Experiences with the Design, Fielding, and Evaluation of a Real-Time Communications Agent

Eric Horvitz, Johnson Apacible, Muru Subramani,
Raman Sarin, Paul Koch, JJ Cadiz, Attila Narin, Yong Rui
Microsoft Research
One Microsoft Way, Redmond, WA USA
{horvitz; johnsona; murus; ramans; paulkoch; jjcadiz; attilan; yongrui}@microsoft.com

ABSTRACT
We review the design and evaluation of a fielded communications agent that routes telephone calls for several hundred people at our organization. In an advanced mode of operation, the system performs a cost-benefit analysis to balance the priority of calls with the context-sensitive cost of interrupting users. We discuss the challenges with developing a communications agent spanning client-side software and telephony infrastructure. We focus on metaphors and tools that allow users to specify priorities and costs of interruption. We report on the results analysis of surveys and monitored logs of preferences and activities.

1. INTRODUCTION
Today, people seeking real-time communications, such as telephone conversations with others, make personal decisions about the best timing and channel of the communication. They initiate a call based on their own needs and preferences—as well as on their intuitions about the preferences and context of the person being contacted. Attempts to communicate are often suboptimal for both the initiator and recipient of a communication attempt. Attempts by a caller to establish real-time telephony may interrupt the recipient of the communications attempt during a poor time or frustrate the caller with a voice message capture that may lead to costly delays for both participants. Recipients of calls employ multiple methods to filter incoming communications selectively. Some people may employ well-trained assistants, while others rely on the manual screening of incoming telephone calls. Limiting or deferring real-time communications so as to minimize disruptions and maximize privacy is only one piece of the challenge. Depending on the caller and the situation, contactees may often desire to be reached in real time rather than be missed by a caller. We have all become accustomed the frustration about the volley of attempted communications, referred to as "playing phone tag."

The Bestcom project centers on the development of communication agents that can serve as automated personal secretaries for users. Research in this realm includes efforts to integrate desktop software with networking and telephony infrastructure. Other efforts in the spirit of Bestcom on context-sensitive communication include the work on the Nomadic Radio project [7]. Bestcom differs from earlier work in its focus on developing cost-benefit analyses to identify utility maximizing communication actions.

For telephony, communication actions include putting a call through to an office phone, routing of calls to a best number, recording messages, and the rescheduling of calls for a later time. Our work to develop call-handling agents for telephony includes efforts to develop event-sensing machinery that can provide an understanding of a user's situation, including events drawn from a user's calendar, system activity, and perceptual sensors such as microphones and cameras. We also have addressed challenges with extending legacy communications infrastructure, in an attempt to integrate context-sensitive call-handling with existing corporate PBX systems and voice over IP (VOIP) protocols. We focus in this paper on the critical challenge of formulating designs that allow users to express preferences about communications—preferences that are often viewed by users playing a central role in their personal and professional lives.

We first provide background on the Enhanced Telephony (ET) system, a platform that integrates client-side software with an organizational PBX and discuss the overlay of a call-handling agent named Bestcom-ET on top of the ET platform. Then, we review research on interruptability that has influenced the design and operation of Bestcom-ET. We review designs for tools aimed at acquiring call-handling preferences from users in an efficient manner. Finally, we review the results from surveys and from the analysis of the preferences and logs of Bestcom-ET users that took part in a study.

2. ENHANCED TELEPHONY PLATFORM
To investigate the opportunity of providing an integrated PC-Phone user experience for office workers, we have worked to develop a platform named Enhanced Telephony (ET). To date, over 7,300 employees (more than a quarter of the people on our company's main campus) have installed ET, and ET was being used by more than 4,300 people every day. Details about the basic platform and its core features including dialing, transferring, and placing conferencing calls from a desktop computer are described in [1]. We dwell in this paper on Bestcom-ET, a call-handling system with components that have been integrated into the ET client software and overall communications infrastructure.

Figure 1 shows a high-level overview the ET system and overlaid Bestcom-ET communications infrastructure. The ET system works with the Intecom PBX, the telephone system used by our organization. Intecom provides an application interface toolkit that allows external software to control individual telephones and to receive notifications about events for specific phones. ET client software communicates with a server which in turn communicates with the PBX. The PBX then sends the appropriate signals to the phone to execute commands. Beyond use of ET on an office-based desktop system, common scenarios include running the ET clients on a laptop with a wireless connection to the corporate

network and on a home computer that is connected to the corporate network.

We undertook an effort to develop an automated call-handling communications system, named Bestcom-ET, in parallel with the development of the basic ET platform. The Bestcom and ET teams worked in a tight collaboration on the integration and refinement of Bestcom-ET client-side and infrastructural components. A Bestcom-ET environment was created within the ET client and Bestcom components were added to the infrastructure, including Bestcom servers that store and communicate a user's preferences about call handling with multiple ET clients and with the ET server. The Bestcom server accesses events emitted directly from client software, and from the PBX via the ET server, and sends information to clients and the ET server for mediating call-handling actions such as the forwarding and rescheduling of calls based on situations and calendars.

The Bestcom-ET client components run on multiple personal computers. In one prototypical usage configuration, users execute the client on their main desktop system and also on a laptop computer. The system automatically senses which computer is currently being used by a user and uses the state sensing information from the active computer. Thus, when a user is active on their laptop computer in a mobile, networked setting, the laptop Bestcom client code and state sensing is relied upon for call handling operations.

