The Family Intercom: Developing a Context-Aware Audio Communication System

Submitted to Ubicomp 2001

Kris Nagel, Cory D. Kidd, Thomas O'Connell, Anind Dey, & Gregory D. Abowd

College of Computing & GVU Center
Georgia Institute of Technology
Atlanta, GA 30332-0280 USA
{kris, coryk, thomas, anind, abowd}@cc.gatech.edu

Abstract. We have begun an exploration of how ubiquitous computing technology can facilitate different forms of audio communication within a family. We are interested in both intra- and inter-home communication. Though much technology exists to support this human-human communication, none of them make effective use of the context of the communication partners. In the Aware Home Research Initiative, we are exploring how to augment a domestic environment with knowledge of the location and activities of its occupants. The Family Intercom project is trying to explore how this context can be used to create a variety of lightweight communication opportunities between collocated and remote family members. It is particularly important that context about the status of the callee be communicated to the caller, so that the appropriate social protocol for continuing a conversation can be performed by the caller. In this paper, we will discuss our initial prototypes to develop a testbed for exploring these context-aware audio communication services.

1 Introduction

Human-human communication is an essential part of our everyday lives. Advances in communication technology have enabled anytime-anywhere connection between people, but not always in the most socially acceptable fashion. In this paper, we focus on communication between family members, both collocated and distributed. While there are many tools enabling direct communication, there is little support for the appropriate social mediation of communication between persons with a trusting relationship, such as a family. Dad may use an in-home intercom to call Mom in her home office, telling her to come quickly and see the castle the children have built in the upstairs playroom. If Mom is in the middle of a phone call in her office, however, would Dad really want to interrupt her? An adult son may want to phone his mother living in another city to see how she is doing, but he does not necessarily want to wake her up from a nap in the process.

In the Georgia Tech Aware Home Research Initiative, we are investigating domestic services leveraging off perception technologies that deliver information on the whereabouts and activities of members of a family. For human-human communication, we are interested in how awareness of location and activity can facilitate both intra- and inter-home communication. We want to provide a variety of lightweight interfaces that facilitate a human's ability to decide whether a proposed conversation should be initiated or not. The Family Intercom project proposes a testbed for exploring a variety of intra- and inter-home communication scenarios. Our intent with the Family Intercom project is to explore how context-aware communication can support family communication, but before we can embark on authentic evaluation we need to provide a flexible testbed to support our experimentation. This paper discusses the challenges of producing an appropriate testbed and describes two prototype implementations that will lead to a more general platform for exploration.

We begin with a discussion of the challenges associated with family communication and the various technologies used to support this communication. We identify a design space for family communication based on the level of context available in the environment of the caller and callee. Our initial prototypes focus on conversations between individuals in a context-rich environment, such as the Aware Home [12], and between an initiator who is in a context-poor environment trying to communicate with an individual in a context-rich environment. We describe the design of communication prototypes for each of these two situations, addressing how an appropriate separation of concerns between the various phases of communication allows for the exploration of a variety of physical interfaces, modes of interaction, and mediation strategies. One of the important design considerations is to allow for the use of relatively sophisticated context sensing in advance of technologies that provide that context. We discuss how use of the Context Toolkit facilitates the simulation of complex context-sensing. We conclude with a discussion of related work and a description of our future work in this area of lightweight, context-aware family communication.

2 Understanding Home Communication

Longitudinal studies of census data on home computer ownership and usage show an increased use of personal computers across all demographic groups since 1984 and that the dominant use of personal computers is communication via the Internet [29]. Similarly, email was found to sustain ongoing relationships and strengthen some weaker relationships in an analysis of home Internet usage [15]. This quantitative data shows the value of technology in supporting communication at home, but does not provide insight into how technology facilitates more direct, synchronous human-human communication, especially within the home itself. However, the Casablanca project has explored new forms of home communication through lightweight ethnographic studies and by deploying prototypes within homes [9]. They report new approaches to home communication were welcomed, especially those using simple, lightweight interactions. Home users want technology to facilitate existing communication needs, while not obligating them to extended social interactions.

Another ethnographic study in the home highlights the value placed on communicative activities between collocated household members, although these are often in small time blocks and dispersed over multiple spaces within the home [17]. There is the problem of "space overload," which states that when technology is fixed in a particular location, as opposed to being distributed throughout the home, problems can arise over shared use of the space [21]. This localization does not afford the interaction and coordination of activities common in the everyday routine of a household. These intra-home communication patterns and the desire to decentralize technology both suggest that any computationally enhanced communications system should be accessible throughout a home, instead of being centered on the few places that might contain a personal computer or even localized to special-purpose appliances hung on the walls. The social interactions within the household are better served by distributed technology throughout the environment, that is, through ubiquitous computing technology.

Interpersonal communication within and between homes can also be examined from the psychologist's perspective of human interaction. Communication is viewed as the instrument of human transactions, including linguistic and nonverbal messages [13]. Messages across the different channels may be consistent or contradictory with each other. The message alone is not sufficient, the situation in which the interpersonal communication takes place equally defines the interaction [6]. With our current audio tools, there is not sufficient contextual information provided to the participants to support more complex interpretation of communication within relationships.

2.1 Limitations of Existing In-Home Intercom Systems

We have looked at several intercom systems to inform our exploration of intra-home communication. Intercom systems may be built into home telephones or operate as a completely separate system, sometimes integrated with a home security system. Intercoms most often support one-way audio connections, forcing an awkward protocol on the users in order to use a half-duplex technology for a conversation. More expensive intercoms incorporate two-way audio and monitoring capability. Some intercom systems have hands-free operation, either through a headset, a foot control, or preset voice commands, such as "talk".

Intercoms directly support place-to-place communication. While this is appropriate for some communication patterns (for example, monitoring or broadcasting), it is not suitable for person-person communication that is very common in the home. To reach a particular person, the intercom requires either a broadcast to all stations or a search through a sequence of stations in an attempt to reach the desired person. This polling through the stations usually requires the caller to press a button for each station (even in most hands-free systems); the call recipient is then able to respond hands-free, as the caller will have already activated the audio connection. Once this connection between two individuals is made, they are tied to those station locations for the duration of the communication; they may not roam from station to station without explicitly resetting the communication path.

