Proceedings of the Second European Conference on Computer-Supported Cooperative Work
Bannon, L., Robinson, M. & Schmidt, K. (Editors)
September 25-27, 1991, Amsterdam, The Netherlands

PEPYS: Generating Autobiographies by Automatic Tracking

William M. Newman, Margery A. Eldridge and Michael G. Lamming

Rank Xerox EuroPARC, England

This paper presents one part of a broad research project entitled 'Activity-Based Information Retrieval' (AIR) which is being carried out at EuroPARC. The basic hypothesis of this project is that if contextual data about human activities can be automatically captured and later presented as recognisable descriptions of past episodes, then human memory of those past episodes can be improved. This paper describes an application called Pepys, designed to yield descriptions of episodes based on automatically collected location data. The program pays particular attention to meetings and other episodes involving two or more people. The episodes are presented to the user as a diary generated at the end of each day and distributed by electronic mail. The paper also discusses the methods used to assess the accuracy of the descriptions generated by the recogniser.

Introduction

Human memory is far from perfect. Most people can recount numerous occasions when someone has had to remind them of some of the circumstances of a long-forgotten event. Consider the following scenario:

Person A:	"Do you remember what I said about changing our paper?"
Person B:	"No."
Person A:	"Remember we were working on section two on your workstation?"
Person B:	"No—was John in the office?"
Person A:	"Yes—and you had just had a phone call from the conference organiser."
Person B:	"Oh yes, he said it was too long, and you said....."

and so begins the recall of the main part of the conversation.

All human activities take place in some context: the discussion to which A is referring here took place in a particular room (B's office), in the presence of a third person (John), and after a telephone conversation (from the organiser). Reconstruction of such a context can assist recall of episodes that took place in the context (Smith, Glenberg & Bjork, 1978). The basic hypothesis of the research described in this paper is that if contextual data about human activities can be

automatically captured and later presented as recognisable descriptions of past episodes, then human memory of those past episodes can be improved. The collection and analysis of data related to the context of various human activities is a part of a broad research programme into Activity-Based Information Retrieval, or AIR (Lamming & Newman, 1991). The present paper describes one part of the AIR programme which deals with information about the locations of groups and individuals. It describes a program, Pepys, that has been designed to make inferences about collaborative work based on location data. Thus information about meetings and other group activities is combined with information about individual activities, and descriptions of these activities are presented in the form of a personal diary outlining the individual user's day.

Although Pepys presents information to people in the form of a personal diary, it is fundamentally different from other diary management and scheduling systems that have been reported in the CSCW literature. The diary presented by Pepys is *retrospective*, and is not directly applicable to diary management or scheduling meetings (for example, the Visual Scheduler reported in Beard et al., 1990). The system for analysing workstation activity described by Thimbleby, Anderson and Witten (1990) resembles Pepys more closely: one of the aims of their system is to help people recall activities they had previously carried out on their workstations. It does this overnight by performing numerous checks on files that have been created or modified during the day, producing a comprehensive retrospective summary of the day's workstation activity.

A particular emphasis of Pepys is on reconstructing collaborative episodes and including them in the diary. Meetings, hallway discussions, encounters around the photocopier, etc., appear to be particularly valuable as contextual cues for recall. Monitoring workstation activity, as proposed by Thimbleby et al., is also being pursued as part of the AIR project (see Lamming & Newman, 1991), and the aim is to incorporate this information into the diaries generated by Pepys. A wide variety of episodes, many of them collaborative, contribute to the working day and should ultimately find their place in the worker's daily record.

EuroPARC: The Research Site

The AIR project is based at Rank Xerox EuroPARC in Cambridge, England. The EuroPARC building is a multi-media environment, where each office and meeting room contains an audio/video node including a camera, monitor, microphone and speaker (Buxton & Moran, 1990). This system is used to support various multi-media applications, such as video conferencing, video-phones and non-speech audio announcements. The laboratory infrastructure also includes a tracking system that allows people to record their movements automatically. This tracking system, developed at Olivetti Research Labs (Want, 1990), provides location data that serve as a database for the project reported here.

