Giving the Caller the Finger: Collaborative Responsibility for Cellphone Interruptions

Stefan Marti
MIT Media Lab, E15-384C
20 Ames Street, Cambridge MA, 02139
stefanm@media.mit.edu
+1 (617) 253-8026

Chris Schmandt
MIT Media Lab, E15-368
20 Ames Street, Cambridge MA, 02139
geek@media.mit.edu
+1 (617) 253-5156

Abstract
We present a system in which a cell phone decides whether to ring by accepting votes from the others in a conversation with the called party. When a call comes in, the phone first determines who is in the conversation by using a decentralized network of autonomous body-worn sensor nodes. It then vibrates all participants' wireless finger rings. Although the alerted people do not know if it is their own cellphones that are about to interrupt, each of them has the possibility to veto the call anonymously by touching his/her finger ring. If no one vetoes, the phone rings. A user study showed significantly more vetoes during a collaborative group-focused setting than during a less group oriented setting. Our system is a component of a larger research project in context-aware computer-mediated call control.

Author Keywords
Mobile communication, cellphone interruption, social polling, collaborative responsibility, social intelligence.

ACM Classification Keywords
H5.2. User Interfaces: Interaction Styles; Haptic I/O.

INTRODUCTION
In the midst of a quiet romantic dinner with your spouse, your mobile phone rings. Perhaps you would like to destroy it? But the next day a long-winded colleague corners you at the water cooler; a call would be a relief and an excuse to terminate the conversation.

The mobile phone is a device we love to hate; yet we cannot live without it. Its interruption is important to our productivity at work and social and familial availability, yet we detest a distracting call breaking up an important face-to-face conversation. Those around us, co-located conversation partners or strangers who happen to share the physical space, are also impacted. These third party eavesdroppers can become uncomfortable and annoyed by the interrupting call that has nothing to do with their ongoing activity, and may behave so as to assert their own physical presence [5].

CHI 2005, April 2–7, 2005, Portland, Oregon, USA.
ACM 1-59593-002-7/05/0004.

In this paper, we describe a system that allows a phone that acts as an intermediary between caller and callee to anonymously and subtly poll the participants of a conversation (including the owner of the cellphone) in order to assess the appropriateness of a cellphone interruption. This is a component of a larger research project in context-aware computer-mediated call control.

DESIGN RATIONALE AND SYSTEM
Mobile phones are interrupting to others around us as well as to ourselves. Recent research proposes detection of interruptibility [4] and the possibility of adding the recipient's context to the phone agent's decision matrix [7]. The recipient's social context should be incorporated into the decision as well, but we doubt that sensor-based approaches will ever be able to understand the tone or importance of a conversation. But local others are part of the user's social setting and generally have accurate insight about the user's interruptibility. Our system exploits people's expertise in social intelligence. Most humans know well what is socially appropriate in a given situation, especially how to interrupt a conversation when something important comes up, and not to interrupt when it is not important enough. Furthermore, humans know exactly what kind of social situation they are in and if it is appropriate to take phone calls.

Therefore, we suggest that whether a phone call should interrupt a group setting should not get decided by the user only, but also by co-located people. In order for a phone to "harvest" local other's input, it needs to solve the following problems:

1. Determine who is part of the user's conversation
2. Notify all involved in the user's conversation with subtle "pre-alerts." (*Alert* is the term used in telephony for the ring, hence the "pre-alert")
3. Get input from all in the conversation via "vetoing"

If there is no veto after 10 seconds, the cellphone will alert.

(1) is implemented in a system called Conversation Finder nodes, (2) and (3) are part of our Finger Ring system. We will describe these pieces in the following sections.

Conversation Finder nodes

In order to determine who is part of the user's conversation, we have developed short-range radio sensor nodes (38x33x15mm) with dual microprocessors and microphone (Figure 1), worn close to the neck. They observe the wearers' conversational turn taking. Basu [1] has showed that alignment of speech is a reliable indicator for conversational groupings. Our approach is novel, however, because the groupings are detected in real-time, and each sensor node is autonomous and comes to a conclusion independently.