Another feature of today's intercom is the monitor mode, where every station may listen to a designated station. Any speaker will automatically be connected to the monitored station. In this mode the entire intercom system is listening to the specified location and is not simultaneously available for conversation between other household members.

2.2 Limitations of Existing Between-Home Audio Communication

As we consider communication between family members in different homes, we see the same issues arising, only the need for context is more pronounced. Telephones provide a point-to-point connection, but do not provide location of the recipient to the caller. Mobile phones connect to the current location of the phone, usually collocated with the person owning the phone. Although a person can be reached, the caller has even less knowledge of the situation, since the location is dynamic. Call forwarding can also disguise the location of the recipient, along with cordless phones whose range is well into the yard. How can the caller determine if this is an appropriate time to call?

Instant messaging applications, such as ICQ [10], though almost exclusively used for textual communication, provide various context modes to indicate a person’s availability, such as away, busy, or available. ICQ can automatically set the context for several modes, but it ultimately requires from the need for explicit user action. The context provided to family members may be different than that shown to work colleagues, and information may vary based on the family hierarchy (e.g., adult to child versus adult to adult). Some instant messaging services support a hierarchy of buddy lists to provide varying information to encourage the use of this lightweight non-audio communication.

3 Requirements for the Family Intercom

3.1 Using Context to Mediate Communication

One problem with existing audio communication systems is that they require too much explicit activity on the part of the caller and the callee to establish and maintain a one- or two-way audio connection. Another problem is that the caller has no choice but to make an announcement to a particular location, via intercom in the house or phone call between homes. This can be a disadvantage, possibly disturbing an individual whom the caller would otherwise not wish to disturb. The caller is unaware of two pieces of important context:

- Where is the intended recipient located?
- What is the recipient doing, and is it OK to interrupt the recipient at this time?

A context-aware communication tool that could provide the caller with this kind of information would enable the caller to initiate only desirable conversations. Such a system can use context to automatically or manually mediate appropriate audio conversations. Other context, such as knowledge of ambient sound levels in the locations of the caller and recipient, can be used to set appropriate parameters for the conversation, such as an initial volume. Finally, during a conversation, a change in context for the caller or callee can be used to adjust important parameters of the conversation. All of these examples demonstrate the potential advantages of building communications services that leverage off of automatically sensed context.

3.2 Design Space of Context-Mediated Communication

These homes may be instrumented, as is our Aware Home [12], to provide richer contextual information or simply provide a communication portal, with little sensing available. We use the two extremes to define the communication tool design space:

Recipient Environment	**Conversation Initiator Environment**	
	Sensor-rich (e.g., Aware Home)	Sensor-deprived
Sensor-rich (e.g., Aware Home)	(1) Within-Home collocated Use context to mediate	(2) Caller identity context Receives rich context of recipient
Sensor-deprived	(3) Challenge to cue recipient	(4) Challenge to cue at both endpoints

Table 1: Design Space of Context-Mediated Communication

Here sensor-rich includes identity, location and activity information for each person, such as laying down, preparing food, talking with someone. The sensor-deprived site may have identity and limited location knowledge, such as in the home or beside a communication portal. When communication is viewed as a message sent from one person to another, each communication endpoint may reside in either type of environment, yielding a four-cell design space. Our first intercom prototype is deployed on one floor of the Aware Home. Our second prototype explores communication initiated by a family member from a remote, non-instrumented home to the Aware Home. The asymmetric sensing was also reflected in the feedback and interface provided at each endpoint. We have two cells of the design space not yet explored, when the non-sensored home is the recipient.

3.3 Motivating Usage Scenarios

People communicate with family members within a home and across households in many different ways. To better portray the communications patterns we hoped to support with a context-aware Family Intercom, we describe some relevant usage scenarios in detail before presenting our prototypes that support all of these capabilities. The first two scenarios occur within the Aware Home, where our model of interaction is made hands-free by providing voice interaction with the house intercom system. The final scenario describes inter-home communication, in which the initiating home has only caller and callee identity information. The recipient's home is the Aware Home, instrumented to support hands-free interaction and to provide context of activity, location and identity.

3.3.1 Scenario 1: Monitoring activity

Often, we want to monitor activity in a remote part of the house. A classic example is the baby monitor. In this case, the parent would initiate the monitoring by asking, “How is the baby doing?” or “What’s going on in the baby’s room?” When the intercom receives this request, it determines which room the baby is currently in and creates a one-way audio connection from the microphone in the baby’s room to speakers nearest the parent. If the parent moves to a different room, the baby monitor audio channel follows along, moving to the speakers closest to the parent who requested the baby monitor. When the parent no longer needs to listen in on the baby, he tells the house, “Stop the intercom.”

3.3.2 Scenario 2: Having a conversation within the home

From the kitchen, Mom sends Sally down to the basement to get some items from the pantry. Once Sally gets down to the pantry, she cannot find the items Mom sent her down to retrieve. Sally wants to ask for some clarification from Mom, but Mom cannot hear even if Sally yells. So, Sally instructs the house intercom, “House, I want to talk to Mom.” Meanwhile, Mom has set up a baby monitor connection to her younger son, Joey. She can hear Joey crying, so she departs to the family room to care for him. When the house recognizes the request from Sally down in the basement pantry, it then locates Mom, who has now moved to the family room. The house knows that baby Joey is also in the family room, so tells Sally, “Mom is now in the living room with Joey. Do you still wish to speak with Mom?” Sally guesses that Mom is changing Joey’s diaper because he was heard crying before she went down to the basement. Though Mom’s attention will be divided, Sally still wants to speak with her, so she responds, “Yes.” A two-way audio connection is established between Sally in the basement pantry and Mom in the living room. Sally asks Mom to help her determine which items to bring up to the kitchen. During the course of the conversation, Mom finishes with Joey and returns to the kitchen to see what else she needs Sally to bring up from the pantry. The conversation between Sally and Mom continues uninterrupted as both move about the house. As Sally finally returns to the kitchen where Mom is, the house determines that their remote conversation has ended and automatically terminates the audio connection between them.