3. COST-BENEFIT CALL HANDLING

Bestcom-ET provides to users several different levels of call-handling features, involving different amounts of user configuration effort. At the highest level, users can choose whether to specify call handling policies by making quick statements in a *Basic Bestcom* mode, or they can enter an *Advanced Bestcom* environment that contains additional tools for building a personal call-handling agent. The different modes are displayed as a top-level set of choices when entering the Bestcom Preferences environment, as displayed in Figure 2.

In the basic mode, users can made statements such as *Forward calls to my mobile phone or directly to my voicemail when my Instant Messenger is set to Busy and/or Away, my main office system is locked,* or *when my screen saver is running.* Users seeking more sophisticated control of call handling can invoke the consideration of groups of callers, mult-number forwarded by

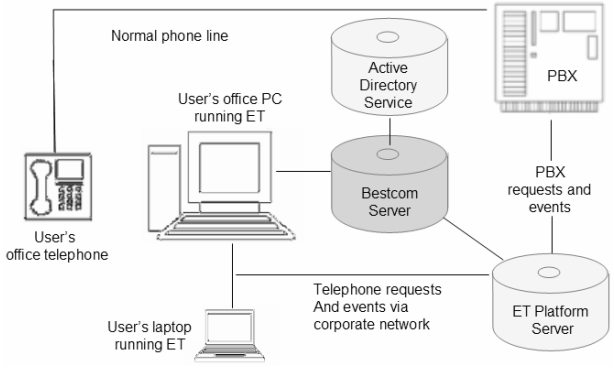

Figure 5: A high-level overview of the implementation architecture for the ET, showing the overlaid Bestcom service. The field deployment consisted of four PBXs and four Windows servers.

time, and a more expressive cost-benefit approach to call handling, by selecting *Advanced Bestcom* settings. We shall focus on the cost-benefit features and usage in this paper.

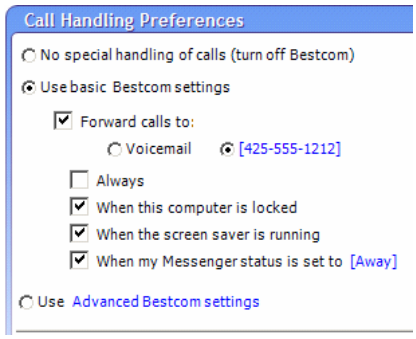

Figure 2. Screen for specifying basic policies and for entering the more advanced call-handling options.

3.1 Vision of Cost-Benefit Call Handling

Our long-term vision of building a call-handling agent has centered on developing methods that consider the costs and benefits of incoming communications. We have explored real-world incarnations of cost-benefit communications in several projects. In a research prototype, named Bestcom-X (for Bestcom-Experimental), fielded to a small group of people on our team, we are exploring the use of formal decision analysis to handle telephony. Bestcom-X employs personalized Bayesian models, learned via an intensive process of training via labeling data, to compute the expected cost of interruption for users. Such models take as inputs calendar information, the real-time monitoring of desktop events, and information gathered via acoustical and vision-based sensing. In Bestcom-X, inferences about a user's current interruptability and predictions about when a user will be available are used in decisions about relaying an incoming call to users, taking a message, or deciding if and when to reschedule a call.

For wider-scale fielding of a call-handling agent within the Extended Telephony platform, we developed a qualitative cost-benefit approach that harnesses key ideas from formal decision analysis, while bypassing the use of detailed inferential models. The approach centers on the direct assessment of policies about call priority and cost of interruption, rather than on the use of probabilistic inference. Nevertheless, the system's design leverages results gleaned about the cost of interruption via machine-learning analyses. We shall pause to examine these results.

3.2 Prior Studies of Interruptability

Interest has been growing over the last several years on methods for endowing computing systems with an understanding of users' interruptability in different settings. Researchers have investigated the cost of interrupting people in various ways in different situations [2,4,6], and have probed the cost of interruptions to people in office settings [3,5]. A study of the cost of interruption from case libraries of meetings and associated properties mined from calendars demonstrated that the interruptability of users is sensitive to several appointment properties [4]. Properties found to be useful for discriminating between low and high cost of interruption include the number of attendees at meetings, meeting location, relationship between the organizer and the user, and duration of meetings. In a set of studies on the use of machine

learning to build and field online models that can predict a user's interruptability [3], investigators found that the cost of interruption in office settings is sensitive to whether or not conversation is detected in an office, whether a user is currently interacting with the computer, whether the user is typing, and the software application that is active and in focus. In related work, a Wizard of Oz study of interruptability [5] demonstrated the importance of noting whether a user is speaking, writing, sitting, or interacting with objects such as a keyboard or phone, the presence and activities of occupants, whether the user's office door is open or closed, and positions and configurations of people.

3.3. Cost-Benefit Approach in Bestcom-ET

Rather than insist on deploying inferential models directly in Bestcom-ET, we have explored the value to users of tools that allow them to make explicit statements about costs and benefits of different call-handling actions. We allow users to make assertions about their interruptability based on observations about their context.

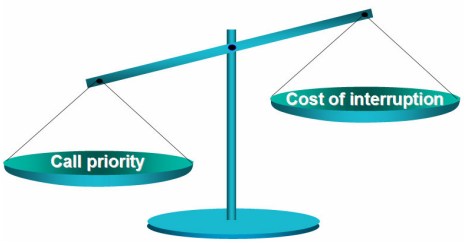

Figure 2. Perspective of cost-benefit call handling, centering on weighing a call priority with current cost of interruption.

