Proceedings of the Third European Conference on Computer-Supported Cooperative Work
13-17 September, 1993, Milan, Italy
G. De Michelis, C. Simone and K Schmidt (Editors)

Design for Privacy in Ubiquitous Computing Environments

Victoria Bellotti* and Abigail Sellen*†
* Rank Xerox EuroPARC, Cambridge, UK
bellotti@europarc.xerox.com; sellen@europarc.xerox.com

†MRC Applied Psychology Unit, Cambridge, UK

Abstract: Current developments in information technology are leading to increasing capture and storage of information about people and their activities. This raises serious issues about the preservation of privacy. In this paper we examine why these issues are particularly important in the introduction of ubiquitous computing technology into the working environment. Certain problems with privacy are closely related to the ways in which the technology attenuates natural mechanisms of feedback and control over information released. We describe a framework for design for privacy in ubiquitous computing environments and conclude with an example of its application.

Introduction

Information technology can store, transmit and manipulate vast quantities and varieties of information. Whilst this is critical to government, public services, business and many individuals, it may also facilitate unobtrusive access, manipulation and presentation of personal data (Parker et al., 1990; Dunlop & Kling, 1991).

The term "Big Brother" in the context of computing technology, seems to imply two classes of problem. The first is due to the fact that computer technology may be put to insidious or unethical uses (e.g., Clarke, 1988). All information systems, and particularly distributed systems, are potentially vulnerable to covert subversion (Lampson et al., 1981) and, although it can be made extremely difficult to tamper with data in computing systems, protection mechanisms "are often only secure *in principle*. They are seldom secure *in practice*." (Mullender, 1989).

Deliberate or poorly considered design resulting in invasive applications and sophisticated subversion of supposedly secure systems are discouraged by cultural censure and law (although these forces trail behind the advances in sophistication of the technology). However, software must still be secure in order to reduce the risks of covert abuse of personal data and this is an important area of research. There are already a number of useful software protection models and standards which are designed to reduce the risks (see e.g., Lampson et al., 1981; and Mullender, 1989).

The second class of problem is related to very different concerns about a fast growing, less well understood set of issues. These arise from user-interface design features which interfere with social behaviour. These features may foster unethical use of the technology but, more significantly, they are also much more conducive to inadvertent intrusions on privacy (Heath & Luff, 1991).

Mediated interactions between people via technology are prone to breakdowns due to inadequate feedback about what information one is broadcasting and an inability to control one's accessibility to others. This disrupts the social norms and practices governing communication and acceptable behaviour. Our concern in this paper is tackle the latter kind of problem in the context of systems design.

In attempting to design systems which reduce perceived invasions of privacy, it would be useful to have a practical working definition of the concept. Unfortunately, although privacy is widely considered to be an important right, it is difficult to define this notion in more than an intuitive fashion (Anderson, 1991). Attitudes to what is and what is not private data vary between people in different contexts and roles (Harper et al., 1992). Codes of practice and the law offer inadequate guidance on what actually counts as violation of privacy in technologically sophisticated environments (Clarke, 1988) and it may take lengthy court proceedings to determine what the case may be (Privacy Protection Study Commission, 1991).

Any realistic definition of privacy cannot be static. With the introduction of new technology, patterns of use and social norms develop around it and what is deemed "acceptable" behaviour is subject to change. Naturally evolving social practices may interact with organisational policies for correct usage (Harper et al., 1992). In addition, people are more prepared to accept potentially invasive technology if they consider that its benefits outweigh potential risks (e.g., Ladd, 1991; Richards, 1991). In recognition of these facts we take privacy to be a personal notion shaped by culturally determined expectations and perceptions about one's environment.

The social practices and policies that determine any *rights* an individual has to privacy interact with the technical and interface design aspects of the technology they use. Technology is not neutral when it comes to privacy. It can increase or reduce the extent to which people have control over personal data. Our concern is to ensure that privacy should be a central design issue in its own right.

We present a framework for addressing the design of *control* and *feedback* of information captured by multimedia, ubiquitous computing environments. These two issues are fundamental to successful communication and collaboration amongst users as well as to maintaining privacy. We ground our examples largely in the domain of networked audio-video environments and in particular in experiences

with one such environment. However, our framework may also be related to the design of CSCW systems and distributed computing environments in general.