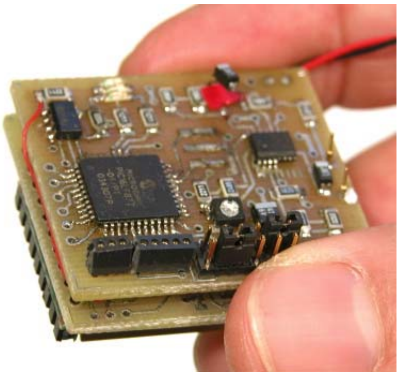

Figure 1: Conversation Finder node prototype

A Conversation Finder sensor node can transmit two types of messages: a periodic 'heartbeat' message to indicate mere presence, and a 'talk' message indicating that its user is currently speaking.

Simultaneously, each node listens for incoming radio messages from nearby nodes. Upon receiving a 'heartbeat,' the other node is classified as *Listener*. Detecting a 'talk' message will upgrade its status to a *Talker*. Each node continuously determines if the detected nodes might be part of its owner's conversation or not. If the node's microphone determines that its user is talking, and simultaneously receives 'talk' messages from another node for more than a three-second window, it excludes the other node for a 30-second period by tagging it as *Excluded*. If a node classified as a *Talker* stops sending 'talk' messages, it will get re-classified to a *Listener* after a period of time. Similarly, if a node fails to send out 'heartbeat' messages, it will get tagged as *Absent* by the other nodes. This continuous process of classifying all other nodes is done in each sensor node independently, and our prototype nodes (on breadboards) have proven to be a reliable and fault tolerant source of conversational status information.

Our prototype phone has a compatible radio transceiver and uses the Conversation Finder's state to determine how to handle a call. We anticipate that commercial deployment could be done through implementing this protocol on Bluetooth headsets. Each Conversation Finder sensor node will provide three pieces of information: (1) how many people are in the user's conversation; (2) for each participant, what is the ratio between listening and talking; and (3) is the user talking right now.

Finger ring

The Conversation Finder nodes' information about conversational groupings is used to create a socially intelligent mobile communication device. The people who are involved in a face-to-face conversation with the user are 'polled' in a subtle way, using a wirelessly actuated finger ring that can vibrate, indicating as to whether an interruption from a cellphone would be appropriate. Upon this pre-alert, all involved are given the possibility of anonymously vetoing the call by simply touching their respective rings. This shifts the burden of deciding whether to interrupt away from the phone and towards the humans who are actually involved in a conversation.

Since no one knows which mobile communication device is about to interrupt, this system of 'social polling' fosters collective responsibility for controlling interruption by communication devices.

Protocol

In our wireless system, the network protocol hides the identity of the vetoing person from the phone that queries all participants for their input (Figure 2).

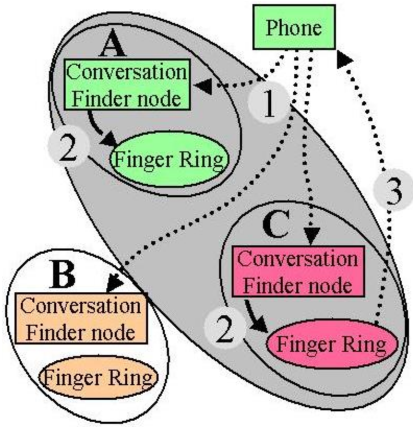

Figure 2: Communication protocol

The querying phone of user A broadcasts a message to all Conversation Finder nodes in range (1). If they 'think' they are in a conversation with user A (in our example: user A and C, but not user B), they are asked to send a directed message to their respective finger rings, which will cause these rings to vibrate as a pre-alert (2). At that point, all participants who received pre-alerts (A and C) can veto the upcoming interruption by pressing the finger ring's micro switch. If a user presses the switch (user C), the ring will broadcast an anonymous veto message that will be picked up by the querying phone of user A (3). Note that A's phone never knows who else is in a conversation, much less who vetoed (but can still count the number of vetoes received).

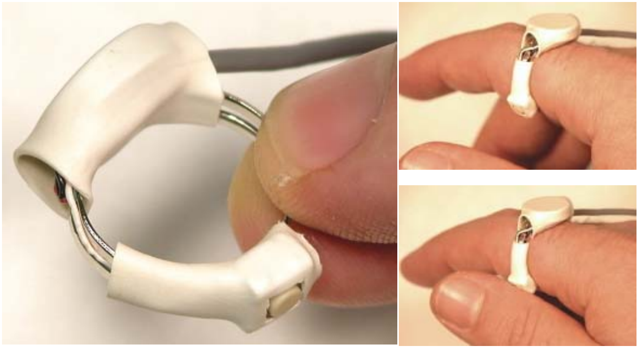

Figure 3: Wired finger rings used for user studies

We have developed both wireless and wired finger rings to test the communication protocol and for user studies.

