

# Bayesphone: Precomputation of Context-Sensitive Policies for Inquiry and Action in Mobile Devices

Eric Horvitz, Paul Koch, Raman Sarin, Johnson Apacible,  
and Muru Subramani

Microsoft Research, One Microsoft Way Redmond,  
Washington 98052 USA

{Horvitz, Paulkoch, Ramans, Johnsona, Murus}@microsoft.com

**Abstract.** Inference and decision making with probabilistic user models may be infeasible on portable devices such as cell phones. We highlight the opportunity for storing and using precomputed inferences about ideal actions for future situations, based on offline learning and reasoning with the user models. As a motivating example, we focus on the use precomputation of call-handling policies for cell phones. The methods hinge on the learning of Bayesian user models for predicting whether users will attend meetings on their calendar and the cost of being interrupted by incoming calls should a meeting be attended.

## 1 Introduction

Over the last decade, there has been increasing research on the use of probabilistic user modeling for inferring user goals and states of the world [6,7] under uncertainty. The user models have been applied typically in desktop settings, where designers can assume that a personal computer is available for performing inferences. We focus in this paper on the precomputation of ideal decision-theoretic policies from probabilistic user models and the caching of the policies on a cell phone for decision making in a mobile setting. We believe that such precomputation and caching of policies will enable probabilistic learning and reasoning to be applied to the large and growing number of devices and appliances in the world with limited computational abilities.

We focus on the example of using probabilistic models to guide the handling of telephone calls, so as to deliberate about the cost of interruption versus the cost of deferral of an incoming call. Such decisions can be made locally at cell phones, based on a consideration of context and multiple properties of meetings. We analyze the case of local decision making based on sensed properties of meetings on a user's calendar and on properties of callers based on caller identification, as well on real-time sensing of motion and ongoing conversation.

We first discuss the learning of predictive models of attendance and of interruptability. We present the computation of the expected cost of interruption from the output of these models. Then we discuss the computation of value of information to reason about the value of acquiring additional information from users in real time. We review how we can precompute and cache policies on cell phones that consider whether calls should interrupt users. We discuss how the methods have been used to field a prototype call handling system that we call Bayesphone.

## 2 Learning Models of Interruptability and Attendance

Efforts over the last several years have demonstrated that relatively accurate models can be constructed for predicting the interruptability of users from such contextually relevant observations as sensed activity and calendar properties [2,3,4,5,8]. We shall focus first on the construction of two Bayesian network models via supervised machine learning. One model predicts the interruptability of a user. More specifically, we build a model that is used to infer a probability distribution over the cost of interruption of users. The second model outputs the probability that users will attend meetings that appear on their electronic calendar. Inferences from both models are used to predict the expected cost of interruption at different times for a user. Additionally, we show how we compute from the output of the models, the value of information associated with asking users in real time about their situation. Such an analysis considers the inferences from the models as well as the frequency and types of calls coming in over time.

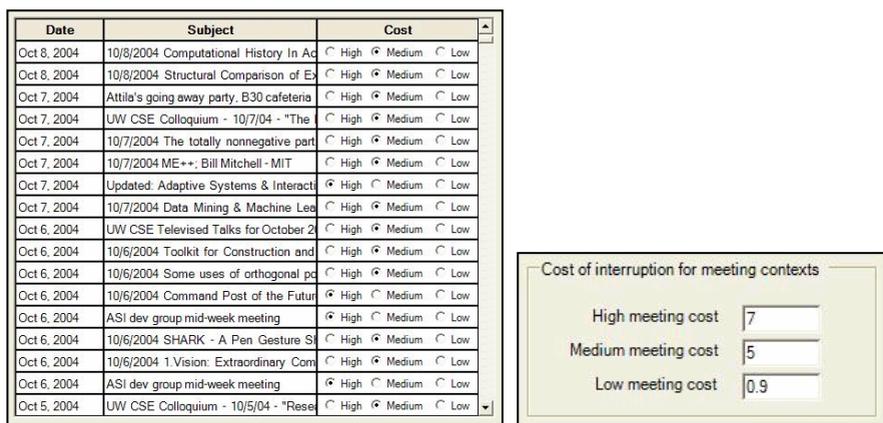
### 2.1 Models of a User's Interruptability