In the following sections we first introduce the context and nature of the technology which is the focus of our interest, we then go on to outline our design framework and provide a brief example of its application.

Maintaining privacy in a media space

The need to understand and protect personal privacy in sophisticated information systems is becoming critical as computing power moves out of the box-on-the-desk into the world at large. We are entering the age of *ubiquitous computing* (e.g., Weiser, 1991; Lamming & Newman, 1991; Hindus & Schmandt, 1992) in which our environment comes to contain computing technology in a variety of forms.

Increasingly, we are seeing such systems incorporate sensors such as microphones, cameras and signal receivers for wireless communication. These sensors have the potential to transmit information such as speech, video images, or signals from portable computing devices, active badges (Want et al., 1992), electronic whiteboards (Pederson et al., 1993), and so on. These devices can be networked so that multimedia information can be stored, accessed, processed and distributed in a variety of ways. Services include audio-video (AV) interconnections, information retrieval, diary systems, document tracking and so on (e.g., Lamming & Newman, 1991; Gaver, 1992; Eldridge et al., 1992).

Ubiquitous computing usually implies embedding the technology unobtrusively within all manner of everyday objects which can potentially transmit and receive information from any other object. The aims are not only to reduce its visibility, but also to empower its users with more flexible and portable applications to support the capture, communication, recall, organisation and reuse of diverse information. The irony is that its unobtrusiveness both belies and contributes to its potential for supporting potentially invasive applications.

In light of these developments, it is dangerously complacent to assume that social and organisational controls over accessibility of personal information are sufficient, or that intrusions into privacy will ultimately become acceptable when traded against potential benefits. Such a position could leave individual users with a heavy burden of responsibility to ensure that they do not, even inadvertently, intrude on others. It also leaves them with limited control over their own privacy.

"Media spaces" (Stults, 1988) are a recent development in ubiquitous computing technology, involving audio, video and computer networking. They are the focus of an increasing amount of research and industrial interest into support for distributed collaborative work (e.g., Root, 1988; Mantei et al., 1991; Gaver et al., 1992; Fish et al., 1992). EuroPARC's RAVE environment is just one of several media spaces which have been set up in various research laboratories around the world.

In RAVE, cameras, monitors, microphones and speakers are placed in every office, to provide everyone with their own personal RAVE *node*. This allows one to communicate and work with others and to be aware of what is going on in the build-

ing without leaving one's office. Various kinds of flexible video-only and AV connections between nodes are set up and broken by central switching devices which are controlled from individual workstations.

Whilst media space technology improves the accessibility of people to one another, some may feel that their privacy is compromised. The very ubiquity of such systems means that many of the concerns with existing workstation-based information systems are aggravated. A much wider variety of information can now be captured. People are much less likely to be "off-line" (inaccessible) at any given moment. Further, the design of many of these systems is such that it may not be clear when one is off- or on-line and open to scrutiny (Mantei et al., 1991; Gaver, 1992). People also express concern about their own intrusiveness to others when they try to make contact without being able to determine others' availability (Cool et al., 1992). Concerns about such problems have strongly influenced the installation and ongoing design of RAVE, as well as the way in which people use it.

Feedback and Control in RAVE

At EuroPARC people generally do not worry much about privacy. They feel that the benefits of RAVE outweigh their concerns. This is because the design has evolved together with a culture of trust and acceptable practices relating to its use. Individual freedom was fostered to use, customise, or ignore the technology. Design was informed by studies of how collaborative work is socially organised and how such technology impacts it (e.g. Heath & Luff, 1991; 1992). Users' views and reactions were obtained via questionnaires and interviews. The varied individual feelings about privacy were accommodated by ensuring that users could decide how accessible they were to others via the media space (Dourish, 1991; Gaver et al., 1992; Dourish, 1993).

In designing for privacy in RAVE, two important principles have emerged (Gaver et al, 1992). These are *control* and *feedback*, which we define as follows:

Control: Empowering people to stipulate what information they project and who can get hold of it.

Feedback: Informing people when and what information about them is being captured and to whom the information is being made available.

RAVE users can control who may connect to them and what kind of connections each person is allowed make. If they omit to do so, automatic defaults are set to reject connections. User control via the workstation is supported by "Godard", the software infrastructure which provides the primary interface to the complex AV signal-switching and feedback mechanisms (Dourish,1991). These mechanisms comprise the kinds of connections which can be made between people, to different public areas, and to media services (e.g., video-players).

