A Preliminary Study on Interruptibility Detection Based on Location and Calendar Information

Hermann Stern, Viktoria Pammer, Stefanie N. Lindstaedt

Know-Center GmbH
Inffeldgasse 21a
8010 Graz, Austria
{hstern, vpammer, slind}@know-center.at

Abstract. Modern communication technology - such as mobile phones - increases our level of availability, but also raises the risk of being inappropriately interrupted. In this paper, we present our on-going research on automatically detecting a user's interruptibility. This is the first step towards (i) contextualizing the modus of message notification, i.e. making the notification more or less attention-grabbing, and (ii) contextualizing message ranking, i.e., prioritizing messages according to their relevance for the user's current level of activity. We describe our approach of automatically detecting a user's interruptibility based on location and calendar information. Both kinds of data are easily available in a mobile setting using smartphones. Second, we present a preliminary study that evaluates (i) whether GPS information is available sufficiently for our purpose in a real-world setting, and (ii) whether the computed interruptibility corresponds to a user's own perception of interruptibility.

Keywords: Interruptibility, mobile devices, context, services, service recommendation.

1 Introduction

Modern communication technology brings many benefits such as ways for fast communication or high availability, but also raises the risk of being inappropriately interrupted. According to [1], 24% of cell phone users feel the need to answer a phone call when they are in a meeting. The most common way to ensure no interruption could happen is to turn the device off. However, users then are afraid of missing important calls or information [1]. We assume, in line with existing research (see also Section 3 on related work), that a user's interruptibility with regard to incoming messages on a mobile phone, i.e. the degree to which she will feel disturbed if interrupted by an incoming message, can be computed in a mobile setting based on location and calendar information. Assuming that we can correctly compute a user's interruptibility, this interruptibility is the first step towards contextualizing the modus of message notification, e.g., adaptively muting a mobile phone when recognizing that a user is currently busy. In addition, interruptibility could be used as one of two aspects in message ranking, where the second aspect would the content-wise relevance of an incoming message to a user's current activity.

2 Problem Statement: Interruption and Interruptibility

Interruptibility

Interruptions are “externally generated, randomly occurring, discrete events that break continuity of cognitive focus on a certain task” [2] or “disturbance[s] to the normal functioning of a process in the system” [3]. The common denominator of most definitions of interruption is that an external event is in some form drawing attention from the current primary task. While for humans social conventions indicate when it’s appropriate for a person to interrupt each other, computer systems are not aware of those social conventions and the impacts their actions may have [4]. We define interruptibility as “the current state of a user regarding her receptiveness to receive messages”, and focus in this work on receiving messages sent automatically by computer services. Low interruptibility means that the user is focused on her current task and does not want to be disturbed. High interruptibility on the other hand states that the user will not feel disturbed by incoming messages.

Motivation and Vision

Our overall goal in studying interruptibility is to prioritize incoming messages on mobile phones based on user context, and subsequently to adjust the modus of message notification. In our vision, we focus on computer-generated messages, such as ads, notifications concerning upcoming appointments, or updates on travel information. Incoming messages should be prioritized according to the user’s current level of activity and the relevance of the incoming message to her current activity. The modus of notification could then be adapted for instance between “silent”, “visual notification”, “vibration”, or “sound”. Automatically detecting interruptibility is only the first step towards this goal. In order to develop the full functionality of message prioritization (as envisioned earlier in [5]), additional software components that rate the message’s relevance with respect to the user’s current task, or translate the overall message prioritization into adaptive message notification, are necessary.
In this paper, we are concerned with using interruptibility as a way to measure the “level of activity” and not with detecting the “content of activity”, i.e. which project meeting a user is currently attending, whether it is personal or work time, etc.

Hypothesis

The hypothesis that we investigate in this paper is that a user’s interruptibility can be computed automatically using location and calendar information, which are typically available in today’s mobile phone settings. We assess this hypothesis once from a technical and once from a user-oriented point of view:

- Does GPS recording work accurately enough for interruptibility calculation (indoors, in special places)?
- Does the computed interruptibility value correspond to the users’ own perception of interruptibility?