The Active Badge System

EuroPARC shared an interest with Olivetti Research Labs in investigating potential applications of tracking technology, and therefore in 1989 installed the badge system shown in Figure 1. Badges measure about 5 cm square and are worn like normal security tags. Each badge has a different identity code, which it emits approximately every 20 seconds using infra-red signalling similar to the method employed in TV remote controls. This code is picked up by receivers installed in the rooms, hallways and stairwells of the building. Roughly once per second, a polling computer interrogates each receiver for recently detected badge codes. When a badge is detected at a new location, this event is stored in a record that includes the date and time, the badge ID code and the old and new locations. A location server makes these records available to applications that rely on current badge data, including the data logging program that builds log files for Pepys. The data in log files are encoded and are kept secure from unauthorised applications. When decoded, the contents of the files are as shown in Figure 2. Several different types of record may appear in the files: besides the basic records of movement, there may be lost-badge records, where a person has not been detected by the system for several minutes, and attention records, where the person has pressed a button on the

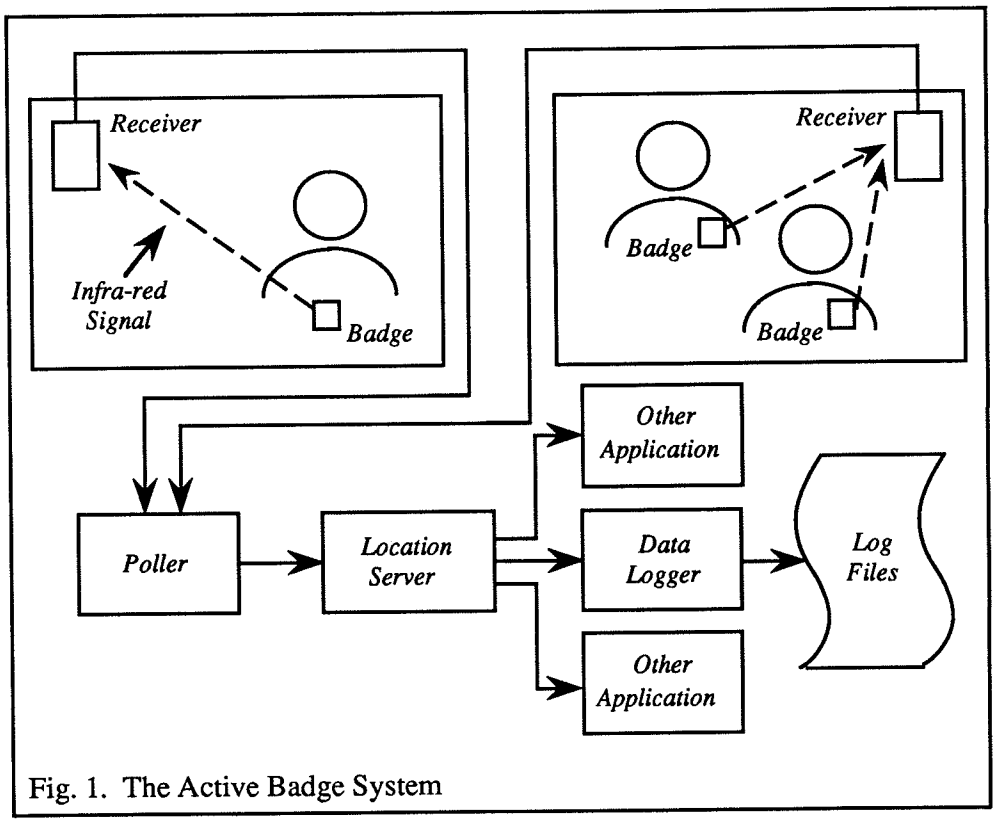

Fig. 1. The Active Badge System

8:26:40 am	30-Oct-90	Morton	move	Commons > Kitchen
8:28:02 am	30-Oct-90	Morton	move	Kitchen > 3rd floor corridor
8:29:05 am	30-Oct-90	Morton	move	3rd floor corridor > Kapp's area
8:29:55 am	30-Oct-90	Morton	move	Kapp's area > Fax/Copier room
8:25:46 am	30-Oct-90	Price	lost	Wilson's office
8:37:30 am	30-Oct-90	Little	attention	Little's office
8:30:18 am	30-Oct-90	Morton	lost	Fax/Copier room
8:41:04 am	30-Oct-90	Little	attention	Little's office
8:42:27 am	30-Oct-90	Morton	move	> Stairwell
8:42:48 am	30-Oct-90	Morton	move	Stairwell > Morton's office
8:45:39 am	30-Oct-90	Morton	move	Morton's office > Reception
8:48:36 am	30-Oct-90	Morton	move	Reception > Morton's office
8:52:09 am	30-Oct-90	Little	move	Little's office > Commons
8:53:27 am	30-Oct-90	Andrews	move	> Rear porch
8:53:27 am	30-Oct-90	Andrews	attention	Rear porch
8:53:38 am	30-Oct-90	Morton	move	Morton's office > Clark's office
8:53:46 am	30-Oct-90	Andrews	move	Rear porch > Stairwell
8:54:24 am	30-Oct-90	Little	move	Commons > Kitchen
8:54:28 am	30-Oct-90	Morton	move	Clark's office > Morton's office
8:54:30 am	30-Oct-90	Morton	move	Morton's office > Kitchen
8:54:35 am	30-Oct-90	Little	move	Kitchen > Stairwell
8:55:02 am	30-Oct-90	Little	move	Stairwell > Little's office
8:55:10 am	30-Oct-90	Andrews	move	Stairwell > Reception

Fig. 2. An example of a decoded portion of a Log File.

badge in a particular location. Each of these records appears in the sample of Figure 2. Pepys takes account of movement and lost-badge records, but ignores all other types of event.

Several other applications have been implemented using the Active Badge system. One system allows the administrative support staff to forward phone calls to EuroPARC personnel by reading their location off a screen (Harper, Lamming & Newman, 1991). Another system allows badge-wearers to gain access to the building by pressing the button on the badge and thus automatically unlocking the main entrance door. These applications have contributed to the Pepys project by encouraging staff to wear badges.

Social Implications of Active Badges