We have been investigating predictive models of the cost of interruption from evidence associated with a user's context, including a stream of sensed data generated by a user's interaction with a desktop computer and properties of items on a user's electronic calendar [4,5]. Online calendars are central for coordinating meetings in many enterprises. For example, the Outlook calendaring subsystem is used universally at our organization for extending invitations to meetings, monitoring responses about planned attendance, and scheduling and tracking daily agendas. As part of the Coordinate project, we constructed an appointment crawler and assessment tool that searches through users' online calendars, as represented in the Microsoft Outlook messaging and calendaring application [3]. The appointment crawler sifts through online appointments and records sets of properties associated with each appointment. The appointment crawler notes, for each appointment, a set of properties drawn from the Outlook application, including the time of day and day of week of the meeting, meeting duration, subject, location, organizer, the response status of the user (responded yes, responded as tentative, did not respond, or no response request was made), whether the meeting is recurrent or not, whether the time is marked as busy or free on the user's calendar, whether the user was required or optional, the number of invitees, the organizational relationships of the invitees to the user, and the role of the user (user was the organizer versus a required or optional invitee). The system accesses the Microsoft Active Directory service to identify the organizer of the meeting and invitees and notes whether the organizer and attendees are organizational peers, direct reports, managers, or managers of the user's manager.

The crawled data is used to build an assessment view that displays a form to users. The form consists of a list of titles of meetings and provides fields for indicating the state of interruptability of users. Fig. 1 shows the assessment palette for assigning a cost of interruption to each crawled calendar item and a form used to define the meaning of high, medium, and low cost of interruption states. Users use this form to assign scalar values to each state of interruptability. For this assessment, we ask users to estimate the cost associated with a ringing phone during states of high, medium,

and low cost of interruption. To ground the semantics of cost throughout the system, we consider the decision-analytic notion of *willingness to pay*; and assess dollar values that users would be willing to pay to avoid a call in each setting.

Given the interruptability tags and appointment properties, we build a library of cases, and then employ a Bayesian structure search procedure, based on methods developed by Chickering, et al. [1], to build a Bayesian network. The methods employ a greedy search across different structures to identify the probabilistic dependency structure that best explains the data, based on a score known as the Bayesian Information Criterion. The resulting Bayesian network can be used later to infer probability distributions over the states of interruptability for previously unseen meetings, based upon a consideration of a set of observations consisting of the properties of meetings.



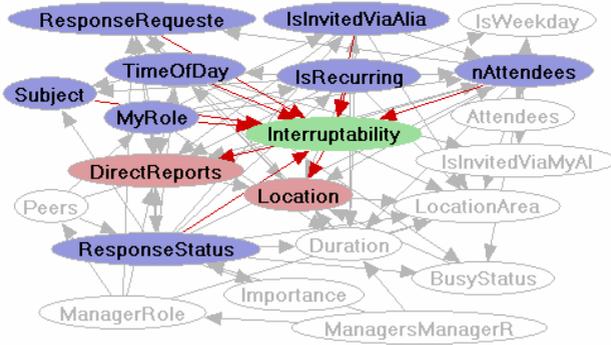
**Fig. 1.** Cost of interruption assessment palette that enables users to view a list of prior appointments and to assess the cost of a phone call during each meeting (left). Overall dollar-valued costs are assigned to each state (right)

Fig. 2 displays a Bayesian network model learned from a set of cases tagged by cost of interruption. The model can be used to infer a probability distribution over states of interruptability, outputting for each previously unforeseen appointment the likelihood that the meeting has a high, medium, or low cost of interruption. A study of a model constructed from the same 559 appointments and tested on 100 hold out cases showed a classification accuracy of 0.81 for assigning interruptability.