3.3.3 Scenario 3: Having a conversation outside the home

Kim notices the digital portrait of his mother on the end table. Kim wonders how she is feeling today and touches the portrait to initiate a conversation. The portrait is actually a portal providing communication to the pictured family member and context from their remote home. The portal includes a flat touch screen enabling any household member to initiate a two-way audio connection to the remote family member and to see the communication status. The portal is also able to identify household members located near the portal. The portrait is replaced by a representation of this communication request, Figure 1 (transition A). Kim's mother is at home and available to talk. In the mother's home, a background chime alerts mother to the conversation request. Kim says, "Hi, Mother, this is Kim". "Oh, Kim, good to talk with you." The conversation continues until either party explicitly stops it, Kim by touching the stop sign on the display or mother via voice command.

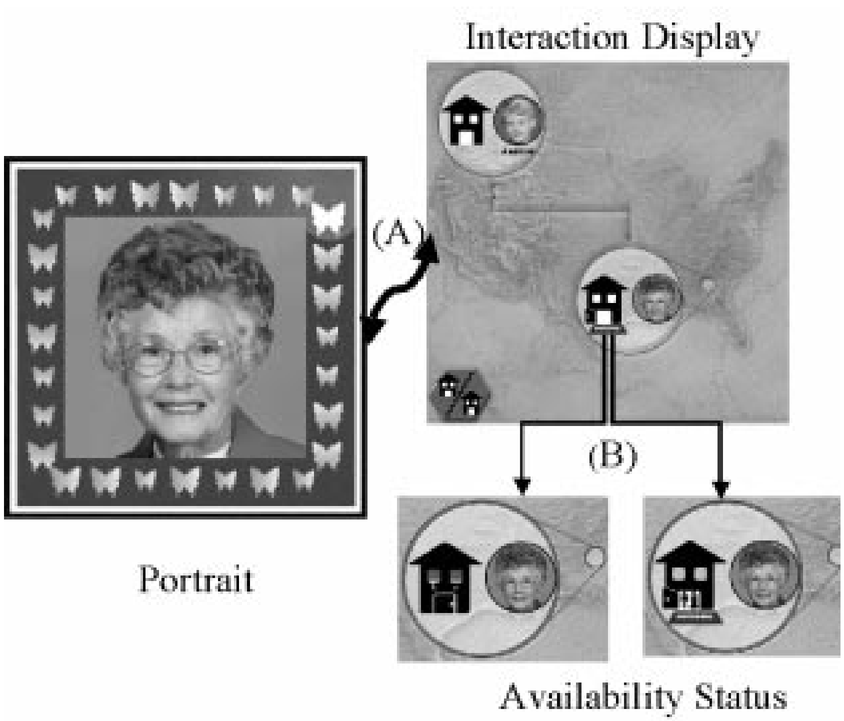

Figure 1 - (A) Touching portrait opens communication
(B) Stylized House indicates availability, not available, wait a moment

3.4 Highlights from the Usage Scenarios

These scenarios highlight the major context-aware features of the Family Intercom. The family intercom should provide support for:

1. Initiating communication

a. distinguish who is the initiator of a request (the caller) and who is(are) the intended recipient(s);
b. gathering and interpreting pertinent context information about the recipient and their environment to help the caller determine whether it would be appropriate to continue with the set-up of the audio connection;
c. mediating the initiation of an audio connection by providing adequate feedback through the best interface mode to the caller;
d. setting up a variety of one- and two-way audio connections between locations within the house and/or between homes, including multiple connections at once;

2. Mediating communication
   a. create audio connections that move appropriately with the movements of the caller and recipient(s) in instrumented settings;
   b. remove a superfluous audio connection when the two callers move within earshot (e.g., in the same room)
3. Terminating the audio connection
   a. explicit request;
   b. implicit cues.

4 Exploring the Within Home Intercom

The initial context-aware intercom prototype has been installed in the Aware Home. We will discuss it in terms of environmental instrumentation, software design, and the evolving prototype.

4.1 Environmental Instrumentation

The hardware design consists of an electronically configurable audio system with the following components:

- Speakers: There are pairs of speakers mounted in the ceiling of each room throughout the house. This allows audio output of a conversation to any room where a person may be located.
- Microphones: Currently each person wears a wireless microphone while walking around the house. We have ceiling microphones placed in each room, but we require proper echo cancellation, not currently implemented, to address audio feedback for two-way conversations. Once addressed, we will move to the open-air microphone infrastructure.
- Audio Switch: The speakers and microphones feed into an audio switch. This allows us to control the input and output routing through a serial connection to a desktop computer. The audio switch supports simultaneous point-to-point connections (two people speaking or one person listening to another), as well as one-to-all broadcast. We are looking into other available hardware and software

solutions that will give us more flexibility in the amount and type of connections we can create within the house (such as multi-party conversations across three or more rooms).

- Positioning system: We initially used a positioning system from PinPoint Corporation, the 3D-*i*D® Local Positioning System [23]. This system is deployed over an entire floor of the house, giving readings at a room-level accuracy for tags worn by individuals. We have since explored other research and commercial positioning systems.

4.2 Software Design

We are interested in rapidly prototyping and easily evolving the context-aware features of the intercom application. To facilitate these goals, we designed the intercom software using the Context Toolkit [25,3]. The Context Toolkit provides several useful abstractions (widgets, aggregators, services, and interpreters) for organizing the functionality of the intercom software and greatly eases the incorporation of sensed context. The overall organization of the software is shown in Figure 2.