Tracking technologies tend to raise alarms about invasion of privacy. There was considerable concern over the introduction of active badges at EuroPARC, and one of the aims of ongoing research has been to understand these concerns better and to address them. The ultimate goal is to learn how to design and build systems incorporating novel technologies with due attention to their social implications. Many of the technology-based artefacts which support tracking and are now widespread in society, such as credit cards and cellular telephones, have presented similar

problems when first introduced. The same problems are now being raised by collaborative technologies used in CSCW, such as video and shared workstation environments. An informed approach to design can avoid the two undesirable extremes of forcing such technologies on an unwilling public or abandoning them altogether in order to avoid a negative backlash.

Preliminary studies of the social impact of badges have helped identify some of the issues at stake. In a study reported by Harper, Lamming and Newman (1991), an application called the Locator was developed at two different sites to enable people to be located more easily, and the outcomes at the two sites were compared. While some of the findings are specific to the Locator application, others can be generalised to apply to other CSCW technologies and applications such as Pepys.

The Locator project raised the issue of what kind of information within an organisation is 'private' and what is 'public'. The balance between these two categories of information is a delicate matter, normally kept at a tacit level. At one of the two sites, system support people felt that their location was public information, since their social role within the organisation depended upon other people being able to find them when needed. In contrast, researchers at this same site, having the freedom to organise their time as they wished, felt that information about their location was not public—their social roles within the organisation did not include being found by others. Introduction of badges appeared to affect the individual's control over the boundary between public and private.

The Locator also showed very clearly the relationship between privacy concerns and utility of technology. Some users found the Locator extremely useful: it meant that they could receive phone calls wherever they were, and that they spent less time answering other people's calls and going in search of them. Administrative staff valued the Locator because it enabled them to do their job better and to respect the individual wishes of staff for privacy or availability. These people seemed relatively unconcerned about the badges' invasive characteristics. Researchers gained little benefit from the Locator, because they worked primarily alone, received few phone calls, and disliked being disturbed. These people tended to view the badges as socially undesirable. There was a strong negative correlation, therefore, between the perceived usefulness of the Locator and concern over the badges' invasiveness: utility appeared to have the effect of attenuating concern. This phenomenon was seen in its most extreme form among researchers who had no access to the Locator, and therefore had no prospect of benefiting from it

A similar comparative study of Pepys at two sites is presently under way. The conclusion of this paper presents some interim observations about possible extrapolations of the Locator results to the Pepys project.