3 Related Work

Mobile phones are the ‘perfect’ devices to interrupt the user in her current activity, both per incoming messages or phone calls. Thus they are widely used to examine interruptibility issues. In one study, [1], 1201 incoming phone calls were examined regarding the characteristics users tend to look for when deciding whether to pick up a phone call or not. The authors found that in 87.4 % of the cases, the study subjects use the caller ID (“who”) as decision support, in contrast to the current local social (34.9%) or cognitive contexts. The authors of the study examine two fundamental paradigms of interruption management based on their goals: (i) reduce negative impacts of interruptions and (ii) harness the utility of interruptions. One of the key findings is that users prefer getting unnecessary information to potentially missing important messages or calls. The framework presented in this paper follows this approach, and therefore encourages the use of a “show all messages” tab to allow the user to double check the incoming messages or calls. We plan to follow these findings by making sure that all messages arrive, and just modifying the modus of notification. We note however, that we do not deal with phone calls, but with textual messages, which are per definition asynchronous. Those messages are coming from an existing service framework.

Interruptibility can also be computed based on physiological user data. Chen et al. [6] describe a study on the use of physiological measurements for predicting interruptibility. The authors measure Heart Rate Variability (HRV) and Electromyography (EMV) signals during tasks like reading, typing, solving puzzles etc. Adding additional sensors may allow better identification of the user’s current task or status. However, physiological sensors are not highly usable in daily life; and even quite popular and promising technologies like wearable sensor devices as presented in [7] were pushed away by the increasing success and availability of smartphones. Phithakkitnukoon and Dantu [8] use the sensors embedded in mobile devices for their context-aware mobile computing model (called ContextAlert) to configure the mobile phone alert mode accordingly. Rosenthal et al. [9] present a model for approximating interruptibility costs by learning users’ preferences for receiving audible notifications, to later on turn on or off phone volume automatically.

4 Computing Interruptibility

We use an interruptibility value (also referred to as interruptibility) to denote how disturbed a user will feel if interrupted. We chose a normalized value between 0 and 1, where 0 (low interruptibility) means “will feel very disturbed by incoming messages” and 1 (high interruptibility) means “will not feel disturbed by incoming messages”. We derive interruptibility from location and calendar information by using a rule based approach to categorize user behavior, and then to derive a user’s interruptibility. We primarily focus on interruptibility during daily work, including the way to or from work, lunch breaks, meetings etc.

Location Information is available on most modern smartphones by using the GPS sensor. Note that we currently use only plain GPS information, and not (yet) map information.

Taking the user history into account, location-related aspects like average speed, transport type currently used, current occupation, or movement levels can be derived. Based on location information and information inferred from location history, different types of movement (resting, walking, or running) or ways of transportation (by foot, by car or public transport) are estimated. Such estimations contribute to the computation of an interruptibility value. For instance, if a user is not moving, the interruptibility increases, because she apparently tends to be more available for messages. When the user is running, the interruptibility drops a bit. Additionally, location information is also used to draw conclusions about the user's current activity (e.g., heading towards the appointment, doing nothing of relevance, and heading away from appointment). To eliminate short-term effects caused by for instance blind alleys or other obstacles, the history of several past movements is taken into account. Thus, "moving towards the appointment" does not automatically mean that the user exclusively heads towards the appointment in all recorded data points, but gives an indication about her general direction of movement in the past few minutes. Since this approach is tailored to an existing platform which at the moment only exploits GPS information, other location based technologies like NFC, Bluetooth or WiFi are under discussion, but not used yet.

Calendar Information (Appointments) are the second aspect we consider in our interruptibility computation; currently we use appointments and not (yet) tasks. Appointments are central when dealing with interruptibility, since they state clear user constraints regarding time and ideally also location (which is especially relevant in a mobile setting). We see the calendar as a central metaphor for managing personal time. Time and location information of appointments informs us of constraints. For instance, a user can only be at one location at a time, spatial distances between appointments put constraints on temporal distances between appointments. We consider only appointments in the present or in the near future.