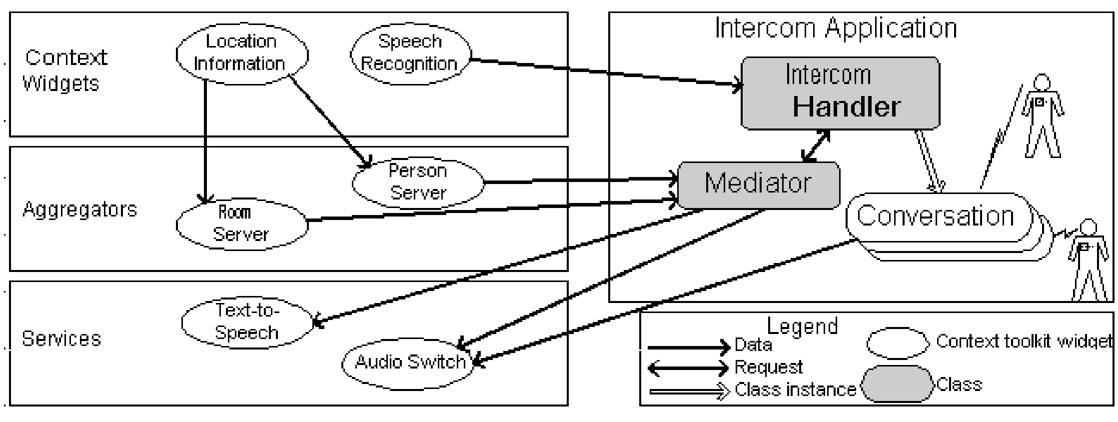

Figure 2 – IntercomApplication Software Design

Context widgets represent abstractions over sensors that hide details of how sensing and interpretation of the environment occurs. In our intercom application, the following widgets are used:

- Location: This widget was initially a wrapper around the entire PinPoint 3D-*i*D® Local Positioning System, providing an interface that delivers information automatically to interested software components when individuals leave or enter rooms in the house. As we have incorporated other ways of getting location information in the house, we have added more instances of location widgets.

- Speech Recognition: In the current prototype, each member of the household wears a wireless microphone that is connected to a dedicated PC running the IBM ViaVoice® software. We use the Java Speech Grammar Format [11] to designate grammars that the ViaVoice® engine will recognize. The Speech widget is used for other applications in the home, and for intercom requests it produces formatted context messages of the form: ID: “Jamie” COMMAND: “intercom,” and DATA “2-way, Kim” or ID: “Dad,” COMMAND: “intercom,” and DATA: “broadcast”.

Aggregators collect information for relevant entities of an application. In the case of the intercom system, the relevant entities are rooms in the house (Room Aggregators) and residents of the household (Person Aggregators). The Room Aggregators currently collect information on who is in a given room. For example, the Kitchen Aggregator may know that both Mom and Sally are in the kitchen at the end of the Scenario 2 above. A Room aggregator can also hold other information in the room, such as appliance status, ambient noise level, or even an interpretation or prediction of a high-level activity (for example, dinner preparation). Person aggregators currently hold information about where a person is in the house. The Joey Aggregator, for example, will contain the information that he is in the living room in Scenario 2 above. In that same scenario, the Mom Aggregator would indicate that she is currently engaged in a conversation with Sally.

Services provide actuation to the intercom system. In this prototype, we take advantage of two services, the Audio Switch and a Text-to-Speech engine. The Audio Switch service allows us to send commands to the switch in the form of **Route(input, output)** to connect the microphone of one person to the speakers in some room. This service also allows us to adjust the volume of individual connections and remove connections when a conversation is finished. The Text-to-Speech service allows us to send spoken feedback to the users during their use of the system. For example, when a conversation is being requested, the system can tell the initiator “Go ahead” when the connection has been established and “Good-bye” when the conversation has been completed.

Every software component described so far can be shared by any context-aware application. The intercom application itself supplies the appropriate context-aware functionality, and consists of three components, the Intercom Handler, the Mediator and Conversation widgets.

The Intercom Handler subscribes to the Speech Recognition context widget with a filter so that it receives only commands recognized as intercom requests. When an intercom request is received, the Intercom Handler will query the Person Aggregator of both the caller and recipient to figure out which rooms are involved in the connection. The Mediator component then determines whether the connection should be established. If the connection is approved, a Conversation context widget is created to manage the connection.

A Conversation widget requests the Audio Switch service to make the appropriate one- or two-way audio connections. This widget also subscribes to the relevant Person aggregators of the caller and recipient to be informed when either changes rooms. If a room change is detected, the Audio Switch is instructed to alter the connection path in order to follow the caller and recipient. It also informs those Person Aggregator that each person is now engaged in a conversation by setting the value of an activity context variable. The Conversation widget subscribes to the Speech Recognition widget to be informed when either the caller or the recipient requests termination of the connection. When caller and recipient become collocated during a 2-way conversation, the conversation is automatically terminated.

The Mediator uses available context and simple heuristics and negotiation with the caller to determine whether an intercom request should be approved. Currently, it can query the Room Server where the recipient is located to determine if others people are also located in that room. If not, the intercom request is approved. If the recipient is not alone in the room, then feedback of this context is spoken back to the caller, as illustrated in Scenario 2. The caller is asked whether to proceed with the intercom request. This is an example of informing the caller of relevant context so that she might apply an appropriate social protocol for continuing with the initial request.

4.3 Evolving the Prototype

The first prototype context-aware intercom is a research vehicle for exploring intra-home communication patterns. The design of the prototype with the Context Toolkit facilitates the addition of new and different contextual information. In addition to the PinPoint 3D-*i*D® system, we have installed the WEST WIND environmental tracking system [16] and the Texas Instruments RFID® System [28]. Each new positioning system is encapsulated in a unique Position widget in the context layer of the house. All available Position widgets can be linked to any subscription request for positioning information, resulting in no change to the intercom application. While the location systems are engineered and installed, there is still a need to provide location information for demonstration and development work. Just as any number of position widgets can provide location, a simulated position widget can be used in place of a real system. The design and use of the Context Toolkit infrastructure has enabled the prototype to evolve as the sensing emerges.

We currently have created a hands-free interface to the intercom using voice recognition of simple commands. However, we have seen situations when an alternate interaction modality is more appropriate. For example, in a crowded room with lots of ambient noise, initiation of a conversation can occur through different modes of interaction. Any different interaction modalities for initiating a conversation can be added simply by wrapping the functionality as a context widget that produces the standard intercom message, as defined above. The intercom application does not have to change at all, and all of these interaction modalities can co-exist. Similar variety is possible for terminating a conversation.

