resources, the cultural norms, the context and so on.

Awareness information assumes a much more important role in the context of the work-related activities of close collaborators – even more so if they are geographically distant (Herbsleb, Mockus et al. 2000). Since we are interested in supporting collaborative work of globally distributed teams, we will focus on studying awareness systems and mechanisms encountered in this scenario.

2. Privacy

Privacy is currently one of the most highly publicized and hotly debated topics. Yet, due to the complexities involved, there exists no commonly agreed upon, precise definition of privacy. The difficulty of precisely defining what privacy is probably stems from the fact that privacy is a highly *situated*, context-dependent concept. Not only that, but even in the *same* situation, different individuals involved may have different opinions and expectations of what privacy means. This fuzziness, context dependency and individual variability makes dealing with privacy a rather difficult task.

Bellotti (Bellotti 1996) points out that two types of privacy definitions are common, to which she refers as *normative* and *operational*. Normatively, Warren (Warren and Brandeis 1890) defines privacy as “freedom to be left alone”. Stone et. al. (Stone, Gardner et al. 1983) offer an operational definition of privacy as “ability of the individual to personally control information about oneself” whereas Samarajiva (Samarajiva 1997) extends the definition to “the control of outflow of information that may be of strategic or aesthetic value to the person and control of inflow of information including initiation of contact”.

In the physical domain, a variety of mechanisms and artifacts seem to have evolved over time to make privacy management easier. These embody certain social protocols based on some shared assumptions. For example, locking the door to prevent access to others, or knocking on a door before entering even when the door is partially open etc. However, when the shared assumptions behind the embodied social protocols are no longer applicable, for whatever reason – individual, cultural, contextual, task-specific – privacy management once again becomes problematic and privacy violations occur.

Given the inherent complexities involved in privacy management, it is possible that people always harbor some concern regarding potential violation of privacy. The consequences and risks involved may determine the amount of (explicit) effort and time devoted to managing privacy. When the consequences are potentially severe, people may devote considerable attention to preserving privacy. If, despite their efforts, a violation of privacy does occur, individuals typically *negotiate* until a commonly agreed upon state of privacy is reached for everyone involved.

3. Relationship between awareness and privacy

The above discussion regarding awareness and privacy makes the inherent interrelation between the two apparent. The general perception is that there is an inverse relationship between privacy and awareness: more awareness leads to less privacy and vice versa. Even though this may typically be the case, the reverse may also be true, i.e. providing more awareness provides more privacy. For example, maintaining a personal web page allows faculty members to limit the intrusion by requests for copies of their publications (Palen and Dourish 2003). Given the highly situated and context dependent nature of both awareness and privacy, it should be no surprise that the precise manner in which awareness and privacy are dependent on each other is also context dependent. However, regardless of the exact relationship between the two,

it is certainly true that they influence each other greatly.

The question, then, is how do people manage the relationship between awareness and privacy – both in the physical domain and in the digital domain. Answering this question involves addressing various sub-questions. Some of these include:

- What are the possible benefits to be derived from awareness?
- What are the possible benefits to be derived from privacy?
- What are people’s expectations regarding privacy?
- What mechanisms do people use to manage privacy according to these expectations?
- How do people deal with violations of privacy?
- How do people seek awareness of others?
- How do people provide awareness about themselves to others?
- How do people deal with conflicts between awareness and privacy?
- How do the various domains (physical, social, digital) differ in terms of the affordances they offer for management of awareness and privacy?

4. Privacy in current awareness systems

The current focus in awareness research lies mainly on the awareness of presence or physical activity of others (e.g. talking on the phone, reading email, etc.). In contrast to this, our primary focus is on the awareness of task-related activities, particularly in the context of distributed software development (e.g. progress on a program module, completion of documentation, reporting of a bug etc.). Nonetheless, both foci have underlying similarities that make it instructive for us to study the privacy mechanisms in current awareness systems.

Designers and builders of collaborative awareness systems frequently tend to treat privacy either as a secondary consideration or as an issue for future exploration. This may be due to the underlying assumption that individuals who collaborate with

each other have less stringent privacy expectations. The result is often systems with privacy mechanisms that are either too tight or too loose, and have minimal flexibility for modification.

Current awareness systems provide for privacy management through a combination of a large number of mechanisms. The essence of these mechanisms seems to revolve around controlling access (to oneself and one’s artifacts) through proper authorization. Different mechanisms differ in terms of who has control, who is authorized and how the process of authorization works. Some examples of privacy mechanisms include access control (e.g., password-protected login), permissions (e.g. UNIX file permissions), assignment of groups and roles, summary and distortion (e.g. abstracting a document, blurring of a video stream (Boyle, Edwards et al. 2000)). These mechanisms may be enacted and enforced in a variety of ways including provision of defaults, generation of *feedback*, enforcing of *reciprocity*, policies and procedures, social consensus and so on.