Combining Location and Calendar Information

The possibility of reaching an upcoming appointment (derived from combining current location and location of appointment) is affecting interruptibility, as well as the user's current actions (change of location). Our system recognizes three kinds of user movement with respect to the next appointment:

(A) User does nothing of relevance to her next appointment.
(B) User moves towards the appointment location.
(C) User moves away from the appointment location.

Our system also determines the user's urgency with respect to her next appointment:

(1) User has plenty of time to reach the next appointment.
(2) User has just enough / or not enough time to reach the next appointment.

The final interruptibility value is a combination of interruptibility history and detected categories. The interruptibility computation is initialized with a default starting value; from then on, each subsequent interruptibility value is computed by raising or lowering the previous value based on which categories have been detected. Both single categories and combinations of categories (there are six combinations of "user movement with respect to reaching the next appointment(s)" and "urgency") lower, raise or keep the interruptibility value at the same level.
For instance, "User has not enough time to reach the next appointment" lowers a user's interruptibility (we assume she is under pressure to reach the next appointment).
The combination A1 does barely affect the user's current interruptibility, C2 on the other hand decreases interruptibility quite significantly since we assume that the user is in a hurry to reach the appointment. The parameters (how much to lower/raise individual values; upper and lower boundaries for each category and combination) of the algorithm presented here were determined using a simulation environment in which we can (re)play predefined scenarios and thus refine the interruptibility computation.

5 Evaluation Prototype

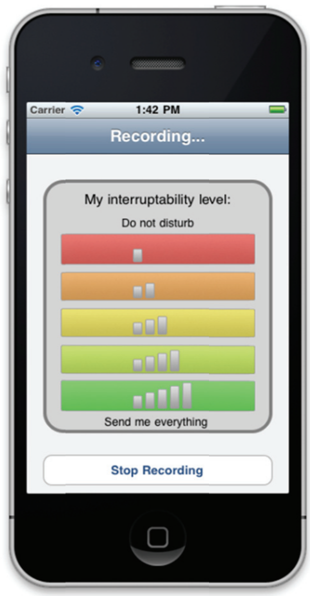

Figure 1: iOS evaluation client

The prototype for our preliminary study consists of a mobile client (iOS) and a server component for collecting and processing the incoming data. The client automatically submits location information to the server component every 30 seconds. Missing network coverage leads to data points not transferred to the server component, but still available in a local (on the mobile client) log file that can be exported after recording. This functionality would also be part of a client intended for real usage. Moreover the mobile client helps in creating a "gold standard" data set by asking the user for her personal estimation of her interruptibility every five minutes. Figure 1 shows the evaluation client's interface asking the user to estimate her interruptibility on a five point scale. The scale is visually enriched by the color scheme also used at traffic lights and icons well known for displaying the current reception level. One bar means "low reception, only the most relevant information", while 5 bars mean "send me everything". The users' manually entered interruptibility as well as the time needed to actually answer the request are sent to the server. The server component receives information from the mobile client, and computes the user's interruptibility in real time. In a prototype intended for real usage, the interruptibility would be communicated to interested components; in the evaluation prototype, the interruptibility is only stored alongside the user session data. Convenience tools are available for tracing and inspecting different aspects of interruptibility calculation, and it is also possible to reuse the recorded user data against different settings of interruptibility computation.

6 Study

Study Setup

We devised a scenario that was followed by all participants in our study: The fictitious use case person is an employee in a knowledge management company.
She starts her lunch break at 14:30, leaves the office and heads towards a restaurant for lunch. This restaurant is approximately 10 minutes walking distance away, which gives here about 30 minutes for lunch, since her next appointment at the office is scheduled 15:30 o'clock. The whole scenario takes a little more than one hour, and she should reach the appointment easily. Since the lunch takes place outside the office, she moves away from the location of the next appointment, and after lunch she returns. All participants were asked to follow this scenario, and to enter their own perception of being available / not available whenever the evaluation client asked for. After recording the scenario, the participants were interviewed using a graphical visualization showing both their manually set interruptibility values and the computed ones. In this interview, the participants were asked to explain the situations in the chart where the manually set interruptibility changed significantly. Finally, the study participants filled out a questionnaire concerning their general attitude regarding interruptibility. A central part of this questionnaire was a matrix where the participants were asked to state how they think their interruptibility changes based on the two dimensions mentioned in Section 4.