Figures 2 and 3, extracted from first section of the Bestcom-ET user manual, are meant to familiarize users with the overall approach to building a cost-benefit call-handling agent. Figure 3 shows a schematic view of Bestcom-ET decision making, showing the changing cost of interruption with accepting a real-time telephone over time, and different people having different call priorities, representing the value of taking a call in real-time or, equivalently, the cost of deferring a communication until a later time.

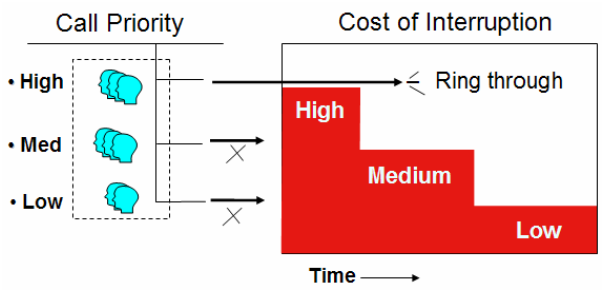

Figure3. Graphic from Bestcom user manual, as part of a discussion attempting to familiarize users with a qualitative cost-benefit perspective on call handling.

With the cost-benefit approach, users assign values of low, medium, and high priority to callers based on identity or situation. They are asked to consider their own definitions of low, medium, and high as a standard currency for value and cost, when assigning low, medium, and high costs of interruption for different contexts. In operation, the priority of a caller must be at least as high as the current cost of interruption for the caller to break through to the user. If not, the call is either shunted to voicemail or rescheduled depending on the privileges assigned by the call recipient to the caller. Thus, a caller needs to be in a group assigned a high priority to break through to users in a state of high cost of interruption, and a medium priority to break through to a user in a medium state of interruption, and so on.

3.4 Constructing a Communications Agent

The Bestcom-ET user manual steps users through phases of setting up a cost-benefit agent. We shall now review the key phases of the construction of a personal call-handling agent.

Static and Computed Caller Groups. Bestcom-ET provides tools that allow users to abstract callers into groups of callers and to express call-handling policies in terms of these groups. Groups include custom-tailored *static groups* and dynamically assigned *computed groups* that assign call-handling properties to users

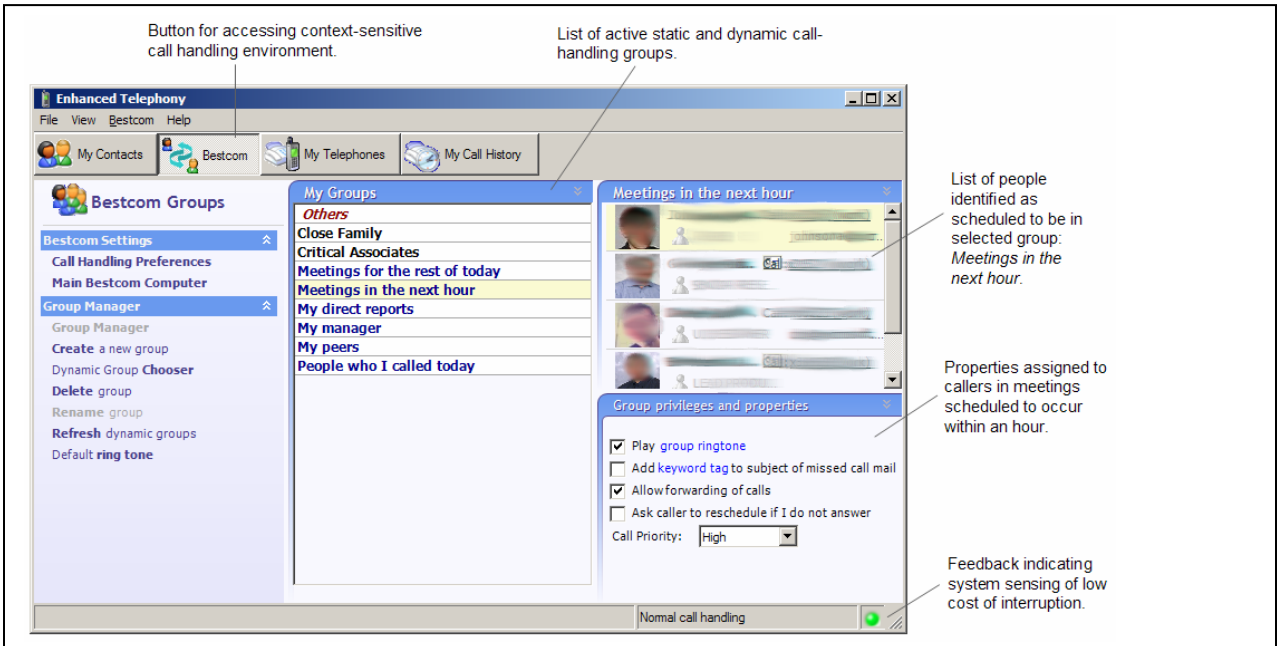

Figure 4. View of ET client in Bestcom-ET Group Manager mode, showing groups and associated properties under consideration (names blurred for anonymity).

based on based on relationships and situations. Users can assign properties to the groups, including *caller priority, forwarding privileges, rescheduling preference,* and group *ring tones*.

