Sensor Redundancy and certain Privacy Concerns

Position paper for the UbiComp 2003 workshop on Privacy as Boundary Negotiation

James Fogarty
Human Computer Interaction Institute
Carnegie Mellon University
jfogarty@cs.cmu.edu

In our work on situationally appropriate computing, we have been examining how systems can observe people in the real world, model their behaviors, and interact with them in ways that are socially appropriate, or simply polite. Specifically, we have been focusing on using sensors to automatically estimate the interruptibility of office workers, so that systems might use interruptibility estimates to determine whether or not it is currently appropriate to initiate an interaction [5, 7]. This position paper provides a very short review of relevant portions of this work, discusses a notion of sensor redundancy that we have developed to address a particular sort of privacy concern raised by this work, and introduces some thoughts on how this notion of sensor redundancy might relate to the boundary negotiation interpretation of privacy presented by Palen and Dourish [11].

One of the major results in this line of work has been to show the predictive power of using a single microphone in an office to determine whether or not anybody in that office is talking. The population we studied was most non-interruptible when socially engaged, such as when on the telephone or having a meeting in the office. Because talking is highly correlated with these activities, a single microphone is actually a very good sensor for estimating interruptibility. Our work has found that a sensor to determine whether anybody is talking in an office can, by itself, provide estimates of interruptibility that are as accurate as those made by people viewing audio and video recordings of the office working environment.

The predictive power of a single microphone in an office environment makes it very reasonable to think about deploying systems that automatically estimate human interruptibility. We have been using simple $10 microphones plugged into the existing audio ports of commodity desktop computers and have been experimenting with using the microphones that are already a part of most notebook computers. As such, there is little or no cost associated with sensing infrastructure. This is in sharp contrast to approaches to sensing that require the installation of cameras or other sensors. While any individual camera may be inexpensive, it can quickly get very expensive to run network cables for cameras throughout a building (or even just power cables if a wireless network is used). Video processing is also computationally expensive enough that dedicated computers will probably be required, while our audio processing is done on a person's existing computer.

While the microphone-based approach has many advantages in terms of the cost and practicality of deployment, it has the same problems that have previously been reported when deploying cameras or microphones in office environments [1, 2, 4, 6, 9, 14]. Many people object to such deployments because they feel they may be monitored without their knowledge or consent. In the framework presented by Palen and Dourish, this rejection of cameras and microphones can be at least partially explained by their negative impact on the reflexive interpretability of action – people feel they cannot anticipate how the information collected by the cameras and microphones might be used.

Work on media spaces has explored a variety of technical approaches for mitigating some of these concerns. These include enforcing the reciprocity normally found in shared spaces [4], as well as processing audio [12, 13] or video [3, 8, 15] signals to reduce their information content in ways designed to address certain privacy concerns. The privacy community has also explored methods for limiting the flow of information that is collected [10], providing certain technical (or social and legal) guarantees about how information will be used once it is collected.

We believe that certain privacy concerns related to cameras and microphones cannot be sufficiently addressed by guarantees about how information will be used or shared once it is collected. Some people have strong emotional reactions to cameras or microphones, and they may either not understand or not trust whatever additional technical mechanisms are in place to protect their privacy. We do not mean to say that such mechanisms are not important, only that some people will see a camera and believe that somebody can use the camera to watch them.

This position paper suggests that, instead of attempting to mitigate these emotional reactions to cameras or microphones, designers of ubiquitous computing systems can instead take advantage of redundancies in human behavior to entirely remove the need for cameras and microphones. Our microphone sensor is effective because talking is highly correlated with many other activities we might want to sense. While investigating the possibility of predicting human interruptibility without using cameras or microphones, we found that other sensors are highly correlated with social engagement. For example, our data shows that visitors to an office often sat down, rather than standing the entire time they were present. Using conductive foam in chairs to determine when people are sitting in them seems like it might draw less of a negative emotional reaction than a camera or microphone. Because it is clear that the only thing that can be sensed by this foam is whether or not something heavy is currently on the chair, it seems that such a sensor is more conducive to supporting reflexive interpretability of action.