Study Participants

The preliminary study was carried out with three participants, male and between 28 and 35 years old. All of them work as researchers and developers in the authors' research group. This first study also serves as testing environment for architecture, prototype and experiment setting and delivers insights to further develop and refine a consistent interruptibility model.

7 Results

7.1 GPS Recordings and GPS-based Inferences

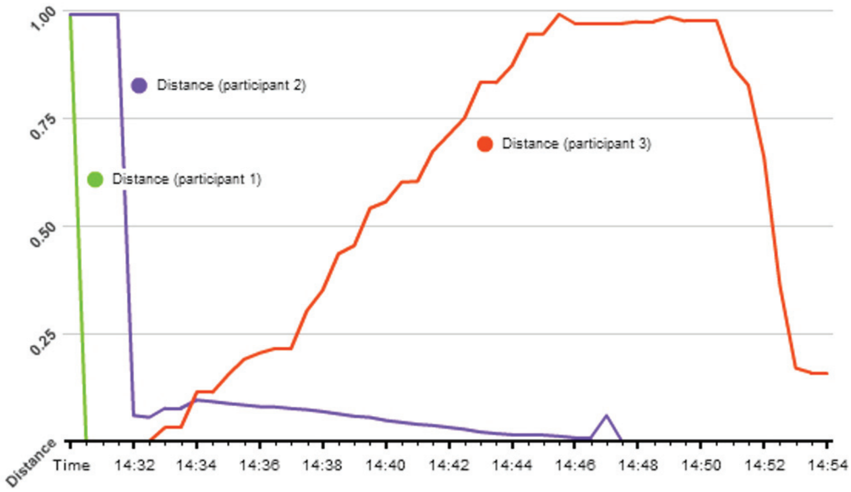

Figure 2: Visualization of “distance to appointment”

GPS Recordings

The raw GPS values were regularly erroneous. Figure 2 visualizes “distance to appointment” for the three user scenarios recorded. Since all three participants started their lunch break at the office, where also the next appointment was scheduled, the initial, correct, distance to appointment should be close to zero. However, only for Participant 3 the distance to the next appointment behaves as expected, i.e. it slowly increases; then the user heads back towards the office, i.e. appointment and finally reaches. For both other recordings (data from participants 1 and 2) the GPS signal at the beginning of the scenario was not correct, i.e. the GPS data showed the participants to be a few kilometers away from the appointment. However, the bogus location data at the beginning only affects a few data points, so the overall computation of interruptibility is correct. We can detect incorrect location information by comparing new values with user activities in the past (e.g., calculate average speed and ignore new values which exceed this average speed by predefined percentage). Even if unusual jumps between data points are found, this does not automatically mean the system has delivered wrong data; the user could have entered a car or public transport etc.

GPS-based Inferences

Since we infer the transportation type (resting, walking, etc.) from GPS values, erroneous GPS values also impact these inferences. This problem was dealt with by assuming a pedestrian scenario and thus setting boundary values on average speed, distance to appointment, etc. For instance, if the average speed is suddenly several hundred meters per second, this value is rejected as implausible.

General Discussion
The GPS recording interval should not be too small, such that (i) the user has enough time to clearly move into one direction, but also (ii) battery and mobile network are not overused. During the study we asked for a GPS value every 30 seconds. This is quite battery consuming, and since GPS-based inferences work sufficiently well with 30 seconds, one could experiment with larger intervals for GPS sensing. An intermediate solution would be to record GPS values every 30 seconds but to send them to the server only in larger intervals. We also use an experiment setting without fixed intervals, in that case the system queries the user interruptibility each time the computer generated interruptibility dramatically changes; this means that the system queries for each major change whether it is also true in reality or not.