The Design of Pepys

Background

Badges offer a number of possible ways of supporting information retrieval; several of these were investigated before the system described here was built. For example, a simple prototype system was built to demonstrate the feasibility of indexing into recorded audio or video using a database of location data gathered from badges. This was a useful exercise in many ways, helping to identify technically difficult areas and to show where activity-based information retrieval might be most useful. Initially it had seemed that an AIR application, in order to be useful, would need to include at least an interactive user interface and a means of indexing into existing forms of data such as video or electronic documents. The Pepys project showed that neither of these was strictly necessary: that a useful application could consist merely of diaries generated automatically from badge data and issued each day via electronic mail. Indeed, Pepys may be considered an example of an interactive system with an extremely 'low-intensity' user interface (Newman, 1990) requiring almost no effort on the part of the user.

The first stage of the project was to understand how to process badge data into a recognisable reconstruction of the day's events. Part of this stage involved understanding how to organise the data. At first, data were collected in separate files for each user, on the assumption that a fairly simple analysis of each file would yield the events of the user's day. However, the analysis turned out to be far from simple, requiring inspection of all of the users' data together. A different approach was required to data logging, creating a single file for each day's traffic. Although this decision simplified the design, it had some unexpected negative effects on users' attitudes to Pepys, and these are discussed in the Conclusion.

At one stage in this early research, an entire day's badge data were collected and analysed by hand in complete detail. Diagrams were drawn to represent gatherings of several people at one location; from these it was possible to identify the formation and dispersal of meetings at various locations in the building throughout the day. It was possible from the data to recognise journeys made by individuals between events: trips to the coffee pot, tours of the building looking for people, guided tours by visitors. Also recognisable were the periods that users spent alone in their offices. From this analysis it became clear that the program would need to look for three basic types of episode: gatherings of two or more people, travel between locations, and periods spent alone. The analysis also helped in the design of algorithms for automatic recognition of these episodes, which were built into the two principal software modules of Pepys—the so-called Quorum Spotter and Travel Agent.

Quorum Spotter

The Quorum Spotter generates a log of the day's gatherings from the location data log gathered from badges. The term 'gathering' is used in preference to 'meeting' which is inappropriate for casual events. Each gathering is defined in terms of its location, overall start and finish times, and the attendances by the individuals taking part, that is, the periods of time they spent at the meeting. A gathering is normally considered to start when a 'quorum' of two people has formed, and to finish when such a minimal quorum no longer exists. It is necessary, however, for at least one member of the quorum to be a non-resident of the room in which the gathering is taking place; otherwise people sharing offices would appear to be in continuous meetings! Attendances are recognised as periods spent at locations where gatherings exist. It is possible for an individual to make several attendances at the same gathering, and the program identifies these instances so that they can be spelled out in the diary: each attendance record names the gathering involved, by location and start time.

The Quorum Spotter attempts to distinguish between the different types of gathering that may form, and to attach type descriptions appropriate for inclusion in diaries. In particular, it looks for episodes that appear to justify the description 'meeting.' The criteria on which it decides include the length of the episode, the number of people present, speed with which the quorum formed and then dispersed, and the location. Thus it would not apply the description 'meeting' to a gathering in which attendees continually arrived and departed throughout, or to a gathering in a corridor.

Travel Agent

The task of the Travel Agent is to process the data on changes of location and to distinguish between important and unimportant changes. An active user may clock up several hundred such changes during a single day, including many that result in spending only a few seconds at each location. The Travel Agent must discard such transient records, and concentrate on recognising 'stopovers'—significant periods at a single location. Some stopovers are easy to recognise, e.g., lengthy periods in a gathering or alone in the user's office. Others are more subtle, e.g., stopping for three minutes by a photocopier, or spending five minutes in a meeting that had been in progress for two hours.

The three main criteria on which the Travel Agent selects stopovers are: (1) the length of the stopover itself, (2) the nature of the stopover event, and (3) the distance travelled. If the user spends more than a certain number of minutes at a location, this is considered significant on its own. If a meeting was taking place there, this is considered memorable even if the user joined the meeting for only a short time. If neither of these was the case, but the user travelled a significant distance to be at a particular location, that on its own justifies inclusion of the episode. The

Travel Agent therefore requires two additional types of information not present in the badge data files: meetings attended by the user, which it obtains from the Quorum Spotter, and a floor-plan of the building giving distances between pairs of locations. These enable it to build a list of significant stopovers, each specifying a start and stop time, a location, and a pointer to an attendance record if appropriate.