Feedback depends on the type of RAVE connection being made. Three kinds of interpersonal connection are "glance", "v-phone call" and "office-share". Glance connections are one-way, video-only connections of a few seconds' duration. V-phone and office-share connections are longer two-way AV connections. For

glances, audio feedback (Gaver, 1991) alerts users to onset and termination of a connection and can even announce who is making it. For the two-way office connections, reciprocity acts as a form of feedback about the connection (if I see you, you see me) and, in the case of an attempted v-phone connection, an audio "ringing" signal is given and the caller's name is displayed on the workstation, whereupon the recipient can decide whether to accept or reject the connection. Office-shares, being very long term, do not require such a protocol.

Public areas have cameras which can be accessed by a glance or a "background" connection which is indefinite, one-way and video-only. We provide feedback about the presence of a camera in a public place in the form of a video monitor beside the camera which displays its view.

Control and feedback also figure strongly in the design of RAVE's architecture. Connection capability lists and an access control model define who can connect to whom and provide long term, static control over accessibility. Providing distinct connection types can also allow users to exercise discriminating dynamic control over their accessibility as in the v-phone call (for a fuller description of these features see Dourish, 1993). Our concern, however, is with the moment-to-moment continuous control that people exercise over how they present themselves in public as respectable, social beings (Goffman, 1963). In the next section we indicate why people, especially newcomers and visitors to places with media spaces and other kinds of ubiquitous computing technology, can feel uneasy about their ability to monitor and control their self presentation and consequently their privacy.

Disembodiment and dissociation

A number of problems with RAVE relate to the extended duration of v-phone calls and office-shares. Users tend to forget about their existence and associated implications. Even seasoned users can get confused about the nature of their connections. For example, if a colleague with whom you have an office-share switches off their camera or moves out of shot, it is easy to forget that they can still see you.

Problems in public areas include the fact that monitors next to cameras only suggest (and then only to those familiar with a media space) that a video image may be being broadcast to many people, via background connections. They cannot inform people when or to whom the image is being sent. For most EuroPARC "regulars" this is not a major concern, but for newcomers to the building, it may be.

The underlying causes of such problems lie in the fact that the technology results in *disembodiment* from the context into and from which one projects information (Heath & Luff, 1991) and *dissociation* from one's actions. These phenomena interfere with conveying information about oneself or gaining information about others.

Conveying information: In the presence of others you convey information in many ways. These include position, posture, facial expression, speech, voice level and intonation, and direction of gaze. Such cues influence the behaviour of others. For example, they can determine whether or not others will try to initiate communication with you.

In CSCW and ubiquitous computing environments disembodiment means that these resources may be attenuated. So you may not be able to present yourself as effectively to others as you can in a face-to-face setting. For example, in an AV connection, you may only be a face on a monitor (Gaver, 1992) with your voice coming out of a loudspeaker, the volume of which you may not be aware or able to control. You may only appear as a name (McCarthy et al., 1991) or a name associated with a room displayed on a screen (e.g., Harper et al., 1992). At worst (e.g., in a RAVE background connection) you may have no perceptible presence at all. On the other hand, disembodiment also means that you may be unaware of when you are convey information to others because of a lack of feedback from the technology.

Dissociation occurs in CSCW applications when only the results of actions are shared, but the actions themselves are invisible. In other words when you cannot easily determine who is doing, or did, what (e.g., ShrEdit, a shared editor with no telepointers; McGuffin & Olson, 1992).

Gaining information: In face-to-face situations, cues given by others influence your judgements about whether to attempt conversation, what to say and how to act.

In media spaces, there is usually no way to gauge how available someone else is before connecting to them (Louie et al., in press). Once connected, awareness of the person at the other end of the link or their actions is likely to be limited to the fixed and narrow field of view of a camera, and whatever a microphone picks up (Gaver, 1992). That person also has a reduced, disembodied presence. In turn, you are likely to receive fewer cues when someone is observing you, or your work.

Breakdown of Social and Behavioural Norms and Practices

The effects of disembodiment and dissociation manifest themselves in a variety of breakdowns in behavioural and social norms and practices. For example, breakdowns associated with disembodiment include a tendency for users to engage in unintentional, prolonged observation of others over AV links (Heath & Luff, 1991). Users may intrude when they make AV connections, because they cannot discern how available others are (Louie et al., in press). Furthermore, the intuitive principle that if I can't see you then you can't see me, does not necessarily apply to computer mediated situations where one person may be able to observe others' activities without themselves being observed.