In reality, control and authorization considerations change dynamically with context. Incorporating this context dependence into the capabilities provided by present systems is problematic, to say the least. Our goal is to study the adequacy of these mechanisms for privacy management and the manners in which they are utilized in current awareness systems. If we know what works (and to what extent) and what does not work, we can look into the *why*, and then use the findings to inform the design of privacy management mechanisms in a general awareness framework.

5. Comparisons of popular Instant Messengers

One of the most popular and widespread contemporary awareness mechanisms is Instant Messaging (IM). IM allows people to indicate their presence to others who are on their “buddy lists”. At the same time, it allows checking for the presence of “buddies”. It is possible to provide finer-grained information than merely online/offline, by indicating one’s current status through various

predefined “modes”, e.g. “busy”, “on the phone”, “out to lunch”, “away from the desk”, etc. Even a simple system such as IM presents a multitude of issues regarding privacy. The importance of managing this influence on privacy, even in situations presumed to involve familiar “buddies”, is evident from the fact that all popular IM systems provide a “Privacy” menu with settings and options to allow individuals to manage their privacy.

Although IM emerged in a non-work context, it is being increasingly used and studied in the context of supporting collaborative work (Herbsleb, Atkins et al. 2002). Given the growing appeal of IM for the workplace, we started by comparing the awareness, privacy and other relevant features provided by the four most popular IM systems: AOL Instant Messenger (<http://www.aol.com>), MSN Messenger (<http://messenger.msn.com>), Yahoo! Messenger (<http://messenger.yahoo.com>) and ICQ (<http://www.icq.com>).

The following features were considered in our review:

Sound notification: This refers to the capability of associating sound alerts to various events, such as incoming messages or someone logging in.

Grouping: Grouping functionality allows various contacts to be organized into different groups, such as “Family”, “Friends”, “Coworkers”, “Project X Members” and so on.

Privacy menus: Privacy menus allow individuals to modify and customize various settings in order to manage their privacy. For instance, it may allow a person to specify whether she wishes to reveal status information on the web, or whether her phone number should be available to people on the contact list.

Blocking: Blocking a contact allows individuals to prevent their awareness information from being provided to the blocked contact. The blocked contact will always see the individual as being “offline”.

Customized Status: The ability to set customized status, such as “Working on Documentation for

Project X”, improves the limited flexibility of the pre-set status modes provided by the system.

Auto Reply: Auto reply functionality allows an individual to reply to an incoming message with an automatic reply message when she is away. The message may be chosen from ones provided by the system or may be custom defined by the individual. This functionality is analogous to an unpersonalized or personalized answering machine greeting.

Offline messaging: Offline messaging refers to the ability to receive messages from contacts even while being “offline”. Offline messaging capabilities allow the messages to be stored on the server (akin to email) and delivered upon the next login.

Popup notification: This refers to the capability of receiving small, ephemeral popups in the corner of the screen to serve as notification of events such as a contact signing in or a new message session being started.

Individual settings: This refers to the ability to specify various settings on a per-individual basis. For instance, a person may wish to be always “Available” to a certain contact, regardless of what his actual status is for other contacts. Similarly, he may not want a particular contact to have access to his cell phone number.

Group settings: Group settings refer to the ability to specify settings on a per-group basis. Changing a setting for a group affects the individual settings for all contacts in the group. Thus, if a person chooses to always appear “Away” to the “Friends” group, all contacts who are grouped under “Friends” will see her as always being “Away”.

Video connection: Video connection allows one to broadcast live video images of oneself with a computer-attached camera. The video may either be a continuous stream or a series to snapshots taken at regular (small) intervals.

Reciprocity: Reciprocity refers to whether or not the system enforces policies in a reciprocal manner. For instance, if person A blocks person B, a reciprocal policy will require that person A is also automatically blocked by person B.

Web status integration: Integrating status information with the web allows publishing of status information to a web site to allow others to view it from the web without having to log into the IM system.

Permission to add: This refers to whether or not an individual needs explicit permission from others to add them to her contact list and vice versa. Users, who wish to avoid getting multiple individual

requests for permission to add, may choose to set a global option to allow anyone to add them without explicit permission.