Diary Episode Builder

A further stage of processing is carried out to transform the Travel Agent's stopover list into a list of episodes suitable for inclusion in the diary. As will be seen, Pepys diaries provide a continuous sequence of episodes, from which the user can see what Pepys thinks she or he was doing at any time between the first and last sighting of the day. The stopover list, on the other hand, has gaps in it caused by travel between stopovers or by other activity excluded because it failed to meet the stopover criteria. These gaps can easily be accounted for, but the resulting list is very detailed, often including insignificant entries such as 40 seconds spent going from one location to another.

The main roles of the Diary Episode Builder are to fill in the gaps in the stopover list and then to reduce its level of detail by scanning through the list repeatedly, looking for particular patterns. Gap-filling is done by inserting a period alone, if the stopovers before and after were at the same location, or a period of travel if they were not. Detail-reduction involves replacing sequences of two or more elements of the list by a single, new element. The sequence is retained as a sub-list of the new element so that the details can later be included in the diary. Figure 3 shows an example of pattern-matching by the Diary Event Builder.

alone in 2nd floor pod from 13:04 to 13:08	13:04 In 2nd floor pod [4 mins]
attending event in 2nd floor pod from 13:09 to 13:34	13:08 Gone from 2nd floor pod
alone in 2nd floor pod from 13:34 to 13:35	[1 min]
alone in Smith's office from 13:35 to 13:40	13:09 Attended discussion in 2nd floor
attending event in Smith's office from 13:43 to 13:53	pod; with Tim [25 mins]
alone in Smith's office from 13:53 to 13:54	13:35 Mostly in meetings in office
attending event in Smith's office from 13:54 to 14:12	[37 mins]
alone in Smith's office from 14:12 to 14:13	13:35 In office [4 mins]
attending event in 2nd floor pod from 14:13 to 14:18	13:40 Gone from office [3 mins]
alone in 2nd floor pod from 14:18 to 14:20	13:43 Meeting in office; with Alice
end of travel at Watson's office from 14:20 to 14:21	[29 mins]
attending event in Kitchen from 14:21 to 14:22	14:13 Attended event in 2nd floor pod;
attending event in Kitchen from 14:22 to 14:23	with Tim [11 mins]
alone in Commons from 14:23 to 14:24	14:24 Attended event in Commons;
attending event in Commons from 14:24 to 14:26	with Tim, Fred [4 mins]
alone in commons from 14:26 to 14:27	14:28 In and out of event in Watson's
attending event in Watson's office from 14:28 to 15:39	office; with Fred,Tim [1hr 10m]

Fig. 3. Pattern matching by the Diary Episode Builder. The left column shows stopovers, the right column shows the final output after pattern matching.

Diary Composer

The final stage of Pepys is the composing and mailing of diaries. Several alternative approaches were considered, including graphical presentation techniques laying out episodes along a time line. There was no single internal format, however, that could have been received by all badge wearers at their workstations, which included Suns running Xerox Lisp and Unix, and Xerox 6085 workstations running Xerox Lisp, Cedar and ViewPoint. Hardcopy was considered as a distribution medium, but this would have required daily manual intervention.

Pepys therefore distributes diaries as plain text files via electronic mail. The program runs every morning at 4:00 a.m., reading the entire log of EuroPARC users' data for the previous 24 hours. Diaries are also distributed to badge-wearers at PARC, where logs are also maintained; the program prepares these diaries at 2:00 a.m. GMT to minimise the network load caused by accessing the logs from EuroPARC. The diary consists of the hierarchical list of episodes generated by the Diary Event Builder, pruned to a level of detail specified by the user as one of his or her preferences. Other user-settable options include formatting of time and names, inclusion of summaries, and inclusion of comment fields into which the user can type a description of each event. An example of a Pepys diary, showing comment fields partially filled in, is shown in Figure 4.