Dynamic Groups. Static groups are created by users and then populated from a user's predefined contacts or from an online directory. For dynamic groups, a palette of predefined computed groups is made available to users. Dynamic groups include several classes of computed sets of callers that serve to map the callers into groups depending on relationships and contextual information. Categories of groups include *calendar-centric*, *relationship-centric*, *communications-centric*, and *project-centric* groups. Calendar-centric groups include groups that are computed from the user's online meetings encoded in MS Outlook. These include callers in such groups as, *my next meeting*, *meeting in the next hour, meetings today,* and *meetings for the rest of the week*. Relationship centric groups include *my direct reports, my organizational peers, my manager, my manager* and *manager's manager*. Communications-centric groups include, *people who I called today* and *people who called me today*. Project centric groups, include *people who I've co-authored a document with this week, people who have assigned bugs to me,* and *people on my active projects list.* Project-centric groups were displayed in a grayed out mode as they were not yet implemented at the time of the study.

The Group Manager view of the Bestcom-ET environment within the ET client is displayed in Figure 4 (names and images have been blurred for anonymity). For static or dynamic groups, clicking on the group name in the Bestcom-ET group manager reveals the members of the group, with contact information, online status if available, and a picture if the group member had entered an image into the system.

Assigning privileges to groups. The Group Manager allows users to assign privileges to members of groups by entering options in a group privileges and properties region, displayed at the lower right-hand quadrant of Figure 4. Users can check boxes that grant group members forwarding and rescheduling privileges that allows them to be considered for forwarding or rescheduling, respectively, based on an analysis of the user's context. User's can also assign members of groups a call priority, including *breakthrough, high, medium,* and *low priority.* Breakthrough privileges allow the caller to be routed through to the user regardless of the user's context.

Default Interruptability Palette. After defining and activating caller groups and assessing the priorities of callers, users are asked to optionally assess their background or *default interruptability* for a typical week. The default interruptability represents the cost of taking phone calls at different times of day and days of the week in situations where there is no further statement about context. Users assert their background cost of interruption via a *time-pattern palette*., displayed in Figure 5. The palette allows users to sweep out regions of low, medium, and high cost of interruption over a seven-day period. Users can also indicate which periods of time should be set to block calls. At these times, only users assigned *breakthrough* privileges can get through to the user. Users are instructed that they can bypass this palette, thus informing Bestcom-ET to assume a background *low* cost of interruption for all times.

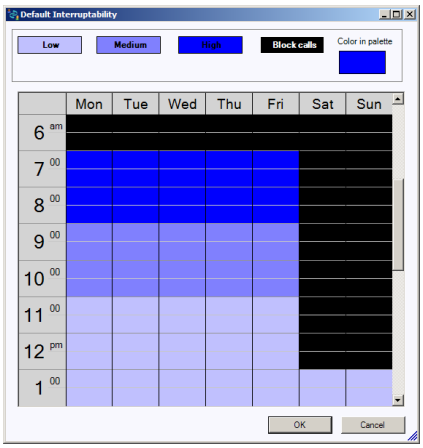

Figure 5. Time-pattern palette for specifying background costs of interruption by time, showing regions marked as low, medium, high cost of interruption (increasing shades of blue) and times when calls should be blocked (black).

Special Contexts. Next users are asked to specify sets of events that define context-sensitive changes in their interruptability. Figure 6 shows the *special context* assessment tool. Users can select and instantiate the values of observed events for desktop activity, calendar information, and sensors. Users select and instantiate events and then drag the events into low, medium, or high cost of interruption. Users are informed that the system will consider all events that they have specified and select the highest cost of interruption possible for all observed events.

Activity-centric events include any desktop activity, typing, using one or more applications, and instant messenger presence status of one or more of *busy, away,* or *online*. *Calendar events* include any meeting currently in progress, meeting duration, location, organizer, subject, attendees, and number of attendees. For attendees and organizer, users can specify lists of individuals as well as predefined abstractions including direct reports, peers, manager, and manager's manager. Currently, for the *sensors* class of events, users can instruct the system to consider them to be in a state of high, medium, or low cost of interruption when a conversation is detected in their office.

Activity-centric events are evaluated by a system event-monitoring component. Calendar events are gleaned from the MS Outlook application, using a periodic caching procedure to minimize computational effort. Conversation is detected with a module developed at our laboratory that detects acoustical energy in the audio spectrum in the human-voice range. The component was built to distinguish sound live conversation versus voices coming from speakers, e.g., a broadcast from a radio. Bestcom-ET provides a separate audio configuration, accessed when users first activate this event while assessing the special contexts. A microphone is required take advantage of the conversation-detection capability.

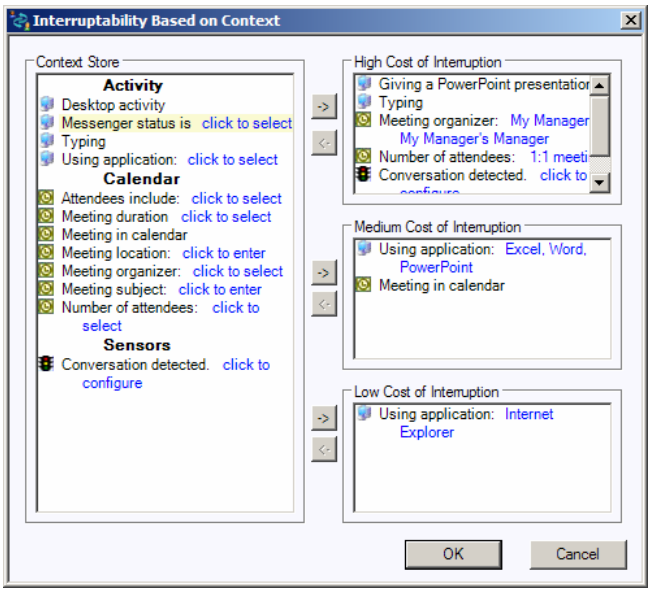

Figure 6. Assessment tool allowing users to specify sets of monitored events to define low, medium, and high costs of interruption as a function of activity and meetings.