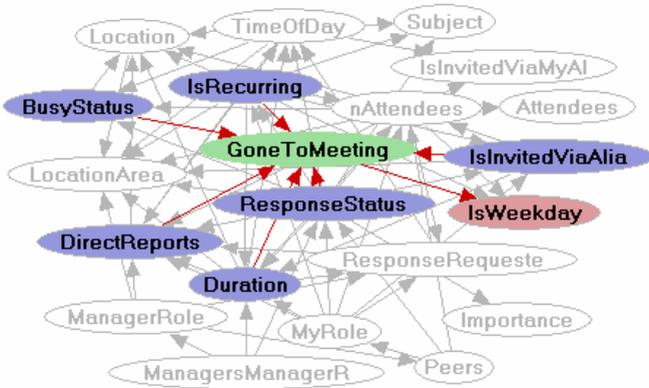
## 2.2 Models of Meeting Attendance

Beyond models of the cost of interruption associated with a context associated with the attendance of a meeting, we also assess and learn in an analogous manner Bayesian network models that predict the likelihood that the meetings will be attended, based on meeting properties. Fig. 3 shows a sample Bayesian network learned from training data for inferring the likelihood that a user would attend meetings, based on

meeting properties. The model was trained with the same appointments as were used to train the model for the cost of interruption. In use, the personalized attendance model generates, for previously untagged meetings, the likelihood that users will attend the meetings. For this model, a study of the accuracy on 100 cases held out for testing found that attendance was classified at an accuracy of 0.92.



**Fig. 2.** Bayesian network learned from case library that can be used to infer the probability distribution over states of a variable representing the interruptability of a user, given attendance of a meeting with particular properties. The most influencing variables and their probabilistic dependencies are highlighted with shading

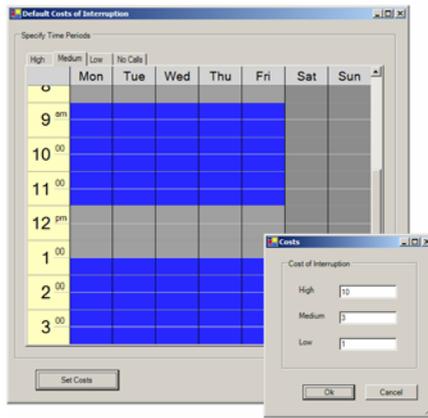


**Fig. 3.** Bayesian network learned from case library that can be used to infer whether a user will attend a meeting or not, based on meeting properties. The most influencing variables and their probabilistic dependencies are highlighted with shading

### 3 Computing Expected Cost of Interruption

We can employ the Bayesian networks for predicting attendance and the cost of interruption to compute the *expected cost of interruption* (ECI) associated with calls that ring through to users who are attending different kinds of meetings. To perform

the computation of expected cost of interruption, we consider the probability distribution over the cost associated with the meeting at hand, as provided by the interruptability model, and the likelihood that a user will attend the meeting indicated on the user’s calendar, as provided by the attendance model. To compute the ECI of interrupting a user when a meeting on a user’s calendar is recognized as being in progress, we need one additional piece of information—the *default cost* associated with receiving a phone call when a user does not attend a meeting indicated on a user’s calendar. Such a default cost is typically a function of the time of day, as receiving a call during the early hours of the morning or very late at night is likely to be different than receiving a call during business hours, and the cost of interruption may also be dependent on the day of week. To assess default costs of interruption, we allow users to sweep over default high, medium, and low cost regions within a seven day by twenty-four hour time palette, and to define default costs of interruption to each value. Fig. 4 displays the palette for assessing the default of interruption by time.



**Fig. 4.** Time palette for assessing costs of interruption by time via a sweeping out of regions of time, and assessing default costs for non-meeting times assigned high, medium, and low costs