A major problem related to dissociation is one's inability to respond effectively to a perceived action because one does not know who is responsible for it. A familiar example of this problem exists with telephones where it is impossible to identify nuisance callers before picking up the receiver.

Problems of disembodiment and dissociation receive far less attention in the literature than insidious exploitation of technology. This is unfortunate as they are also problems for social interaction and communication mediated by technology and likely to be much more pervasive, particularly because they often relate to purely unintentional invasions of privacy. Furthermore, by addressing these problems through careful design, we may reduce the potential impact of system abuse.

It must be pointed out that the technology itself is not inherently problematic. Resources used in face-to-face situations can be exploited, simulated, or substituted through design. For example media space systems can embody the principle "If I see your face, you see mine," which is natural in face-to-face situations (e.g., VideoTunnels; Buxton & Moran, 1989) or they can supply means to convey availability (e.g., Louie et al., 1992). Dissociation problems in CSCW systems have been reduced by means of conveying gestures, or even body posture (e.g. Minneman & Bly, 1991; Tang & Minneman, 1991).

Our ongoing research assumes that problems of interaction, communication and privacy in ubiquitous computing systems, can be reduced through technological design refinements and innovations. Disembodiment and dissociation may be reduced through the provision of enriched feedback about the state of the technology and information being projected about users. Users must also have practical mechanisms of control over that personal information. We now present a framework for systematically addressing these issues. Although it may have general use for designing CSCW technology to support social interaction and communication, we focus in particular on how it helps us confront privacy as a central design concern.

A design framework

Based on our experience with privacy issues in RAVE and other similar systems, we have developed a simple design framework aimed at counteracting the kinds of problems we have outlined.

Addressing the Problems

Much of the mutual awareness, which we normally take for granted may be reduced or lost in mediated interpersonal interactions. We may no longer know what information we are conveying, what it looks like and how permanent it is, who it is conveyed to, or what the intentions of those using that information might be. In order to counteract problems associated with this loss, our framework proposes that systems must be explicitly designed to provide feedback and control for at least the following potential user and system behaviours:

Capture: What kind of information is being picked up? Candidates include voices, actual speech, moving video or framegrabbed images (close up or not), personal identity, work activity and its products such as keypresses, applications used, files accessed, messages and documents.

Construction: What happens to information? Is it encrypted or processed at some point or combined with other information and, if so, how? Is it stored? In what form? Privacy concerns in ubiquitous computing environments are exacerbated by the fact that potential records of our activity may be kept and possibly manipulated, and used at a later date and out of their original context. This leads to numerous potential ethical problems (Mackay, 1991).

	Feedback About	Control Over
Capture	When and what information about me gets into the system.	When and when not to give out what information. I can enforce my own preferences for system behaviours with respect to each type of information I convey.
Construction	What happens to information about me once it gets inside the system.	What happens to information about me. I can set automatic default behaviours and permissions.
Accessibility	Which people and what software (e.g., daemons or servers) have access to information about me and what information they see or use.	Who and what has access to what information about me. I can set automatic default behaviours and permissions.
Purposes	What people want information about me for. Since this is outside of the system, it may only be possible to infer purpose from construction and access behaviours.	It is infeasible for me to have technical control over purposes. With appropriate feedback, however, I can exercise social control to restrict intrusion, unethical, and illegal usage.

Figure 1. A framework for designing for feedback and control in ubiquitous computing environments: Each cell contains a description of the ideal state of affairs with respect to feedback or control of each of four types of behaviour.

Accessibility: Is information public, available to particular groups, certain persons only, or just to oneself? What applications, processes, and so on utilise personal data.

Purpose: To what uses is information put? How might it be used in the future? The intentions of those who wish to use data may not be made explicit. It may only be possible to infer what these are from knowledge of the person, the context, patterns of access and construction.

We now consider each of these four classes of concerns in relation to the following two questions:

What is the appropriate feedback?

What is the appropriate control?

We thus have eight design questions which form the basis for a design framework (Figure 1) with which we can analyse existing designs and explore new ones, with respect to a range of privacy issues. This framework is a domain specific example of the QOC approach to design rationale in which design issues, couched as

questions, are explicitly represented together with proposed solutions and their assessments (for more details see MacLean et al., 1991; Bellotti, 1993).