Table 1 presents information on some activities that could be sensed to determine whether or not an office worker is currently in their office. These values are based on manually simulated sensors that are discussed in [5, 7]. Active badges can provide this information with a very high degree of reliability. Location tracking systems can also be based on the wireless access point being used by a person's computer. But both of these systems could potentially expose additional information, such as where the person is when they are not in their office. As Palen and Dourish point out, some office workers

	Keyboard	Mouse	Keyboard or Mouse	Talk	Sit	Desk
Immediate	22.3%	21.9%	40.6%	28.9%	87.2%	67.3%
30 Second	34.4%	35.7%	52.5%	37.2%	92.9%	78.3%
1 Minute	38.5%	40.2%	60.0%	39.7%	93.9%	80.8%
2 Minutes	43.8%	48.2%	61.6%	45.8%	94.9%	84.1%
5 Minutes	55.7%	60.4%	70.8%	57.7%	96.3%	88.1%

Table 1 – Sensors for Detecting Occupant Presence

	Talk	Stand	Sit
Immediate	81.8%	55.3%	39.0%
30 Second	88.7%	65.4%	40.9%
1 Minute	90.6%	69.8%	41.5%
2 Minutes	95.0%	74.2%	42.8%
5 Minutes	97.5%	86.2%	45.9%

Table 2 – Sensors for Detecting Guest Presence

feel that systems that allow a person's location to be determined can sometimes be perceived as a violation of that person's definition of self. Note that one commonly used approximation of presence, whether or not a person has moved the mouse or typed on the keyboard recently, can be used with a 5 minute window to detect 70.8% of occasions when a person is present. A sensor that detects whether somebody is sitting in the office detects 96.3% of occasions when the person is present when used with a 5 minute window. This additional level of accuracy would require the instrumentation of places where a person is likely to sit, and it would be up to the designer of an individual ubiquitous computing system to decide whether the improved accuracy warranted the additional sensing requirements. However, the collection of this sort of information would allow the designer to make an informed decision, rather than just speculating about what sensors might be appropriate.

Table 2 presents similar information on some activities that could be sensed to determine if a guest is present. Note that if guests were present in an office, they had almost always talked during some point in the previous 5 minutes. This is why the microphone emerges as such a good sensor for detecting social engagement. This information also shows that, if guests were present in an office, they had sat down at some point in the previous 5 minutes for 45.9% of our samples. It was this comparison that led us to investigate whether our models of human interruptibility could be created without using a camera or a microphone.

Figure 1 shows confusion matrices for one result of investigating this possibility in our domain. Two models using sensors that could be created without using a camera or microphone both achieve an accuracy of greater than 78%. Neither of these models performs significantly different from the 76.9% accuracy of people shown audio and video recordings of the office working environments of the people whose interruptibility

1	Telephone (Imm)
2	Guest Sit (Imm)
3	Desk (Change-120)
4	Telephone (All-30)
5	Mouse (All-120)

1	Telephone (Count-30)
2	Guest Sit (Imm)
3	Desk (Change-120)
4	Guest Sit (Any-300)

		Naïve Bayes	
		All Other Values	Highly Non
Video Subject	All Other Values	433 64.4%	24 3.6%
	Highly Non	120 17.9%	95 14.1%
		Accuracy: 78.6%	

		Decision Tree	
		All Other Values	Highly Non
Video Subject	All Other Values	432 64.3%	25 3.7%
	Highly Non	122 18.2%	93 13.8%
		Accuracy: 78.1%	

Figure 1 – Naïve Bayesian and decision tree models of human interruptibility constructed without a camera or microphone.

we are modeling. We will not discuss these models in any depth in this position paper, but these models seem to work because they detect social engagement without using a microphone for sensing talking. Specifically, they use a switch that detects whether the phone is physically off its hook and chair sensors to determine if another person is sitting in the office. These sensors seem to detect social engagement reliably enough that models of human interruptibility can be created without a camera or microphone. While these results indicate the possibility of creating such models without using cameras or microphones, it is significantly more difficult to deploy sit sensors than it is to deploy a microphone. In considering sensor redundancy as an approach, it will be necessary to balance privacy concerns with cost and practicality concerns.