Accuracy of Episode Descriptions

The ultimate goal of the Pepys project, and the AIR project in general, is to provide users with an interactive tool which can be used for activity-based information retrieval. In order for Pepys to be used as an interactive retrieval tool, it is necessary that the episode descriptions used by Pepys match those in the users' retrieval requests. One of the first steps required, therefore, is to evaluate the episode descriptions generated by Pepys, and to rectify any errors found in these descriptions. Although more research is required to provide a better understanding of how people describe episodes in their working lives, the project reported here represents the first steps in designing such an interactive retrieval tool.

Interviewing Pepys Users

Pepys has been run on data collected at both Xerox PARC in California and at Rank Xerox EuroPARC. Shortly after the distribution of diaries had begun at PARC, several Pepys users were interviewed, and these interviews were videotaped. Users were first asked to recollect the events of the previous day, and then they were shown their Pepys diaries. After reading their diaries, several users remembered events that they had previously forgotten. However, it was clear that there were some inaccuracies in the episode descriptions in the diaries. For example,

31-Oct-90 Pepys:EuroPARC:RX Diary for Tuesday, October 30, 1990
Date: October 31, 1990 9:11:20 am GMT
From: Pepys:EuroPARC:RX
Subject: Diary for Tuesday, October 30, 1990
To: Little:EuroPARC:RX
Reply-To: Pepys:EuroPARC:RX

Summaries for Tuesday, October 30, 1990

Met in Nathan's office at 15:08 with R. Hatton, W. Nathan [41 mins]
Met in Commons at 16:06 with B. Andrews, M. Morton, R. Hatton [6 mins]
Met in Wright's office at 16:57 with P. Wright [3 mins]
Met in Little's office at 17:12 with I. David, P. Wright [45 mins]
Met in Little's office at 17:58 with P. Wright [15 mins]
Met in Little's office at 18:50 with W. Nathan [6 mins]

2h 10m with others = 43 percent
2h 38m alone = 53 percent
8 mins travelling = 3 percent

4h 57m total

Diary for Tuesday, October 30, 1990

14:14		In office [50 mins] **Writing paper**
15:04		In and out of event in Nathan's office; with W. Nathan, R. Hatton [45 mins] **discussing paper**
15:50		In office [10 mins] **Reading E-Mail**
16:00		In Conference room [4 mins] **checking video set-up**
16:05		Attended part of event in Commons; with B. Andrews, M. Morton, R. Hatton [7 mins] **Comment**
16:13		Mostly in office [44 mins] **Comment**
16:57		Attended event in Wright's office; with P. Wright [7 mins] **Comment**
17:04		Looked in on event in Morton's office; with I. David, M. Morton [1 min]
17:05		Mostly in office [2 hr 3m] **Comment**
	17:05	In office [5 mins]
	17:11	In event in office; with P. Wright, I. David [1h 2mins]
	18:13	In office [36 mins]
	18:50	Meeting in office; with W. Nathan [13 mins]
	19:03	In office [5 mins]
19:09		In 2nd floor rear area [2 mins] **Comment**
19:11		Last seen

v1.13

Fig. 4. An example of a diary generated by Pepys. The first six lines show the standard header added by the electronic mail system; following these are a summary of meetings and a breakdown of time spent; the final section is the diary itself, with comments (in bold) being added by the user on receipt of the diary.example, several users commented that short events were not necessarily unimportant and should not be ignored. Some users were also unsure about what the different episode descriptions meant (e.g., 'event' versus 'meeting', etc.).

several users commented that short events were not necessarily unimportant and should not be ignored. Some users were also unsure about what the different episode descriptions meant (e.g., 'event' versus 'meeting', etc.).

Monitoring Feedback from Pepys Users

Additional detail on the accuracy of episode descriptions was obtained by monitoring feedback from Pepys users at EuroPARC. Users were encouraged to comment on the information contained in their diaries. Of the 26 badge-wearers at EuroPARC, 17 receive Pepys diaries and responses were obtained from just over half of these users. Several users complained about inaccuracies, mainly in the records of their attendances at meetings. Some users also commented on being included as attendees of meetings when they had only passed through the area where a meeting was being held. After analysis of the comments and complaints received from the Pepys users, it was discovered that some of the inaccuracies were caused by the incorrect placement of receivers in some of the rooms in the building, and some other inaccuracies occurred within Pepys itself. Errors in the placements of receivers were corrected, and various changes to the Pepys algorithms were made as a result of the feedback from Pepys users.