In a next step, users are asked to review and refine a summary form that lists the groups that they have enlisted or defined, so as to examine or revise the priority of call groups, whether a group should be granted the privilege to forward calls to other numbers should the priority of a call exceed the cost of the interruption, and whether a caller should be rescheduled or simply shunted to voicemail in the case that the call does not exceed the current cost of interruption. This summary form is displayed in Figure 7.

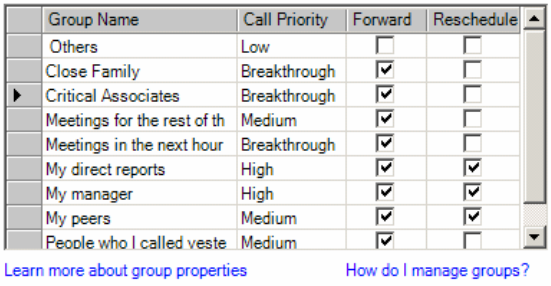

Figure 7. Tool for viewing and refining caller-group—specific privileges and preferences.

Users can specify preferred numbers for forwarding calls that come to their number at the main office PBX at times when they are away from the office phone—should the benefits of the call exceed the costs of the interruption. Multiple numbers are specified on a time-pattern palette for forwarding numbers, displayed in Figure 8. This palette employs a metaphor similar to the seven-day time palette used for assessing default costs of interruption. The forwarding palette allows users to specify times of day and days of week when, for example, a mobile phone should be used versus another office phone, or home phone.

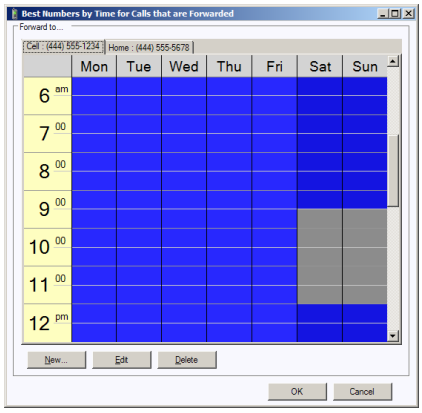

Figure 8. Time-pattern palette of week view showing specification for best forwarding numbers.

3.5 Rescheduling Assistant

If a caller is a member of a group assigned the rescheduling preference, and the priority of their call to the recipient does not outweigh the current cost of interruption for that user, the caller may be actively engaged by an automated rescheduling service. The system checks before engaging the caller if there is time available for the conversation within the tolerances for delay that the user has specified in a rescheduling preference profile. If calls are placed with ET, callers see a pop-up on their computer that indicates that the person they wish to speak with is not currently available and wishes to reschedule the call. By giving the rescheduling assistant a go ahead, the service checks the calendars of both the caller and recipient and recommends a list of potential times for their telephone conversation. Figure 9 displays the pop-ups used to seek convergence on a scheduled time for the call. By clicking on *Details*, the caller can include background information about the reason for the call and specify links to material that will be useful for the conversation. At the completion of the interaction, a tentative appointment is placed on the caller's calendar and an call-appointment invitation is mailed to the recipient.

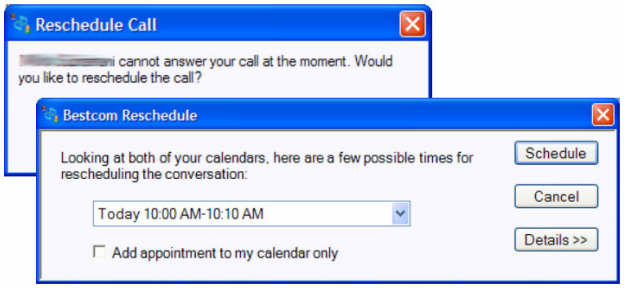

Figure 9. Experience for a user being engaged by the automated call rescheduling action.

Users can custom-tailor the behavior of their scheduling assistant, by asserting minimal times to wait so as to ensure that they will hear about the appointment. They can also assert a maximal delay time, and ask the system to bypass rescheduling if the first available slot will be after the maximal delay time. For example, a user can assert that attempts should be made to reschedule only if an appointment for the call can be made on the same day as the call. In this case, if a call cannot be rescheduled on the same day, the call may be directed immediately to voicemail, bypassing the invocation of the rescheduling assistant, or directly through to the user as a real-time call.

3.6 Assumed Contextual State Indicator

In the current system, information about the interruptability of users is not shared with others. This information is used by the communications agent for making privately-held call-handling decisions. Bestcom-ET shares its assumptions about the current interruptability of the users it is serving by displaying to them an interruptability-status graphic in both the lower right-hand corner of the client application and in the Windows system tray at the bottom of the display. Graphical renditions of a small LED and glowing cell-phone icon, displayed in Figure 10, glow green, yellow, or red to indicate whether the system believes that the user is in a low, medium, or high cost of interruption respectively. Hovering on the state reveals the rationale for the system's assumption.