The following table provides a comparison of the four popular IM applications in terms of these features:

| Feature | AIM | MSN | Yahoo | ICQ |
|------------------------|-----|---------|---------|-----|
| Sound notification | Yes | Yes | Yes | Yes |
| Grouping | Yes | Yes | Yes | Yes |
| Privacy menu | Yes | Yes | Yes | Yes |
| Blocking | Yes | Yes | Yes | Yes |
| Customized Status | Yes | No | Yes | No |
| Auto Reply | Yes | No | No | Yes |
| Offline messaging | Yes | No | Yes | Yes |
| Popup notification | No | Yes | Yes | Yes |
| Individual settings | No | No | Partial | Yes |
| Group settings | No | No | No | No |
| Video connection | No | Yes/No | Yes | No |
| Reciprocity | Yes | Partial | No | No |
| Web status integration | No | No | Yes | Yes |
| Permission to add | No | Yes | Yes | Yes |

Table 1: Comparing the four major IM clients

As the table shows, even though some of the features are common to all the IM systems, there are several that are implemented in different manners in different systems. Moreover, it proves difficult and cumbersome to incorporate even the most preliminary aspects of context into the system. This is evident from the total lack of support for specifying settings based on groups of users and from the sparse support for the ability to specify settings on a per individual basis.

6. Factors and principles involved in privacy management

The above discussion suggests that the following factors and principles seem to influence privacy management:

Reciprocity: Reciprocity ensures that an individual can only request information about others that he is willing to disclose about himself and vice versa.

Feedback: Feedback involves providing individuals with information which suitably informs them of which information about them is being accessed by whom, in which form and at what time.

Context: Which information should be made available to whom and in which form is highly context dependent and keeps changing continually.

Control: Ideally, all information should be under the control of the respective individual(s) involved. This allows the individual(s) to specify how, when and to whom, information about themselves may be revealed.

Norms: Norms regarding privacy can have diverse origins. They may stem from shared cultural understandings, may evolve with growing organizational memory, may have been set through explicit policies or mandates, or may even be a combination of two or more of such influences.

Compilation: People manage privacy at the micro level, focusing on the particular task or issue at hand, and on the current context. However, several such micro-pieces of information could be collected together to form a more macro-level awareness that reveals information which people may not have wanted to divulge intentionally. For example, an individual may only wish to reveal on her shared calendar that she is in a meeting in the conference room without specifying the participants in the meeting. However, if other participants in the meeting also maintain shared calendars, then looking up the various personal calendars may reveal who the participants of the meeting are – something which some or all of the participants may have wished not to divulge.

Overhead: Managing privacy involves overhead in terms of performing tasks which are unrelated to the primary work (e.g. remembering to close and lock the door). These secondary tasks may be time consuming, tedious and/or distracting.

Incentives: Individual and/or group motivations influence how people may manage their privacy. People may be willing to sacrifice privacy if they

receive sufficient benefits. Moreover, the efforts required in privacy management need to be balanced with the direct benefits for the individual from these efforts (Grudin 1988).

Conflicts: People's desires, opinions and expectations regarding privacy may conflict with each other. Consider, for example, two colleagues who share an office, one of whom prefers the office door to be kept shut while the other prefers to leave it open.

Archiving: Archiving of information in any form – paper, digital, organization memory etc. – conserves it over time. As a result, such information may later be available out of context, in a manner different from the way in which it was originally meant to be utilized and to people other than those to whom it was addressed.

7. Privacy in digital domains

Effectively dealing with all the above factors and principles in digital domains such as awareness systems is rather challenging. Part of the reason is that privacy runs into the social-technical gap referred to by Ackerman (Ackerman 2000). Digital systems frequently embody or try to mimic artifacts and concepts from the physical and social domains. However, the underlying assumptions of privacy may be partially or totally lost in the transformation from the physical or social to the digital world. A break in expectations means either too much or too little privacy compared to what is desired and expected.

The other part has to do with specific affordances of the domain itself. In the digital domain it is much easier to mine data and compile separate pieces of information together in such a way that the compiled information is of greater value than the sum of its parts. Additionally, digital information may be easily archived extending its temporal dimension infinitely. Finally, digital information can be easily transmitted across distances making its reach global. This ease of information sharing, archiving and data mining has far-reaching

consequences for privacy management and, as a result, has heavily contributed to the wide-spread concerns regarding privacy in the digital domain.

8. Conclusion

Awareness of the presence and activities of colleagues is valuable for effective collaborative work; all the more so when team members are geographically distributed. Designing and implementing a broad framework in order to capture, maintain, provide and seek pertinent information to provide such awareness needs to deal with the thorny issue of privacy. Due to its inherently fuzzy and complex nature, privacy still remains a concept that is difficult to define. Dealing with these concerns in the digital domain is, however, essential for the design of awareness systems. We have discussed factors and principles that influence privacy and which must be taken into account when developing efficient and effective privacy management mechanisms.