Estimating the Accuracy of Episode Descriptions

To get a quantitative estimate of the overall accuracy of the episode descriptions in the Pepys diaries, a sample of seven of one of the authors' (MAE's) diaries was evaluated. Because the definition of an episode varies depending on the logged events which make up the episode, this estimation of accuracy is quite rough. For the purposes of this estimation, an episode is defined as one time-stamped major entry in a diary (sub-levels under the major entry are not included). To be classified as 'accurate', the following must be true: (1) the starting time of the episode must be correct (within 5 minutes); (2) the duration of the episode must be correct (because of the summation over events that occurs in the Pepys software, the duration of the episode was deemed correct if it was within 5 minutes of the actual duration); (3) the location of the episode must be correct; (4) the description of the episode (i.e., meeting, discussion, gone from, in office, etc.) must be accepted by the author as a reasonable description of that episode; and (5) for events involving other people, all badge-wearers present at the event must be included.

A total of 95 episode descriptions occurred in the sample of seven diaries. Of these, 79 were accurate descriptions using the criteria listed above. This gives an overall accuracy of the Pepys diaries of 85%. Over half of the inaccurate descriptions (9 of 16) were due to not all badge-wearers being listed as present at the event. This inaccuracy is caused by some people at EuroPARC not consistently wearing their badges. Six of the remaining inaccurate descriptions were episodes of the type 'gone from'. This episode description is sometimes quite misleading,

particularly when an exit and subsequent re-entry into the building are obscured by the description. One inaccuracy was due to the length of a partial attendance at a meeting being recorded as being for 10 minutes, when in fact it lasted less than one minute.

Overall, then, the level of accuracy in the Pepys diaries is very high. In fact, if the cases where not all badge-wearers were listed as attending an event are eliminated, the overall accuracy goes up to 90%. This is a reasonable adjusted estimate because nearly all of these inaccuracies are due to individuals failing to wear their badges 100% of time; they are thus not caused by errors in the Pepys algorithms themselves.

Summary of Problems

The inaccuracies in the episode descriptions are the results of the following problems:

(1) Although nearly all personnel at EuroPARC have been issued badges, not all people wear their badges consistently. Some of the social effects discussed earlier in this paper have contributed to some people not wearing badges. Descriptions of episodes involving personnel not wearing badges will necessarily be inaccurate (e.g., 'time spent alone' could be time spent with a non-badge-wearer; meeting rosters may be incomplete, etc.).

(2) Some people wear their badges in non-optimal positions (on hip pockets), where the emissions from the badges can be obscured by clothing or pieces of furniture.

(3) The system can 'lose' people for a variety of reasons in addition to those mentioned in (1) and (2), for example, when someone leaves the building for a brief period. Different reasons for these lost-badge events are very hard for the Pepys algorithms to recognise.

(4) Accurate episode recognition requires an understanding of how people describe episodes in their working lives. Little work has been done on this in the past, and a more thorough understanding is required.

(5) Some of the Pepys algorithms are incorrect, partly due to the problems mentioned above, and partly because the Pepys algorithms were initially designed for the sparse badge system installed at PARC and do not take full account of the complete coverage of the badge system at EuroPARC.

These problems are currently being addressed by ongoing work on the AIR project. The badges themselves are being redesigned, and this should improve the reliability of signal detection. Some of the Pepys algorithms are also being changed to deal with 'lost' badges. More research is being undertaken to understand how people describe episodes in their working lives, and this research should suggest changes to the Pepys algorithms which will make the generated episode descriptions more accurate, and thus more meaningful.

Research is also under way to conduct a comparison of the use of Pepys at several sites, and to see whether this bears out earlier experience with the Locator. Already some similarities have been observed, such as the concern over public versus private information, and the tradeoff between utility and these privacy concerns. The decision to collect badge data into a single log, and to issue diaries from a central service, appears to have led to a view of Pepys as a 'public' system, and of the log files as 'public' files. Designing Pepys as a personal tool, drawing on a private log file, might have produced a different user response. Users appeared very uncertain, at the outset, about the possible benefits of receiving diaries; this may explain why approximately one third of EuroPARC's badge-wearers did not ask to receive them. Lack of perceived utility may have heightened concerns about public access to location data. More applications based on the Active Badge System are being pursued, and as these applications are introduced, there should be an increase in the benefits associated with wearing badges.