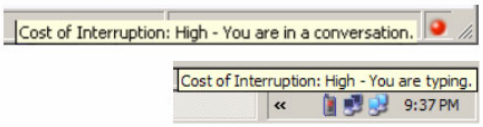

Figure 10. Bestcom-ET feedback about it's assumption about the user's current interruptability, at the bottom of the ET application (above) and in the Windows System Tray (below).

5. FIELD STUDY

To gather data to help answer our major research questions, we conducted a field study of a subset of Bestcom-ET users. Bestcom call-routing features were introduced in several phases. In March 2003, several Bestcom-ET features were made available to users, including the *Group Manager* with time-based forwarding of numbers, forwarding privileges, and ring tones. On September 10, 2003, we introduced the Bestcom-ET cost-benefit analysis. ET users were asked to try out the new features. Users were solicited to participate in a special study of the advanced features, including the filling out of surveys and the building of a cost-benefit call-handling agent. Participants were promised a coupon for a latte and to be included in a drawing for a dinner for two at a well-known local restaurant. Users were provided with a step-by-step guide to setting up their agents, and were given a survey before and after exploring the cost-benefit functionality.

5.1 Data Collection Methodology

We employed two methods to collect data about the usage of Bestcom-ET features. For the first, we instrumented the Bestcom-ET server and ET prototype such that all preference settings and major events (placing a call, receiving an incoming call, transferring a call, etc.) were logged to a database while ET was running. In addition, two separate surveys were conducted of the advanced users.

5.2 Results & Discussion

At the time of the survey, in the middle of October, 396 users were using the basic Bestcom policies, while 519 people were employing more advanced call-handling functionality. Of this group, 77 people were using the basic policies with time-based forwarding of numbers, 286 people were additionally using the group manager, and 157 people were using the cost-benefit analysis. At the time of the writing of this article, 177 people at our organization are using simple call routing, 544 people are using the group-centric controls, and 178 people are using the full cost-benefit analysis.

We shall focus here on the findings for the 157 users who were employing the cost-benefit analysis at the time of the study and the reports of the 87 users completing both surveys.

Use of Dynamic Groups. The proportion of users using different dynamic groups is displayed in Table 1, showing the popularity of relationship-based groups.

Default Palette and Special Contexts. We found that 99 (0.63) of the users assessed the default interruptability palette but that all 157 users assessed special contexts to define cost of interruptability. Table 2 shows the frequency of use of events to describe special contexts as being associated with *high* and *low* costs of interruption. We collected statistics about the particular values selected for each of the variables in the different contexts. The instantiations varied significantly for different costs of interruption. For example, *Messenger Status=Busy* was dominant for *high* cost of interruption but *Messenger Status=Online* was the most popular for the low cost contexts. Meeting property instantiations also showed high variance, with *high cost* associated with more use of attendee relationships, specific locations, and subjects than instantiations of meetings for the low cost state.

Table 1. Use of dynamic groups by users (percent).

My peers	94.3
My manager	89.8
My direct reports	87.3
Everyone above me	18.5
Meetings in the next hour	15.3
People who I called yesterday or today	14.0
Meetings for the rest of today	12.7
My manager and his manager	12.7
Meetings today or tomorrow	12.1
Meetings this week	11.5
People who I called today	11.5
People who called me yesterday	9.6
Meetings for the rest of the week	9.6
My manager and manager's manager	7.6
Everyone below me	5.1
Meetings yesterday	3.8
People who I called yesterday	2.6

Distribution of Interruptability. To better understand the distribution of states of low, medium, and high cost of interruption over work days for users, we analyzed the settings and logs of users for three sequential business days between 10am and 4pm. Based on users' activities and their specifications of definitions of high, low, and medium costs of interruption, we found that users spent an average of 0.67, 0.21, and 0.12 of their time in a states of low, medium, and high costs of interruption, respectively.

Effort to Create Agent. Figure 11 shows the amount of time users believed they had spent creating groups and priorities, and assessing the default and special contexts to define their cost of interruption. Most users reported that the construction of each of these tasks required 5 to 15 minutes.

Results of Usage Survey. In a survey, we sought to understand the value of different features and general feelings about the specific features in offered by Bestcom-ET. Table 3 captures user feedback on the survey inquiring about the value they found in various features, highlighting the importance of features to the users. Table 4 probes user agreement versus disagreement with several statements about the Bestcom-ET system and automated call handling. We found that users would be interested in predefined profiles and in automated tools for filling in their profiles. We were happy to see users showing overall agreement about the sufficiency of the current Bestcom-ET system as a call handling agent. The feedback underscored the importance of providing users with tools that can provide them with a better understanding of the operation of the cost-benefit decision making in summary and in context of particular situations.

Table 2. Percent use of observations to define contexts with high and low cost of interruption (instantiations not shown).