7.2 Human Interruptibility

Comparability of the Three Captured Scenarios
The computed interruptibility values of all three different recordings are visualized along the time in Figure 3.

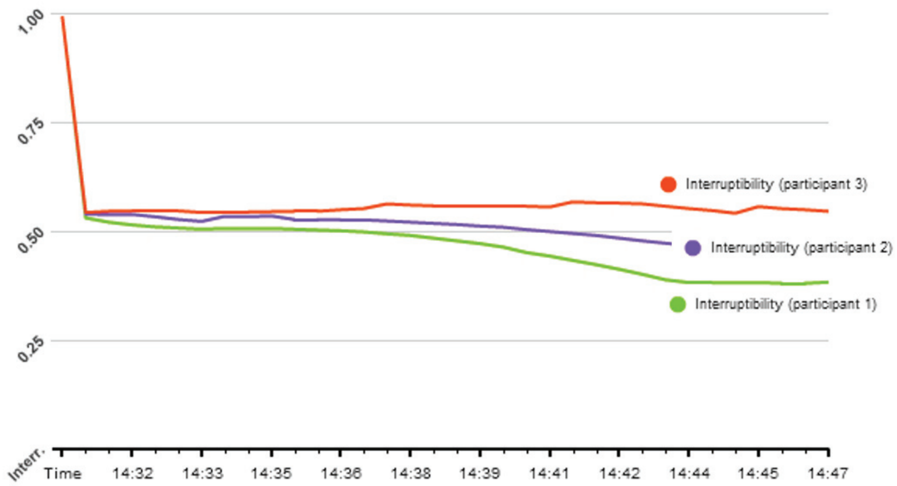

Figure 3: Computed interruptibility value over time
for the three different scenario recordings

In all three scenarios, the computed interruptibility value slowly falls over time. This meets our expectations, because the next appointment was less than one hour away, and therefore is significantly affecting interruptibility. At the beginning of the scenario the participants also walked away from the appointment location, and even though they are returning to the appointment location after that and make it in time, the close distance (in time) to the appointment decreases interruptibility. The differences between the three scenarios lie in different walking speeds and lunch break durations of the three participants. Using the categories presented in Section 4 to identify the various stages of the scenario, the scenario starts with C1: The participants are moving away from the location of their next appointment, but still have plenty of time to reach it. Over time, the decreasing amount of time left until the appointment causes the interruptibility value to sink. After the lunch break, the scenario corresponds to B1, i.e. the participants have plenty of time to reach the next appointment and act to reach the appointment, the interruptibility value still decreases again because of the approaching appointment. It can be seen in Figure 3, that from the point of view of the interruptibility algorithm, the three scenarios are definitely comparable, as expected. This means, that the study participants behaved similarly in the given scenario as far as the observed (and interpreted) data goes.

Agreement of Study Participants with Computed Interruptibility

The first scenario run is visualized in Figure 4. Participant 1 states that he would like to get only the most important messages during lunch break (A1, 14:45, contrary to his questionnaire where he stated that he would like to raise his interruptibility). Contrary to participant 2 and 3, his interruptibility goes up and down when heading for his lunch break (C1), the same behavior after the lunch break (B1). This scenario run heavily shows progressive stages, the interruptibility value goes up at the end.