The most interesting aspect of the context-aware intercom is its ability to use context to mediate the conversations. This occurs during the initiation phase, when a caller tries to connect with some recipient and when a conversation is active and its participants change some part of their context. In the prototype described above, only location information is used to mediate the initiation of a conversation. When the recipient is collocated with one or more other people, the conversation request is not automatically executed, but requires further confirmation by the caller. The prototype intercom provides a concrete means of exploring many types of context and their contribution towards recognizing significant human activities.

5 Exploring the Between Home Intercom

In the second prototype developed for the Family Intercom, we wanted to explore communication between homes. Communication is initiated from the sensor-deprived, remote home to the Aware Home. At each home the intercom is able to:

- Identify the initiator of the request and the intended recipient
- Mediate the initiation of the audio connection, according to availability of the recipient
- Provide feedback appropriate to this initiator, based on family hierarchy
- Set up a two-way audio between the communication portal and the Aware Home, when appropriate
- Terminate the audio connection based upon explicit request or other appropriate implicit cues

5.1 Environmental Instrumentation

The initial remote communication portal was designed to be portable, requiring only an Internet connection for operation. The hardware consists of the following components:

- Microphone: A boundary microphone is place on the tabletop in front of the display, to capture the inputs as the family speaks to the photo on the display.
- Speakers: There is a pair of speakers adjacent to the display unit. This provides audio output of the conversation from the sensored-home. The placement of the speakers is usually beside or behind the display unit, so the voice comes from the image of the remote person. Audio feedback is avoided placing the speakers and microphone judiciously, depending on the furnishings of the environment.
- Identity and Positioning System: We used the Texas Instruments RFID® system [28] to provide the identity of individuals in close proximity to the communication portal.
- Touch Screen: We use a flat-screen monitor with a touch screen for input and feedback.

5.2 Software Design

We again used the Context Toolkit to design and implement the prototype, developing complementary components for each end of the communication at the remote portal and the Aware Home (Figure 3). The remote panel is based on the Digital Family Portrait, that also displays the portrait and a qualitative perception of activity for the family member pictured [20]. The Digital Family Portrait, written in Squeak, was augmented with modules to provide interoperability with the Context Toolkit. Server Module and Status Widget were implemented to emulate the corresponding Context Toolkit interfaces. The Server Module collects information from the relevant Context Toolkit entities: the Family Intercom Status Widget and the Identity Widget. Similarly, the Status Widget was implemented to emulate the Context Toolkit widget interface. The Status Widget automatically delivers the interactions and current status information compatible with the delivery mechanism of the Context Toolkit. The status provided is user identity, user request to initiate or terminate a conversation, and the availability status for the remote family member (as received by the server described below). The Interface Manager handles the user interactions and displays the appropriate interface for the family member initiating the conversation. The augmenting of the Digital Family Portrait with these three modules, is sufficient to provide a communication portal for the Family Intercom.

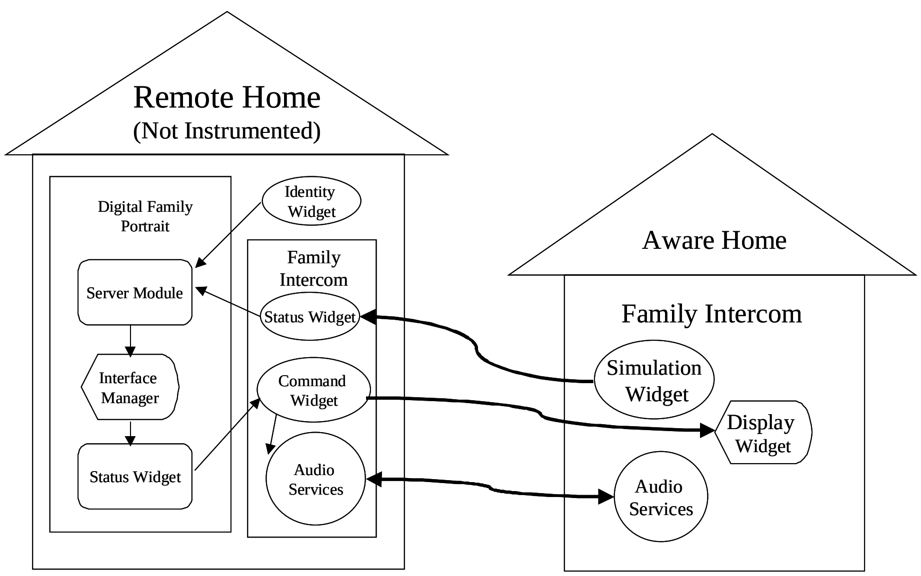

Figure 3 – Between Home Architecture, built into existing Digital Family Portrait

The new Digital Family Portrait modules interact with a collection of Context Toolkit widgets at the remote home. At the remote home, the Family Intercom application includes Status Widget, Command Widget and Audio Services. This application manages the audio communication connections based on the status and commands it automatically receives from the Status Widget in the Digital Family Portrait. The components are:

- Status Widget: This widget provides the status for the family member in the portrait and the command to terminate, as this data is received from the Aware Home Simulation Widget.
- Command Widget: This widget gathers the command and status data from the Digital Family Portrait Status Widget. The intercom application uses this data to request the appropriate audio services to manage the communication.
- Audio Services: This module provides control for the audio connection, including initiating a call, checking the connection status, and controlling the sound level.

Additionally, the Identity Widget at the remote home provides the identity of the family member interacting with the portrait. This widget currently uses Texas Instruments RFID® system [28]. When a family member wearing one of the RFID tags is near the portrait, the Identity Widget informs the Digital Family Portrait Server Module.

The Aware Home provides context to the remote home. The availability status of the callee is simulated. The sensing in the instrumented home does not yet provide the context required to mediate the communication requests. The components in the instrumented home are:

- Simulation Widget: This module receives GUI inputs simulating the availability of the resident for conversation. It automatically delivers updated status to interested software components, where the remote Family Intercom subscribes as an interested component.
- Display Widget: This component receives updates from the remote Family Intercom. The status of the communication (started or stopped) and the identity of the caller are displayed for the benefit of the "wizard" operating the Simulation Widget.
- Audio Services: This module provides control for the audio connection, including controlling the sound level.