Given (1) an inferred probability distribution over the interruptability of a meeting, (2) the likelihood that a user will attend a meeting on their calendar, and (3) the default cost associated with the no-meeting situation, the ECI at any moment is computed by weighting the cost of interruption for the no-attendance and attendance situations in accordance with the likelihoods of these states. Taking the expectation, the ECI is,

$$ECI = p(A|E) \sum_i p(c_i | E) c_i + (1 - p(A | E)) c^b(S) \tag{1}$$

where  $p(A|E)$  is the likelihood that users will attend a meeting, given evidential properties  $E$  associated with the meeting, obtained via Outlook appointment properties,  $p(c_i|E)$  is the probability that users will assign a cost  $c_i$  to the meeting, where  $i$  indexes the meeting as being either in low, medium, or high cost, and  $c^b(S)$  is the background cost of being interrupted in the default situation  $S$ , representing the

case where a user does not attend a meeting, as captured by the time of day and day of week. The default cost can be extended to be dependent on multiple aspects of a user's overall context  $S$ . Also, special mutually-exclusive contexts can be considered as active in a priority-order relationship. In the current Bayesphone prototype, the special contexts of *user driving (stop-and-go* versus *smooth highway driving)* and *local conversation in progress* are sensed from a Bluetooth-based GPS system and headset, respectively. If neither of these situations is sensed, the meeting and default day and time context is considered as active. Otherwise the costs of interruption assessed for the special contexts are assumed.

## 4 Performing Cost-Benefit Analysis in Real Time

We can balance a computed cost of interruption with the cost of deferring a conversation until later. A key piece of the decision is the cost of deferring calls from different callers. We thus obtain from users the dollar-value cost associated with delayed communication when a call is routed to voicemail rather than a real-time conversation. For such an assessment, we allow users to define groups of callers, based on properties of people, so as to provide a manageable set of classes. The deferral-cost assessment tool allows users to create groups of people based on sets of properties of people including organizational relationships and activities. The tool allows users to create such organization-related groups as peers, direct reports, manager, position higher-up in the organizational chart, person within organization, and people identified in a user's list of contacts. Users can also pick activity-based groups, so as to have their device recognize people who are scheduled in meetings in the next hour, on the same day, or later in the same week. Another activity-based group provided by the system is "people I have called today" and "people I have called this week." The tool can also be used to build ad hoc groups like "critical associates," and "close friends."

We employ Equation 1 to precompute the expected cost of interruption based on meeting properties for any time during the day. We shall return to the desktop and mobile device application in Section 5. First, we review the precomputation of value of information for making decisions about when to acquire additional information from users.

## 5 Precomputing Ideal Interactions with Users

Beyond storing policies for making the best decision based on information that is *currently available* to a system, we have extended the basic cost-benefit analysis with precomputation about whether it is worthwhile for the phone to ask users at run time to assist with resolving key uncertainties about the user's situation. More specifically, we precompute the *value of asking users* for information about whether they are attending a meeting that appears on their calendars. Answers to such queries can resolve key uncertainties used in the ECI computation, potentially increasing the value of the call-handling policies.

To identify ideal queries, we compute the expected value of information (EVI) of asking users a question. EVI is a decision-theoretic measure of the value of gathering

additional information that considers the current uncertainties, the likelihood of different answers to a query for more information, and the ultimate influence of the different answers on ideal policies. For the case of Bayesphone, we precompute the value of asking a user if they intend to attend a meeting before a meeting is scheduled to begin. The question itself incurs a cost of interruption that must be balanced with the gains in value based in the new information.

To compute the value of asking the user about attendance, we must consider the ECI before and after asking, and cost of querying the user. Given an answer, the ECI will be either be the expected cost associated with the meeting (the first term in Equation 1), or the background cost of the time of day (the second term Equation 1). To compute the value of information, we introduce the concept of the overall communication cost over a period of time. The *expected communication cost* (ECC) for a period of time is the cost of deferral and cost of interruption for all incoming calls during the period. We wish to interact with a user only if the reduction in ECC is greater than the cost of asking. Bayesphone precomputes the value of information for all meetings and uses this information to drive selective question asking.