The wireless prototype consists of a low-range transceiver (Radiometrix Bim2, same as on the Conversation Finder nodes), a microcontroller, a vibration motor, a micro switch, and a Lithium Polymer battery. Due to the transceiver footprint, its current size is 32x23mm, but could be easily miniaturized by using a smaller transceiver. The wired prototype (Figure 3) used for the user studies consists of a vibration motor and a micro switch, which are connected via flexible 3-conductor cable (attached to the wearer's elbow). Since people have personal preferences as to where to wear a ring, its diameter is adjustable. The switch position is variable as well: some prefer it on the side, some one the lower part of the ring.

USER STUDY AND PRELIMINARY FINDINGS

Is such a system useful or desirable? We cannot really answer that without deployment among a fair-sized group of people and observations into their normal lives. Nonetheless a small study can inform us as to whether people might find it useful at all, and whether, as we would expect, their interruptibility varies with social context.

If participants are given the means to anonymously veto upcoming cellphone interruptions by responding to a subtle pre-alert in the form of slight vibration on their finger ring, will they distinguish between different social settings? Will they more likely *disallow* interruptions in a cognitively demanding group-focused setting, and will they more likely *allow* interruptions from cellphones during 'group downtime'? Will a majority of the participants implicitly agree on when it is appropriate to get interruptions, and when it is not?

Pilot study

The pilot study was conducted to obtain information about parameter thresholds. Different vibration patterns were tested, and it was determined that a single vibration burst of one second on a participant's finger is perceivable yet not disruptive. Furthermore, it turned out that the ratio of collective award vs. individual award during the game (see below) is 1 to 10 in order to balance the behavioral motives. Because some participants had to suppress the reflex to press the ring switch when it vibrates, as in 'picking up a call,' a trial run for the game and ample try-out time was scheduled for the user study.

User study

Methodology

The 45-minute user study involved a simple card game. One experimenter distributed a deck of cards to a group of three participants. Then the cards had to be put down in a specific order, one by one, on a single pile in the middle of the table. Each game lasted 70 seconds, and a clearly visible clock showed the count down. The more cards the group could lay down, the more money each participant earned. For each card on the table, each participant received 5 cents. There were multiple games per session. In between the games, there were pauses for reshuffling and redistribution of the cards. Although the game was simple, it required the full attention of all participants; the pauses in between, instead, were low stress periods.

During the whole session—both during the games and the pauses—participants received short phone calls by a remote experimenter. These calls allowed the participants to earn additional money: they were asked a simple question ("What is 13 times 7?"), and if the participant—and only the participant on the phone—answered correctly, s/he received a 50-cent bonus.

Participants were given subtle pre-alerts in the form of a short vibration of their finger ring when any call came in (not just for their own cellphone). Each participant then had the chance to veto it anonymously by pressing the micro switch on his or her finger ring. Every participant was given the same pre-alerts, and at the time of the pre-alert no one knew who would get the call. The ultimate goal of the game was to earn as much money as possible, either from collective or individual rewards; deciding on which to focus was up to the participants.

All sessions were videotaped with multiple cameras and transcribed to obtain exact timestamps of all events (pre-alerts, calls, vetoes, etc.) In addition to the transcripts, all participants filled out pre and post study surveys (semantic differentials with 17 bipolar scales).

Results and discussion

The study consisted of two group sessions, each with three participants. In total, 30 pre-alerts were issued, 15 during the card games and 15 during off-times. The total length of games and pauses were equal. All vetoes across all groups

Table 1: Results

Setting	Pre-alerts issued	Vetoes received
Group game	15	8 (53%)
Pause	15	3 (20%)

were added up per setting (during card game, during off-times). In all but one case, there was either no veto or one veto per pre-alert. As Table 1 shows, vetoes happened more than twice as often (53%) in the high attention, collective activity setting than in the 'Pause' setting (20%). Even with our relatively low N, the mean differences between the two settings became statistically significant ($p=0.05$, $t(28) = 1.70$, single-tailed t-test): the participants indeed vetoed more during the games than during the pauses.