We used Microsoft's NetMeeting®[18] for the audio communication in both the Aware Home and the remote home. The Audio Service Module provides a wrapper to the NetMeeting functions needed.

5.3 Evolving Between-Home Intercom Prototype

The between-home prototype Family Intercom is a research vehicle for exploring context-mediated inter-home communication. The first iteration was designed as a testbed for investigating communication interfaces from the remote, non-instrumented home, using context gathered from the Aware Home. While the initial interface is a digital portrait, the Family Intercom can easily be modified to accept commands from other modes of interaction. A tablet PC in the kitchen may provide a very different display and use voice recognition of simple commands. We are able to replace the touch screen with a voice command interface, by simply reusing the Speech Widget the Aware Home uses. Each new interaction modality can be wrapped in a context widget that delivers the standard intercom message. This design provides the flexibility to explore alternate interfaces to communication across households.

In addition to evaluating reliable and natural communication interfaces, we use mediation cues in each phase of communication. The adult receives additional textual status information regarding the remote relative, to facilitate human-mediation initiating a conversation. This text information is withheld from the child who may not understand the content or the private, personal nature of messages. The connection is mediated based on three states: available, not available, or momentarily available. This prototype provides a concrete means of evaluating these mediation strategies, but can support other states as additional context is accessible. While the initial prototype is focused on the remote home, the privacy concerns of the Aware Home occupant are also addressed. The callee has a tangible representation of the shared context to gauge their comfort.

The use of simple simulation widgets in place of sensors assist in determining how complex context can be inferred from a variety of sensing. This prototype uses the simple simulation of availability status, but in future versions we will use an assortment of sensors and heuristics to determine availability in each room. At the remote home, the caller identity is used to mediate and is determined by proximity of an RF ID tag to the portrait. In the future, fingerprint or voice identification may be used instead. When either of these systems are available, associated widgets will provide identity to the intercom, with no change to the application.

The audio infrastructure differs between the intra- and inter-home intercom. We intend to deploy both intercoms on a common digital-audio framework. The two prototypes provide the basis for determining common functionality such a framework will need to support. While the audio requirements may vary across environments, the Family Intercom infrastructure should support both.

6 Discussion

The two example Family Intercom prototypes contribute to our overall research agenda in designing applications of ubiquitous computing. This experience has helped us to:

- better understand how to design and engineer an evolving context-aware application,
- extend the use of the Context Toolkit to enable application development in the absence of adequate sensing technology,
- demonstrate ubiquitous computing technologies providing more invisible support of everyday activity.

6.1 Evolving the Application

The current Family Intercom prototypes provide flexible and extensible infrastructure for investigating domestic communication. Though we have not obtained authentic use in a domestic setting, we have gained significant experience. For example, the between-home prototype was part of a four-day exhibit with hundreds of visitors. From this long-term interaction we are able to see the reliability of the infrastructure and how the interaction is perceived as a new approach to communication between homes. While the reactions to using the portrait to talk with a family member were favorable, several short-comings were also revealed. The lower quality audio as compared to the telephone is one such problem, but voice over IP and connectivity is improving rapidly. The immediate connectivity did not provide the initiator a cue to when to begin talking. There were awkward inquiries, "Hi... Are you there"? There is no remote phone to ring and no off-hook notification, the initiator simply awaited an audible greeting from the recipient. This sort of interchange stumble was particularly evident in the "wait" state, when the initiator received no feedback until the recipient was able to respond. In both cases, we need to investigate feedback mechanisms to make sure the user knows the state of the connection. The modification of initiation phase is cleanly separated from the other components, so a change to this start up mechanism will mainly depend upon how to provide feedback to each endpoint.

Each prototype has been designed to accommodate changes in the context data provided. Each uses location from a single positioning system as a form of context. Any new positioning system can be encapsulated in a unique Position widget in the context layer of the house. The room and person servers will automatically link to any subscription request for positioning information, resulting in no change to the intercom application. This same principle is true for the interface. The initial voice recognition of simple commands has been supplemented with a GUI control, suitable when speech is more difficult. We are also investigating the use of a wireless handheld device to provide a portable intercom control unit. All of these interaction modalities for initiating a conversation can be added simply by wrapping the functionality as a context widget that produces the standard intercom message, as defined above. The intercom application does not have to change at all, and all of these interaction modalities can co-exist. Similar variety is possible for terminating a conversation.

6.3 Simulating Context in the Context Toolkit

The Family Intercom uses simulation as an extension of the abstractions the Context Toolkit supports. The Aware Home does not yet have enough sensing in place to provide context for the intercom's mediation. We have used fake sensor data, where the actual instrumentation is not available. In the simplest use of simulation, each sensing widget can be replaced by a simulation widget. To any context-aware application, the simulation widgets are identical to widgets connected to real sensors.

“Wizard of Oz” experiments are more complex simulations, where human operators are able to interpret a situation correctly and set appropriate context attributes. This allows experimentation in the absence of very sophisticated sensing and perception of human activity. The application developer can evaluate the impact of a proposed sensing technology in advance of it being available, a very valuable prototyping advantage in context-aware computing. Simulation widgets can also aid in designing the appropriate sensing, avoiding the costly installation only to find this is not the appropriate context to provide.

While simulation has been very useful, there are limitations to its use. Situations easily assessed by a human for the appropriate social protocol or context may be impossible for the computer to perceive with the necessary accuracy and reliability. For instance, two people are in a room with a television set turned on; is the conversation heard between the two occupants of the room or between digital images on the screen? A person entering the room can easily distinguish who is having the conversation, but this is a difficult problem in general.