While participating in this workshop, I hope to convey the opinion that the interpretation of privacy presented by Palen and Dourish should be applied not only to how information is used and shared in ubiquitous computing systems, but also to decisions about the design of the sensors used in systems. We believe that technical guarantees about how information will be shared and used are not sufficient for deploying ubiquitous computing systems into some environments. Some people either will not understand or will not trust these technical guarantees, and they will have negative reactions to the presence of a camera or microphone in spite of these guarantees. In our domain, we have shown that redundancies in human behavior (such as the fact that guests would often sit when they were in an office for a substantial period of time) mean that we can sense social engagement without the use of a camera or microphone. We believe that such sensors might be better received because their limitations are clear – they better support reflexive interpretability of action. We also believe that the decision to use such sensors should be based on studies of the environments where systems will be used. These studies can provide the designer of a ubiquitous computing system with quantitative information on the impact on reliability that might be expected. This can then be weighed against cost and practicality concerns and the possibility that people will be more receptive to sensors that allow greater interpretability of action.

References

1. Adler, A. and Henderson, A. A room of our own: experiences from a direct office share. In *Conference companion on Human factors in computing systems*, ACM Press, 1994, 209.
2. Bly, S.A., Harrison, S.R. and Irwin, S. (1993) Media spaces: bringing people together in a video, audio, and computing environment. *Communications of the ACM*, *36* (1). 28--46.
3. Boyle, M., Edwards, C. and Greenberg, S. The effects of filtered video on awareness and privacy. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, ACM Press, 2000, 1--10.
4. Fish, R.S., Kraut, R.E., Root, R.W. and Rice, R.E. (1993) Video as a technology for informal communication. *Communications of the ACM*, *36* (1). 48--61.
5. Fogarty, J., Hudson, S., Atkeson, C., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J. and Yang, J. (2003) Predicting Human Interruptibility with Sensors. *In Review, ACM Transactions on Computer-Human Interaction (TOCHI)*.
6. Gaver, W., Moran, T., MacLean, A., Lövstrand, L., Dourish, P., Carter, K. and Buxton, W. Realizing a video environment: EuroPARC's RAVE system. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, 1992, 27--35.
7. Hudson, S., Fogarty, J., Atkeson, C., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J. and Yang, J. Predicting human interruptibility with sensors: a Wizard of Oz feasibility study. In *Proceedings of the conference on Human factors in computing systems*, ACM Press, 2003, 257--264.
8. Hudson, S.E. and Smith, I. Techniques for addressing fundamental privacy and disruption tradeoffs in awareness support systems. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work*, ACM Press, 1996, 248--257.
9. Jancke, G., Venolia, G.D., Grudin, J., Cadiz, J.J. and Gupta, A. Linking public spaces: technical and social issues. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, 2001, 530--537.
10. Jiang, X., Hong, J.I. and Landay, J.A. Approximate information flows: Socially-based modeling of privacy in ubiquitous computing. In Borriello, L.E. and Holmquist, E. eds. *"Ubicomp 2002: Ubiquitous Computing," ACM UBICOMP Conference Proceedings*, Springer-Verlag, 2002, 176-193.
11. Palen, L. and Dourish, P. Unpacking "privacy" for a networked world. In *Proceedings of the conference on Human factors in computing systems*, ACM Press, 2003, 129--136.
12. Schmandt, C., Kim, J., Lee, K., Vallejo, G. and Ackerman, M. Mediated voice communication via mobile IP. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*, ACM Press, 2002, 141--150.
13. Smith, I. and Hudson, S.E. Low disturbance audio for awareness and privacy in media space applications. In *Proceedings of the third ACM international conference on Multimedia*, ACM Press, 1995, 91--97.
14. Tang, J.C. and Rua, M. Montage: providing teleproximity for distributed groups. In *Conference companion on Human factors in computing systems*, ACM Press, 1994, 201.
15. Zhao, Q.A. and Stasko, J.T. Evaluating image filtering based techniques in media space applications. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work*, ACM Press, 1998, 11--18.

[5] is in review and not publicly available, but I will be happy to share it with anybody related to this workshop. Please contact me at jfogarty@cs.cmu.edu for access to it.