The issues in the cells within the framework are not necessarily independent of one another. For instance, in order to be fully informed about the purpose of information usage, one must know something about each of the other behaviours. Likewise, in order to appreciate access, one must know about capture and construction. Understanding construction requires knowing something about capture. Hence there is a dependency relationship for design of feedback between these behaviours. Control over each of them may, however, be relatively independently designed.

For those concerned about privacy, and the potential for subversion in particular, control over, and thus feedback about, capture is clearly the most important. Given appropriate feedback about what is being captured, users can orient themselves appropriately to the technology for collaboration or communication purposes and exercise appropriate control over their behaviour or what is captured in the knowledge of possible construction, access and purposes of information use.

Evaluating Solutions

Our framework emphasises design to a set of criteria, which may be extended through experience and evaluation. Whilst questions about what feedback and control to provide set the design agenda, criteria represent additional and sometimes competing concerns which help us to assess and distinguish potential design solutions. The set of criteria acts as a checklist helping to encourage systematic evaluation of solutions. They have been identified from our experiences with the design and use of a range of ubiquitous computing services. Particularly important in our current set are the following eleven criteria.

Trustworthiness: Systems must be technically reliable and instill confidence in users. In order to satisfy this criterion, they must be understandable by their users. The consequences of actions must be confined to situations which can be apprehended in the context in which they take place and thus appropriately controlled.

Appropriate timing: Feedback should be provided at a time when control is most likely to be required and effective.

Perceptibility: Feedback should be noticeable.

Unobtrusiveness: Feedback should not distract or annoy. It should also be selective and relevant and should not overload the recipient with information.

Minimal intrusiveness: Feedback should not involve information which compromises the privacy of others.

Fail-safety: In cases where users omit to take explicit action to protect their privacy, the system should minimise information capture, construction and access.

Flexibility: What counts as private varies according to context and interpersonal relationships. Thus mechanisms of control over user and system behaviours may need to be tailorable to some extent by the individuals concerned.

Low effort: Design solutions must be lightweight to use, requiring as few actions and as little effort on the part of the user as possible.

Meaningfulness: Feedback and control must incorporate meaningful representations of information captured and meaningful actions to control it, not just raw data and unfamiliar actions. They should be sensitive to the context of data capture and also to the contexts in which information is presented and control exercised.

Learnability: Proposed designs should not require a complex model of how the system works. They should exploit or be sensitive to natural, existing psychological and social mechanisms that allow people to perceive and control how they present themselves and their availability for potential interactions.

Low cost: Naturally, we wish to keep costs of design solutions down.

The first seven criteria are especially relevant to protection of privacy. The final four are more general design concerns. Some of these criteria have to be traded off against one another in the search for design solutions.

Applying The Framework: Feedback and Control for Video Data from a Public Area

We have begun to apply our framework to RAVE to reveal aspects of its design which can be refined. For the sake of brevity, we focus on just one aspect of this media space which involves a video connection from the Commons, a public reading and meeting area at EuroPARC.

In RAVE, people can access the camera in the Commons either in the form of short glance or indefinite, background, video-only connections. The video data can sent to devices such as a framegrabber which takes digitised video snaps. These can be used by various services such as Vepys, a video diary which takes still images of people via media space cameras, every so often, wherever they are, and stores the series of images as a browsable record of their day-to-day activity (Eldridge, 1992).

Providing feedback and control mechanisms over video data taken from this public area is a challenging problem but it is an important one, since the Commons is an area where many visitors to EuroPARC spend their time.

Our framework prompts us to ask the following questions for which we describe existing or potential design solutions (relevant criteria appear in italics):

Q: What feedback is there about when and what information about me gets into the system?

Existing Solutions:

Confidence monitor: A monitor is positioned next to the camera to inform passers by when they are within range, and what they look like. This solution fulfils the design criteria of being *trustworthy, meaningful* and *appropriately timed.*

Mannequin (the Toby): In order to alert people to the presence of the camera, a mannequin affectionately named Toby is positioned holding the camera. Toby draws people's attention because he looks like another person in the room. Originally the camera was in his head, but this concealed it and some visitors thought it was deliberately being hidden. Now Toby holds the camera on his shoulder. This feedback is *less obtrusive* than the confidence monitor (which can be distracting),