The ECC is computed by maintaining a log of incoming calls. Bayesphone records a log of incoming calls by group. This log is segmented into calls that arrive at different time periods. For the current prototype, we consider eight periods: *mornings*, *afternoons*, *evenings*, and *late night* for weekdays and weekends. For each period, we compute the rates at which calls associated with different caller groups arrive each hour. Given this information, we can compute the ECC for any value of the cost of interruption. We simply note the expected number of calls that will be deferred and the calls that will ring through to a user given the computed ECI. The expected numbers of each class of calls are computed as a product of the stored rates for each caller group and the duration of the period. The ECC for a meeting of duration  $t$  based on a consideration of current evidence  $E$  only about properties of the upcoming meeting is,

$$ECC(E, t) = \left( \sum_i f_i c_i^{defer} + \sum_j f_j c_j^{ring} \right) t \quad (2)$$

where  $f_i$  is frequency of calls in each caller group  $i$  that has a cost of deferral lower than the cost of interruption,  $f_j$  is the frequency of calls in each caller group  $j$  that has a cost of deferral higher than the cost of interruption, and  $c^{defer}$  and  $c^{ring}$  are the costs of deferral and cost of interruption of each of these caller classes, respectively. We note that  $c^{ring}$  is just the current expected cost of interruption, ECI, as computed with Equation 1, so we can rewrite Equation 2 as,

$$ECC(E, t) = \left( \sum_i f_i c_i^{defer} + \sum_j f_j ECI(E) \right) t \quad (3)$$

To compute the EVI of asking the user a question, we recompute ECI and ECC separately for the answers of “attending” and “not attending,” identifying the changes in the numbers of calls in the deferral and the ring-through classes for the updated values of ECI, and finally combine these two ECC values together, weighted by the probability of hearing each answer. The communication cost for the answer, “attending meeting” considers the expected cost associated with being at the meeting,

$$ECC(E, a, t) = \left( \sum_i f_i' c_i^{defer} + \sum_j f_j' \sum_k p(c_k | E) c_k \right) t \quad (4)$$

The communication cost for the answer, “not attending” takes as the cost of interruption, the background cost associated with the time of day,

$$ECC(E, \bar{a}, t) = \left( \sum_i f_i'' c_i^{defer} + \sum_j f_j'' c^b \right) t \quad (5)$$

Putting these terms together, we can compute the expected value of asking the user as,

$$\begin{aligned} EVI(E, t) = ECC(E, t) \\ - [p(A|E)ECC(E, a, t) + (1 - p(A|E))ECC(E, \bar{a}, t)] \quad (6) \\ - C^a \end{aligned}$$

where  $C^a$  is the cost of asking the user before the meeting, just the ECI before the meeting begins. The system also considers the added value of directly asking a user about the interruptability of a meeting, given that the user has answered that the meeting will be attended, using an analogous value of information computation. Users are asked to optionally answer a second question about the cost of interruption, if acquiring that information is worth the incremental cost of asking the second question.

## 6 Bayesphone Desktop and Mobile Applications

Bayesphone consists of two applications: (1) a desktop application, running on WindowsXP, that performs inference, cost-benefit analyses, and value of information precomputation of ideal real-time actions and inquiries, and (2) an application running on Smartphones that downloads the precomputed policy file from the desktop via a device synchronization program. The Bayesphone desktop application analyzes each forthcoming meeting, making inferences with the Bayesian network models for both attendance and interruptability. The client application considers these inferences along with the costs of deferral of calls from callers in different groups, the expected cost of interruption with taking calls for each meeting, and the history of incoming calls in the user’s call log, and precomputes the ideal call-handling actions and interactions for each meeting. The desktop system creates an XML-encoded file, which includes for each meeting, the meeting title, date, and time, whether the user should be asked with an alert about meeting attendance before the meeting, and the list of caller groups who are allowed to breakthrough to the user during the meeting for the no-interaction or no-answer case, and for each answer.