7 Related Work

Our exploration of context-mediated human-to-human communication extends current research in computer-mediated informal communication, audio-only social spaces, and use of context to mediate communication. There have been several desktop conferencing tools to support informal awareness, such as Portholes [5] and Montage [27]. While each of these provides some situation awareness and a lightweight interface, each is a desktop application designed for distributed workgroups and requiring explicit user interaction with the desktop box and is not as appropriate in the home. These systems provide computer-mediated communication through the existing network of computers, but our intercom is characterized by its communication through an aware-environment. Our exploration of human-human communication is influenced by the concept of environment-mediated communication (EMC) [8]. The design space of EMC was motivated by the use of mobile devices providing location-independent computer-mediated communication. While our intercom is not a mobile device, it provides mobile conversations between users within a specific locale, in this case the home. The context-aware intercom design is influenced by the positive social interactions with desktop computer-mediated communication and the environment-mediated communication framework.

Thunderwire, a lightweight and always-on audio-only space, promoted informal and impromptu social interactions [1]. However, the interface did not provide sufficient cues to identify who was present or to remind users to disable the system when privacy was desired. The context-aware intercom is not always on; the connection persists for the duration of the conversation, which can be explicitly (through voice command) or implicitly (as the result of collocation or other context-related activities) terminated. This affords a more private conversation, since those hearing the conversation are those in the same rooms of the house as the users. Our intercom is designed to assist communication within a pre-existing social entity, the household, where the identity of those present is usually known. By providing mediated communication based on location and activity information, our intercom provides a more private connection between users.

Talking in Circles, another audio conferencing environment supporting natural interactions, provides spatial cues to ameliorate some of the problems with group membership [24]. It uses lightweight multi-modal communication to focus on social communications, enabling a cocktail-party-like atmosphere. The spatial cues offered in this desktop environment enhance the moving about between groups and the dynamic aspect of the communication. Our intercom is not focused on building a social group, rather it leverages existing social relationships to enable interpersonal communication. Babble is a textual conversation system developed to investigate knowledge in a social context [7]. The Family Intercom shares the design characteristics of social translucence: visibility, awareness and accountability.

In addition to the social interaction within audio environments, the use of contextual information to enhance communication has been used in other research. Calls.calm provides a means for the caller and callee to interact and determine a good choice of time and communication channels [22]. This system leaves the interpretation of availability to the call participants, not to the computer mediating the communication. Context-Call provides the caller with context information for the person they want to call [26]. Once the potential caller has this context information, the application allows the caller to decide the mode of communications, continue the call, leave a message, or even cancel the call. Each person explicitly sets the context information, so the accuracy of the context depends upon the person. People may forget to update their situation information or may provide an inappropriate context. The *live addressbook* is another application providing caller with awareness of the intended recipient's personal presence [19]. It uses a "Buddy list", similar to instant messaging systems, but uses a combination of prompting and explicit user inputs. Our intercom is designed to benefit from both explicit interaction and implicit cues given by sensed and interpreted context. This context is used by the application to promote more socially appropriate mediation of requests.

Our use of the Context Toolkit to support simulation shares certain goals with other prototyping tools. Suede is a Wizard of Oz prototyping tool for speech user interfaces [14]. It provides a graphical framework to quickly develop speech prototypes, supports Wizard of Oz use of the prototype, and records the test data. Rather than support one type of modality of interface, we are looking at a more generic approach to prototype context for any type of sensing. Suede would be useful in developing the one modality of command input to our communication tool, but we need other modes and to simulate data from environmental sensing devices. QuakeSim uses QuakeArena to simulate context normally provided as a real-world service and also builds on the Context Toolkit to deliver the simulated information [2]. It is used to test, evaluate and demonstrate context aware application services. The Quake foundation provides a more realistic simulation environment, using 3D graphic models to portray real-world geographical sites and participants in the situation. However, this system appears to limit itself to replacing real-world sensing data. In addition to simulating actual sensing data, we also see the benefit to simulate devices that have not yet been constructed, to determine if the advantage they will provide is worth the resources used.

8 Conclusions and Future Work

We have presented the design and implementation of two systems to support family-based audio communication in the Aware Home. We have investigated intra- and inter-home scenarios in an attempt to influence a long-term research agenda on family communication. The main contributions of this work are:

- the motivation for a better home communication system than traditional intercom telephony systems;
- the use of context to mediate the initiation, management and termination of one-and two-way audio conversations;
- the development of research prototypes that can endure changes in technology and mediation strategies that are expected to result from future evaluation through real use;
- the identification and use of simulated context to support rapid prototyping and provide direction to sensing research; and
- a demonstration of how ubiquitous computing technology can provide invisible interfaces for everyday activities.

Our immediate research goal is to subject one or both prototypes to authentic use, so that we can learn what changes in mediation strategies and feedback are appropriate and which forms of context we need to more automatically incorporate into the Aware Home. A longer-term goal is to merge our existing prototypes into a single framework for exploring the rich space of family communication services. This space includes the remote/collocated dimension discussed in this paper as well as a synchronous/asynchronous dimension. This testbed will allow research into a variety of interaction styles for communication beyond the simple hands-free and GUI examples shown in this paper. It will also allow us to examine representations of context, specifically those aspects of context dealing with human activity, that support the sophisticated forms of social mediation that appear to be important for managing conversations.

Acknowledgments

This work is supported by the Aware Home Research Initiative, an industrial consortium coordinated by the Future Computing Environments Group at Georgia Tech. We are particularly thankful for the collaboration by Motorola Labs which has influenced the directions we have taken on family communications. The Aware Home research is conducted in the Georgia Tech Broadband Institute Residential Laboratory under the direction of Nikil Jayant and Gregory Abowd. Dr. Abowd's research in ubiquitous computing is partially supported by DARPA/ITO under the Information Technology Expeditions, Ubiquitous Computing, Quorum, and PCES programs, and by the National Science Foundation (grants 9703384, 9818305 and 0070345).