Conclusion

The project reported here was exploratory, and in great part served as a learning exercise. When the project was started, it was not known that the recognition of episodes from location data would be such a major issue. Not only were there several false starts in achieving this episode recognition, but those episodes that were eventually generated were not accurate in all respects. One focus of our future research will, therefore, centre on evaluating and refining the accuracy of the episode descriptions generated by Pepys.

Even though there is still a great deal to be learned about how episodes are described, a number of important things have been learned. Although the diaries present only a fairly coarse-grained description of events, several novel uses of the diaries have been observed. For example, several visitors, who have received diaries for the time of their stay at EuroPARC, have used the information in the diaries to write up their trip reports. Other people have used the information to help in preparing overtime forms or in writing their own personal diaries. Still other people have commented that the diaries were useful in reminding them of meetings and also of possible commitments made during those meetings.

Many people have also suggested enhancements which should serve to make the diaries more useful. For example, several people have suggested adding known calendar events (e.g., tea, seminars, etc.) to the diaries; adding more finely-grained information to the diaries has also been suggested. For example, including details of telephone calls and of informal conversations have both been suggested, and these and other monitoring techniques are currently being pursued (Lamming & Newman, 1991).

Within the AIR project as a whole, there are two major areas of work which are currently being pursued. The first concerns investigations which will help us to

understand the social effects of the technology, and thus to increase the acceptance and the perceived benefits of the technology. The second area of work is concerned with understanding more about how people's working lives are structured, and what sorts of activities people can and cannot remember about their working days. This information will then help to develop and enhance applications like Pepys to make them more useful in retrieving information about work-related activities.

Acknowledgement

We wish to thank Pierre Wellner, Lawrence Butcher, Dan Swinehart, Bart Kinne, Mike Molloy, Ian Daniel and Roy Want for their contributions to designing and building the badge system on which Pepys depends. We are grateful to Mark Weiser for his support of the initial Pepys software development at PARC, and to Bob Anderson for supporting the subsequent research at EuroPARC. Many Pepys users at PARC and EuroPARC provided valuable comments and interviews. During the preparation of this paper, Richard Harper provided valuable comments on analysis of social impact. We are also grateful for advice received from Allan MacLean, Sara Bly, Phil Barnard, Debra Bekerian and Jerry Elkind.

References

Beard, D., Palaniappan, M., Humm, A., Banks, D., Nair, A., & Shan, Y-P. (1990): "A visual calendar for scheduling group meetings", *CSCW '90 Proceedings*, October 1990, pp. 279-290.

Buxton, W., & Moran, T. (1990): "EuroPARC's integrated interactive intermedia facility (IIIF): Early experiences", *Proceedings of IFIP WG8.4 Conference on Multi-User Interfaces and Applications*, Heraklion, Crete, September 1990.

Harper, R., Lamming, M.G., & Newman, W.M. (1991): "The locator: an experiment with active badges", *EuroPARC Technical Report*, in preparation.

Lamming, M.G. (1991): "Using automatically generated descriptions of human activity to index multi-media data", *IEE Colloquium on Multi-Media Communications and Applications*, February 1991, pp. 5/1-5/3.

Lamming, M.G., & Newman, W.M. (1991): "Activity-based information retrieval: Technology in support of human memory", paper submitted for publication.

Newman, W.M. (1990): "Interacting with Ubiquitous Systems." Invited paper, Eurographics '90 Conference, Montreux, 1990.

Smith, S., Glenberg, A.M., & Bjork, R.A. (1978): "The influence of environmental context on recall and recognition. *Memory & Cognition*, vol. 6, no. 4, 1978, pp. 342-353.

Thimbleby, H., Anderson, S., & Witten, I.H. (1990): "Reflexive CSCW: supporting long-term personal work", *Interacting with Computers*, vol. 2, no. 3, 1990, pp. 330-336.

Want, R. (1990): "The active badge location system", *Olivetti Research Laboratories Report*, The Old Addenbrookes Site, 24a Trumpington St., Cambridge, England CB2 1QN, 1990.