References

1. Ackerman, M. S. (2000) The Intellectual Challenge of CSCW: The Gap between Social Requirements and Technical Feasibility. *Human-Computer Interaction* **15**: pp. 179-203.
2. Bellotti, V. (1996) What You Don't Know Can Hurt You: Privacy in Collaborative Computing. In *Proceedings of Human Computer Interaction Conference on People and Computers XI*, Springer.
3. Bellotti, V. and P. Dourish (1997) Rant and RAVE: Experimental and Experiential Accounts of a Media Space. In Finn, Sellen and Wilbur (eds), *Video-Mediated Communication*, LEA, New Jersey: 245-272.
4. Boyle, M., C. Edwards, et al. (2000) The Effects of Filtered Video on Awareness and Privacy. In *Proceedings of The ACM Conference on Computer Supported Cooperative Work*, Philadelphia, Pennsylvania, USA, ACM Press, New York, NY, USA.
5. Dourish, P. and V. Bellotti (1992) Awareness and Coordination in Shared Workspaces. In *Proceedings of The ACM conference on Computer-Supported Cooperative Work*, Toronto, Ontario, Canada, ACM Press, New York, NY, USA.
6. Dourish, P. and S. Bly (1992) Portholes: Supporting Awareness in a Distributed Work Group. In *Proceedings of The ACM Conference on Human Factors in Computing Systems*, Monterey, California, USA, ACM Press, New York, NY, USA.
7. Greenberg, S. and M. Rounding (2001) The Notification Collage: Posting Information to Public and Personal Displays. In *Proceedings of The SIGCHI Conference on Human Factors in Computing Systems*, Seattle, Washington, USA, ACM Press, New York, NY, USA.
8. Grudin, J. (1988) Why CSCW Applications Fail: Problems in the Design and Evaluation of Organizational Interfaces. In *Proceedings of The ACM Conference on Computer-Supported Cooperative Work*, Portland, Oregon, USA, ACM Press, New York, NY, USA.
9. Herbsleb, J. D., D. L. Atkins, et al. (2002) Introducing Instant Messaging and Chat in the Workplace. In *Proceedings of The SIGCHI Conference on Human factors in Computing Systems*, Minneapolis, Minnesota, USA, ACM Press, New York, NY, USA.
10. Herbsleb, J. D., A. Mockus, et al. (2000) Distance, Dependencies, and Delay in a Global Collaboration. In *Proceedings of The ACM Conference on Computer Supported Cooperative Work*, Philadelphia, Pennsylvania, USA, ACM Press, New York, NY, USA.
11. Herbsleb, J. D., A. Mockus, et al. (2001) An Empirical Study of Global Software Development: Distance and Speed. In *Proceedings of The 23rd International Conference on Software Engineering*, Toronto, Ontario, Canada, IEEE Computer Society, Washington, DC, USA.
12. Hindus, D., M. S. Ackerman, et al. (1996) Thunderwire: A Field Study of an Audio-only Media Space. In *Proceedings of The ACM Conference on Computer Supported Cooperative Work*, Boston, Massachusetts, USA, ACM Press, New York, NY, USA.
13. Mark, G., L. Fuchs, et al. (1997) Supporting Groupware Conventions through Contextual Awareness. In *Proceedings of The Fifth European Conference on Computer Supported Cooperative Work*, Lancaster, England, Kluwer Academic Publishers, Dordrecht.
14. Palen, L. and P. Dourish (2003) Unpacking "Privacy" for a Networked World. In *Proceedings of The SIGCHI Conference on Human Factors in Computing Systems*, Fort Lauderdale, Florida, USA.
15. Samarajiva, R. (1997). Interactivity as Though Privacy Matters. In *Technology and Privacy: The New Landscape*, P. Agre and M. Rotenberg (eds.), MIT Press, Cambridge, MA.
16. Stone, E., D. Gardner, et al. (1983) A Field Experiment Comparing Information-Privacy Value, Beliefs and Attitudes Across Several Types of Organizations. *Journal of Applied Psychology* **68**(3): 459-468.
17. Want, R., A. Hopper, et al. (1992) The Active Badge Location System. *ACM Transactions on Information Systems (TOIS)* **10**(1): pp. 91-102.
18. Warren and Brandeis (1890) "The Right to Privacy." *Harvard Law Review* **4**(5).