In use, a user may be asked before a meeting occurs about whether they plan to attend the forthcoming meeting. A special alert tone is used to inform the user about the question, and a screen appears that allows the user to specify whether they will attend the meeting. The maximum likelihood answer is displayed on the device, allowing the user to either confirm or to change the guess. If no answer is available with three minutes, the question times out and the title of the meeting appears on the

screen, along with the groups who can breakthrough. Users can directly change their attendance status or the cost of interruption directly at any time via a menu, and the ideal precomputed policy for the state will be accessed and displayed. Figure 6 displays two screens of the Bayesphone application executing the call-handling policies of one of the authors. In this case, the system has alerted the user to the value of answering a question about attendance before a meeting. The system has guessed that the user will not attend the meeting, and the user confirms this guess. After the interaction, the system shows the users the caller groups that will be allowed to breakthrough. At run time, Bayesphone intercepts incoming calls and takes control of the ringing of the phone. The application checks caller ID, examines the list of callers allowed via the precomputed cost-benefit analysis, and decides whether to ring the phone versus transfer the call to voicemail.

Although the primary intent of this paper is to share with the User Modeling community methods for precomputing user models for fielding ideal policies on mobile devices that do not have the computational power of desktop machines, we are interested also in the value of these methods for the call-handling domain. The initial Bayesphone prototype has been used by two people on our team for four months. We have not yet performed a formal validation of satisfaction with call handling, but the system has been reported by both users to perform well overall in a qualitative survey. Both users provided us with feedback on the effort with setting up the system. The users found that the assessments of caller groups, costs of deferral, and costs of interruption were straightforward, taking under 15 minutes to complete. However, they found the assessment of the Bayesian models to be more burdensome, taking about two hours of time to assess crawled events from their online calendars. We are working on means for easing this burden via experience sampling along the lines of [5], and on the use of lighter-weight, but less precise models. Such an approach includes the reliance on the direct assessments of probability distributions for attendance and interruptability for classes of appointments.

## 7 Summary

We have described a project highlighting the opportunity for precomputing inferences from Bayesian networks and coupling these inferences with cost-benefit policies for fielding policies for action and dialog with users on simple end-point devices like cell phones. We reviewed the construction of probabilistic models that can infer the expected cost of interruption and the likelihood that users will attend meetings on their calendar. We showed how these models can drive a cost-benefit analysis of call-handling policies and reviewed a prototype application. We are now studying the difficulties that users may have in building probabilistic models for the prototype, and the overall experience with using the system. We are also working to extend the evidential considerations beyond meeting properties and time, to include such observations as local sensing of location, motion, and ambient acoustical signals, such as those representing a nearby conversation in progress.



**Fig. 5.** Bayesphone application, showing the case where it is best to ask the user about attendance of a forthcoming meeting. When the meeting starts, the application displays the title of the meeting in progress, the input from the user, and the callers who can break through

Moving beyond the motivating example we selected to explore the precomputation of personalized policies, we are excited about the prospects for precomputing user models for fielding adaptive behavior that can be executed on a variety of small devices—especially on mobile devices that may have minimal computational power.

## References

1. Chickering, D.M., Heckerman, D. and Meek, C. (1997). A Bayesian approach to learning Bayesian networks with local structure. In *Proc. of UAI 1997*, pp. 80-89.
2. Fogarty, J., Hudson, S.E., and Lai, J. (2004). Examining the Robustness of Sensor-Based Statistical Models of Human Interruptability, *Proc. of CHI 2004*.
3. Horvitz, E. Koch, P., Kadie, C.M. Jacobs, A. (2002). Coordinate: Probabilistic Forecasting of Presence and Availability. *Proc. of UAI 2002*, pp. 224-233.
4. Horvitz, E. and Apacible, J. (2003) Learning and Reasoning about Interruption, *Proc. of ICMI 2003*, pp. 20-27.
5. Horvitz, E. Apacible, J. and Koch, P. (2004) BusyBody: Creating and Fielding Personalized Models of the Cost of Interruption, *Proc. of CSCW 2004*.
6. Jameson, A. (1996) *User Modeling and User-Adapted Interaction, Volume 5*, pp. 193-251.
7. Jameson, A. (2003). Adaptive Interfaces and Agents In J. Jacko and A. Sears (Eds.), *Human-computer interaction handbook* (pp. 305-330). Erlbaum Publishers.
8. Mynatt, B. and Tullio, J. (2001). Inferring Calendar Event Attendance, *Proc. of Intelligent User Interfaces 2001*, pp. 121-128. ACM Press.