References

1. Ackerman, Mark, Debby Hindus and Scott D. Mainwaring. (1997) “Hanging on the ‘Wire: A Field Study of an Audio-Only Media Space. *ACM Transactions on Computer-Human Interaction*, Vol. 4, No. 11, (March 1997) pp. 39-66.
2. Bylund, Marcus, and Fredrik Espinoza. (2001) Using Quake III Arena to Simulate Sensors and Actuators when Evaluating and Testing Mobile Services. In Proceedings of Human Computer Interaction, 2001, Seattle, Washington, April 1-4, 2001, pp. 241-242
3. Dey, Anind K., Daniel Salber and Gregory D. Abowd. (1999) A Context-based Infrastructure for Smart Environments. In the *Proceedings of the 1st International Workshop on Managing Interactions in Smart Environments (MANSE '99)*, Dublin, Ireland, December 13-14,. pp. 114-128.
4. Dey, Anind K. and Gregory D. Abowd. (2000) CyberMinder: A Context-Aware System for Supporting Reminders. In the *Proceedings of Handheld and Ubiquitous Computing (HUC2K)*, Bristol, UK, September, 2000, pp. 172-186.
5. Dourish, Paul and Sara Bly, (1992) Portholes: Supporting awareness in a distributed work group. In *Proceedings CHI '92*, pp. 541-547.
6. Duke, Marshall P. and Stephen Nowicki, Jr. (1982). A Social Learning Theory Analysis of Interactional Theory Concepts and a Multidimensional Model of Human Interaction Constellations. Handbook of Interpersonal Psychotherapy, Pergamon Press, Elmsford, New York, pp. 78-94.
7. Erickson, Thomas and Wendy A. Kellogg. (2000) Social Translucence: An Approach to Designing Systems that Support Social Processes. In ACM Transactions on Computer-Human Interaction, Vol. 7, No. 1, March, 2000, pp. 59-83.
8. Gellersen, Hans-Werner, Michael Beigl and Albrecht Schmidt. (1999) Environment-Mediated Mobile Computing. In *Proceedings of SAC (SAC '99)* San Antonio, Texas, pp. 416-418.
9. Hindus, Debbie, Scott D. Mainwaring, Nicole Leduc, Anna Elisabeth Hagström, and Oliver Bayley. (2001) Casablanca: Designing Social Communication Devices for the Home. In *Proceedings of CHI '01*, March 31-April 4, 2001, Seattle, WA, pp. 325-332.
10. ICQ (“I Seek You”) homepage. Available at: http://www.icq.com/
11. Java Speech Grammar Format Specification. Available at http://java.sun.com/products/java-media/speech/forDevelopers/JSGF/index.html.
12. Kidd, Cory D., Robert J. Orr, Gregory D. Abowd, Christopher G. Atkeson, Irfan A. Essa, Blair MacIntyre, Elizabeth Mynatt, Thad E. Starner and Wendy Newstetter. "The Aware Home: A Living Laboratory for Ubiquitous Computing Research" In the Proceedings of the Second International Workshop on Cooperative Buildings - CoBuild'99.

13. Kiesler, Donald F. (1982). Interpersonal Theory for Personality and Psychotherapy. Handbook of Interpersonal Psychotherapy, Pergamon Press, Elmsford, New York, pp. 3-24.
14. Klemmer, Scott R. Anoop K. Sinh, Jack Chen, James A. Landay, Nadeem Aboobaker, Annie Wang. (2000) SUEDE: A Wizard of Oz Prototyping Tool for Speech User Interfaces. Proceedings of the13th annual ACM symposium on User Interface Software and Technology, 2000, Pp. 1 - 10.
15. Kraut, Robert, Trdas Mukhopadhyay, Janusz Szczypula, Sara Kiesler, and William Scherlis. (1998) Communication and Information: Alternative Uses of the Internet in Households. In *Proceedings of CHI '98*, pp. 368-375.
16. Lyons, Kent, Cory D. Kidd and Thad E. Starner. (2000) WESTWIND: Georgia Inst. of Technology GVU Center Technical Report GIT-GVU-TR-00-15. September.
17. Mateas, Michael, Tony Salvador, Jean Scholtz and Doug Sorensen. (1996) Engineering Ethnography in the Home. In *CHI '96 Companion Proceedings*, pp. 283-284.
18. Microsoft Corporation, NetMeeting, http://www.microsoft.com/windows/netmeeting/
19. Milewshi, Allen E. and Thomas M. Smith. (2000) Providing Presence Cues to Telephone Users. In CSCW 2000 Proceedings, December 2-6, 2000, Philadelphia, PA, pp. 89-96.
20. Mynatt, Elizabeth D., Jim Rowan, Annie Jacobs, and Sarah Craighill. (2001) Digital Family Portraits: Supporting Peace of Mind for Extended Family Members. In Proceedings of CHI 2001, March 31-April 4, 2001, Seattle, WA, pp.333-340.
21. O'Brien, Jon and Tom Rodden. (1997) Interactive Systems in Domestic Environments. In *Proceedings of Designing Interactive Systems (DIS '97)*, Amsterdam, The Netherlands, pp. 247-259.
22. Pedersen, Elin Ronby (2001). Calls.calm: Enabling Caller and Callee to Collaborate. In Proceedings of Human Computer Interactions, 2001, April 1-4, Seattle, WA., pp. 235-236.
23. PinPoint Corporation. 3D-iD® Local Positioning System. http://www.pinpointco.com.
24. Rodenstein, Roy and Judith S. Donath. (2000) Talking in Circles: Designing A Spatially-Grounded AudioConferencing Environment. In *Proceedings of CHI '2000*, pp. 81-88.
25. Salber, Daniel, Anind K. Dey and Gregory D. Abowd. (1999) The Context Toolkit: Aiding the Development of Context-Enabled Applications In *Proceedings of CHI'99*, pp. 434-441.
26. Schmidt, Albrecht, Antii Takaluoma and Jani Mäntyjärvi. (2000). Context-Aware Telephony over WAP. Springer-Verlag, London, Ltd., *Personal Technologies*, Vol 4, 4, pp. 225-229.
27. Tang, John C., Ellen A. Isaacs and Monica Rua. (1994) Supporting distributed groups with a Montage of lightweight interactions. In *Proceedings of CSCW '94*, pp. 23-34.

28.Texas Instruments, RFID System, http://www.ti.com/tiris/
29.Venkatesh, Alladi, Eric Shih, and Norman Stolzoff, (2000). A Longitudinal Analysis of Computing in the Home. In *Proceedings of International Conference on Home Oriented Informaics and Telematics, (HOIT 2000)*, Wolverhampton, UK, pp. 205-215.