High cost context		Low cost context	
UsingPowerPoint	55.0	DesktopActive	34.4
MessengerStatus	41.9	Typing	26.9
MeetingPersons	25.0	UsingApplication	19.4
MeetingNow	13.8	MessengerStatus	16.3
MeetingNumPersons	11.9	MeetingPersons	5.0
InConversation	11.3	MeetingSubject	2.5
MeetingOrganizer	10.6	MeetingNumPersons	2.5
MeetingInLocation	9.4	InConversation	1.9
UsingApplication	8.8	MeetingDuration	1.9
Typing	5.0	MeetingNow	1.3
MeetingDuration	3.1	UsingPowerPoint	1.3
MeetingSubject	2.5	MeetingLocation	0.6
DesktopActive	1.9		

Stream of Feedback. Users shared informal feedback through email lists as well as in comments about the features sought through the survey. As a sample message about the excitement in using the system, one user stated to the Bestcom email feedback list shortly after setting up his person communications agent:

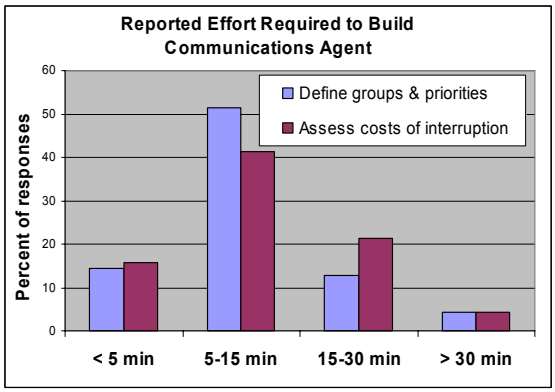

Figure 11. Responses to survey about effort required to assess groups and priorities and to assess costs of interruption.

> Subject: Dynamic Interruptability
>
> This is about the coolest new feature I've seen you put in. Very good stuff, and I'm already finding it useful—I've used this feature for two days now and have already had phone calls sent to VM during times where I was really trying to focus. Now if you can only do this for my door! ;-).

A number of comments centered on the value of having access to additional tools to provide a crisp overview of what would happen in different settings, based on users' assessment of groups, priorities, and costs of interruption. For example, one user stated in a comments on the survey:

> It would be very helpful to learn what will happen when you set a Bestcom feature. ...for example, setting a peer to medium vs. high priority. ...If I could know immediately what my changes affect, it would be beneficial.

We have carefully examined suggestions, kudos, and criticisms gathered during the study and are working to refine the system. As an example, we shall describe in Section 6 how we developed new summarization tools to enhance the transparency of the system.

5.3 Logs of Cost-Benefit Call Handling

We have been logging the cost-benefit call handling. At the time of writing this article, 2236 calls have been handled with the cost-benefit routing. Of these calls, 1705 have been routed to users for real-time conversations and 531 calls have been routed to voicemail. Tables 5 and 6 show the breakdown of the call priority encoded for the incoming caller and users' states of interruptability at the time of the incoming calls.

Table 3: Ratings of Bestcom-ET features by people who felt sufficiently familiar with the feature (*n*). Rated on scale: 1: not useful, 2: useful, 3: extremely useful.

Feature	*n*	avg.	s.d.	med.
Caller groups: User-created	71	2.31	0.52	2
Caller groups: Activity-based	55	2.11	0.65	2
Caller groups: Relationships	75	2.2	0.61	2
Default interruptability palette	61	2	0.68	2
Special contexts (in general)	71	2.27	0.58	2
Activity related	72	2.08	0.64	2
Messenger status	77	2.21	0.67	2
Appointment status	68	2.46	0.6	3
Conversation detection	52	2.1	0.69	2
Rescheduling assistant	29	1.97	0.67	2
Summaries of call handling relayed in email	66	2.5	0.53	3
Feedback on busy status	68	2.06	0.51	2
Time-based forwarding	65	2.32	0.64	2
Basic call handling	74	2.53	0.55	3
Identity-sensitive call handling	69	2.55	0.58	3
Cost-benefit routing	54	2.31	0.57	2
Computer-generated ring tones	44	1.84	0.74	2

Table 4: Reactions to statements about Bestcom-ET and on automated call handling. Rated on scale: 1: strongly disagree, 2: disagree, 3: neutral, 4:agree, 5: strongly agree.

Statement	avg.	s.d.	med.
Control provided by the current version of Bestcom-ET is sufficient for handling my calls.	3.87	0.86	4
The special contexts provided are sufficient.	3.54	0.84	4
The group prioritization control is rich enough.	3.55	0.75	4
Users should expect setting up call preferences will require focused attention.	3.99	1.09	4
Automated tools to help me fill out preference profiles would be useful.	4.16	0.76	4
I would like clear summaries of what will happen when different people call.	4.51	0.64	5
I would prefer a system that learned my preferences simply by watching.	3.62	1.06	4
It would be valuable to choose among pre-built profiles and modify them.	4.08	0.87	4
Context-sensitive call handling can provide great value to people.	4.37	0.61	4
Call-handling features as they exist in the current ET system provide value to me.	4.11	0.72	4

6. TOOLS AND VISUALIZATIONS FOR UNDERSTANDING POLICIES

We have taken particular note of the findings indicating a need for more clarity about behavior of policies and policy changes, added emphasis to our efforts to enhance clarity for users. In this vein, we have worked to develop and field new inspection tools, including means for viewing typical scenarios, summaries showing the handling of calls by the agent in different interruptability states, and the revelation, with a simple gesture, indicating a wish to review "What will happen now?" or "Who can reach me now?" Figure 12 shows an example of the "What will happen now?" summary, that can be accessed with ease by users via clicking on the interruptability status indicator.

Table 5: Context and call priority for calls routed through to users by cost-benefit policies (*N*=1705).

Busy State	Call Priority	Num. calls
Low	Breakthrough	174
Low	Low priority	193
Low	Med priority	837
Low	High priority	56
Med	Breakthrough	53
Med	Med priority	301
Med	Med priority	32
High	Breakthrough	38
High	High priority	17
Block	Breakthrough	4

Table 6: Context and call priority for calls routed to users that were shunted to voicemail by cost-benefit considerations (*N*=531).