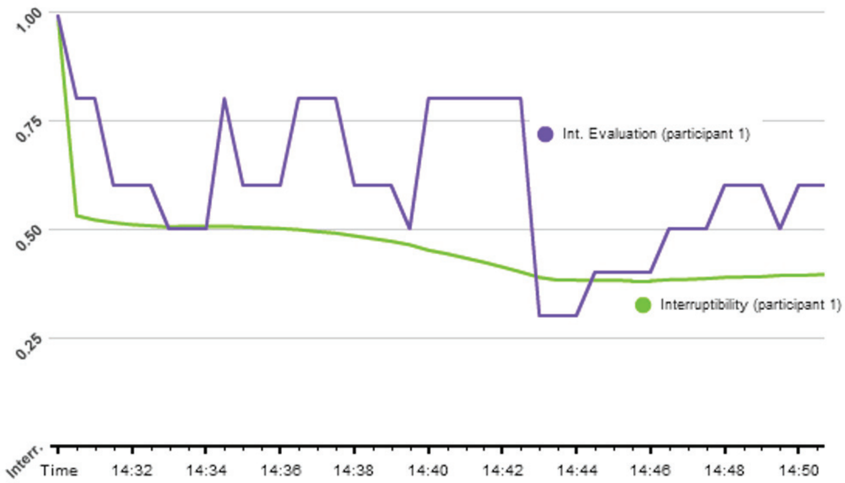

Figure 4: Comparison between computed and recorded human interruptibility of participant 1

Figure 5 visualizes the second user doing the same scenario (participant 2): The user seems to be much occupied during work, because the scenario starts with a very low interruptibility value. This user thinks that messages are quite OK on the way to lunch break (C1, 14:30 to 14:35, but ranked as neutral in the questionnaire), and on the way back to the appointment (B1, 14:36 to 14:45, which he agreed on in the questionnaire), but does not want to be disturbed during his lunch break (A1, which takes place approximately at 14:35, but which he ranked as high interruptibility). The lunch break tends to be very short; this is because the user is queried for his interruptibility every five minutes, and it seems that this is only once the case during the lunch break.

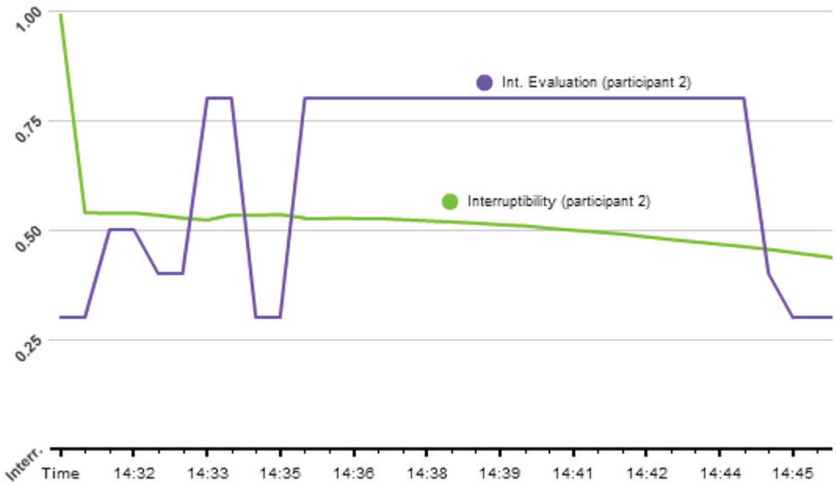

Figure 5: Comparison between computed and recorded human interruptibility of participant 2

Figure 6 compares the computed interruptibility values for participant 3 with his own estimates for his interruptibility.

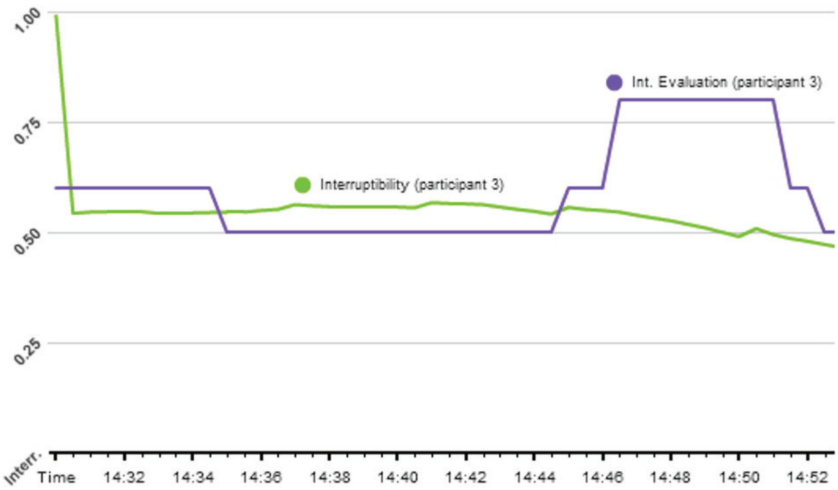

Figure 6: Comparison between computed and recorded human interruptibility of participant 3