The semantic differentials, which attempt to measure the meaning of the general concept "cellphone interruption" (pre study) and the more specific "cellphone interruptions in this study" (post study) will be used later to measure the main connotational difference between generic cellphone interruptions vs. interruptions in our system's specific setting. Our preliminary sample size, however, does not allow for a factor analysis yet.

During the de-briefing, one participant voiced concerns that "random people, like in the bus, could disable my phone." The Conversation Finder nodes, which guarantee that only people in the same conversation with the user can veto and not just any person close by, were not necessary for this study, so the participants did not know about it.

Another participant objected that other people might (accidentally) veto important calls, e.g., from a hospital. It was explained to her that co-located people's veto is just one input of several for our conversational phone system that converses with the caller, and is trying to recognize emergency keywords such as "hospital," "accident," etc. at any point in the conversation, and would override vetoes.

RELATED WORK

Placing private alerting and sensing technology on a user's finger or wrist have been explored in the past; however, they are rarely both combined, and to our knowledge have not been used to manage interruptions collaboratively.

Miner et al.'s [6] Digital Jewelry project describes several versions of finger rings, both as input and output devices. Their LED GlowRing glows upon an incoming email message in varying colors, depending on the importance of the message. When the user touches the face of the ring, it sends a wireless signal to the user's LCD bracelet to display the face identity of the sender, and another signal to the user's earring (serving as a headset) to play back the urgent message. It is not clear, though, how much of this scenario is implemented. Our system uses peripheral alerting similar to GlowRing, but eventually is concerned about collective responsibility for cellphone interruptions.

Fukumoto's "FingerRing" [3] uses ring shaped sensors to detect fingertip typing by measuring the acceleration on each finger. "Whisper" [2] is a wrist-worn handset that is used by inserting a fingertip into the ear canal.

FUTURE WORK AND CONCLUSION

Although our current experimental design is based on an egalitarian approach, variations might be worth exploring: e.g., all participants of a conversation are alerted and allowed to veto except the user who owns the interrupting device; more than one veto is necessary to avoid an interruption (majority approach); different users have different weights in the vetoing process (which would require the identity of the vetoers to be disclosed).

The Conversation Finder nodes and Finger Ring controllers are components of a larger interactive call managing intermediary, which will also converse with the caller, relay voice instant messages in lieu of synchronous calls, and change behavior based on the identity of the caller and the history of previous phone calls to location. In this paper we have discussed how that intermediary can solicit input from the others whom the called party is speaking with, and in a small first step of evaluation confirmed that those others may use this power in an appropriate manner.

ACKNOWLEDGMENTS

We would like to thank Kimiko Ryokai for her help with conducting the user studies, Rachel Kern, Jae-Woo Chung, and Matt Hofmann for their general support, as well as Nena for her long-standing undivided support of our work.

REFERENCES

1. Basu, S. *Conversational Scene Analysis.* Ph.D. Thesis. MIT Department of EECS, September 2002.
2. Fukumoto, M. and Tonomura, Y. Whisper: A Wristwatch Style Wearable Handset. In *Proc. CHI 1999*, ACM Press (1999), 112-119.
3. Fukumoto, M., and Tonomura, Y. Body Coupled FingeRing: Wireless Wearable Keyboard. In *Proc. CHI'97*, ACM Press (1997), 147-154.
4. Horvitz, E., and Apacible, J. Learning and Reasoning about Interruption. In *Proc. of ICMI 2003*, ACM Press (2003), 20-27.
5. Love, S., Perry, M. Dealing with mobile conversations in public places: some implications for the design of socially intrusive technologies. *Ext. Abstracts CHI 2004*, ACM Press (2004), 1195-1198.
6. Miner, C.S., Chan, D.M., & Campbell, C.S. Digital Jewelry: Wearable Technology for Everyday Life. In *Proc. CHI 2001*, ACM Press (2001), 45-46.
7. Schmidt, A., Aidoo, KA., Takaluoma, A., Tuomela, U., Van Laerhoven, K., Van de Velde, W. Advanced Interaction in Context. In *Proc. HUC99*, Springer (1999), 89-101.