Busy State	Call Priority	Num. calls
Low	Block group	18
Med	Block Group	2
Med	Low priority	98
High	Block Group	5
High	Low priority	48
High	Med priority	351
Block	Low priority	2
Block	Med priority	7

The popup displays a list of groups and associated people and the actions that will occur if people contact the users by telephone now. Users can make changes in the policies directly from links provided in this view. The changes are propagated to priorities of callers as groups or as exceptions within groups. The creation of exceptions can be used in spawning new groups of users with special context-sensitive priorities.

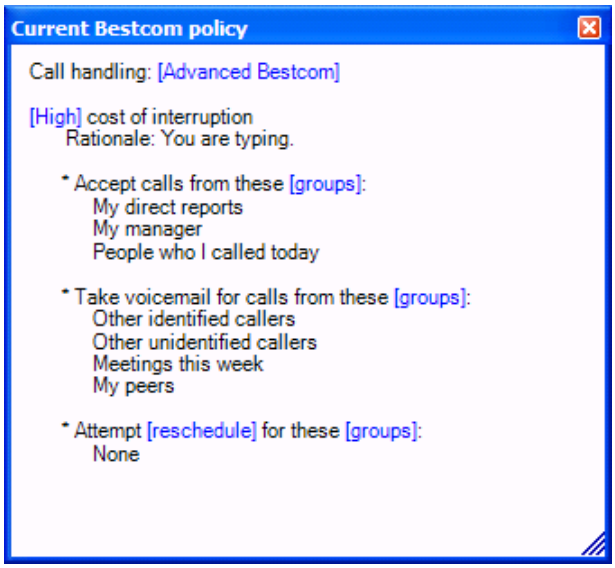

Figure 12. "What will happen now?" display showing current cost-benefit call handling policies.

In other approaches to enhancing the clarity about the operation of the system for users, we now include in an activity log a detailed trace of what happened with each call and why. Missed call notifications include the details for the call at hand. For example, the log may state for a call: "A call from Steve Smith was sent to you because you were scheduled to meet with Steve within an hour, and you were not active on your desktop system. You indicated that if you were away and if it was between noon and 4pm to forward such calls to your cell phone number," or, "The call from Sally Jones was rescheduled until 3pm today as your were speaking with someone when the call came in, a context that you require that callers have a high priority for breaking through—and Sally is in a medium priority group." Users are given controls to make changes directly from links highlighted in these traces.

We shall be studying in future weeks whether these new tools offer the kind of clarity several participants in our survey indicated that they would wish to have in setting up, understanding, and maintaining their call-handling policies.

7. CONCLUSION AND FUTURE WORK

We described the Bestcom-ET prototype, focusing on preference assessment tools for creating a cost-benefit call handler with the ability to balance the cost of interruptions with the assessed priority of callers based on identity, activities, and relationships. We reported several key results derived from online stores of preferences and logs of activity shared out by users as part of ongoing research. We also reported on the statistics of call handling relaying how calls were handled based on the contextual states and caller priorities assigned by users.

We found that users can assess and build an agent that they consider is expressive enough for their needs in a relatively short amount of time, and that a cost-benefit metaphor was understandable and useful.

We have pursued enhancements that provide greater understanding of the agent's decisions, make available new approaches to specifying the value of communications, and integrate additional events for specifying a user's context. We will also continue to explore the use of explicit machine learning, both to assist users with building profiles and in mediating call handling.

We are intrigued with the prospect of providing users with tools for constructing real-time communications agents that can mediate the *if, when, and how* of communications. Our initial investigations make us optimistic that such tools can provide users with valuable automation in the handling of real-time communications.

7. ACKNOWLEDGMENTS

We thank the users who participated in the Bestcom-ET user study, and Bestcom-ET users overall for allowing us to log their interactions with the call-mediation system.

8. REFERENCES

1. Cadiz, JJ., Narin, A., Jancke, G., Gupta, A., Boyle, M. Exploring PC-Telephone Convergence with the Enhanced Telephony Prototype, *Proceedings of the 2004 Conference on Human Factors in Computing Systems* (CHI '04), Vienna, April, 2004.

2. Gillie, T. and Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity and complexity *Psychological Research*, 50, 243-250.

3. Horvitz, E. and Apacible, J, A. (2003). Learning and Reasoning about Interruptions, *Proceedings of ACM ICMI 2003,* Nov. 2003.

4. Horvitz, E. Koch, P., Kadie, C.M. Jacobs, A. (2002). Coordinate: Probabilistic Forecasting of Presence and Availability. *Proceedings of UAI 2002*, pp. 224-233.

5. Hudson, S.E., J. Fogarty, C.G. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J.C. Lee, J. Yang (2003). Predicting Human Interruptability with Sensors: A Wizard of Oz Feasibility Study, *Proceedings of CHI 2003*.

6. McCrickard, D.S., Catrambone, R., Chewar, C. M., and Stasko, J.T. (2003). Establishing Tradeoffs that Leverage Attention for Utility: Empirically Evaluating Information Display in Notification Systems. *International Journal of Human-Computer Studies*.

7. Sawhney, N. and Schmandt, C. *Nomadic Radio: Scaleable and Contextual Notification for Wearable Audio Messaging*. In Proceedings of the 1999 Conference on Human Factors in Computing Systems (CHI '99), Pittsburgh, PA, May 15-20, 1999.