The user leaves the office and heads for his lunch break. Ding his walk towards the location of his lunch breach, he does not want to be disturbed (C1, 14:34 to 14:45), but during lunch break (A1, 14.45 to 14:53), messages are welcome. When returning and heading towards the appointment (B1, from 14:53 onwards), participant 3 would expect his interruptibility to decrease, because he will be occupied soon and wants to only receive the most important messages.

In the questionnaire this user rated the situations C1 and B1 occurring here as neutral, so interruptibility should stay the same; however in the scenario, he decreased his interruptibility on the way to his lunch break and on the way back. A1 (the lunch break) was ranked according to his questionnaire (increase interruptibility). Overall, the user study participants thus did not agree very well with the computed interruptibility values, and also differed amongst each other.

Expectation of Study Participants towards Interruptibility Algorithm
Because study participants were found to have divergent opinions on when they are available for incoming messages, we took a closer look at the expectation of the study participants towards the behavior of an interruptibility computing algorithm. Most of the participants agree that "doing nothing of relevance" and with enough time left (see A1) means that they would accept getting more messages since they have spare time. Most of the other matrix fields differ to some extend; the impact of B1 (acting to reach the appointment with enough time left) for instance heavily depends on the participant. While some users claim that the interruptibility should stay the same in B1, others say that they want more messages, since everything is fine and they are already heading the next appointment. One participant argues that because he is right now heading his next appointment, he has a goal to think of, and so would like to receive fewer messages. All participants agree on the impact of C2 again, "acting towards the appointment" with "user is late" means that they want to receive fewer messages (only the important ones). Since interruptibility calculation is partly based on conclusions drawn based on location information, the actual impact of the different values on interruptibility has to be defined a way it reflects most of the target users. But even the impact of "resting" (which means no movement or at least very low average speed) can't be defined in a generally accepted manner. While at first sight "resting" means that a user has some spare time or at least more time to read messages, our preliminary results show that this heavily depends on the individual user attitude: most participants agree that resting does not automatically mean they are open for all types of messages.

8 Discussion and Outlook

Results show that GPS works essentially accurately enough to categorize type of movement and transportation. This means that the first hypothesis form Section 2 ("Does GPS recording work accurately enough for interruptibility calculation?") can be approved. However, in certain situations (e.g., starting GPS recording within a building), additional algorithms have to check for the plausibility of location information. Our study also shows that the concept of interruptibility cannot be generalized to work for all users the same way, so the second hypothesis ("Does the computed interruptibility value correspond to the users' own perception of interruptibility?") has to be rejected. Indeed, the study participants have significantly diverging opinions on when they are more and when less available for incoming messages. In the follow-up interviews, all participants stated that an exact description of which services or messages would be delivered to them at which individual interruptibility would help them defining which current level to choose.

This, together with the divergence of expectations towards the interruptibility values in similar situations, indicates that interruptibility is not a generally meaningful concept for humans, but needs to be contextualized with regard to which specific messages are incoming. This falls well within our general goal of providing contextualized message ranking for automated messages, although we had hoped to be able to separate interruptibility from the relevance of messages to a user's context.

ACKNOWLEDGMENTS
The Know-Center is funded within the Austrian COMET Program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG. This research has also been co-funded by sayService srl, Trento, Italy. The research leading to these results has received funding from the FFG.

9 References

1. Grandhi, S., Quentin, J.: Technology-mediated interruption management, International Journal of Human-Computer Studies 68, 288-306 (2010)
2. Speier, C., Valacich, J.S., Vessey, I.: The effects of task interruption and information presentation on individual decision making, Proceedings of the eighteenth international conference on Information systems, 21-36, December 14-17, United States (1997)
3. Cooper, R., Franks, B.: Interruptibility as a constraint on hybrid systems, Minds and Machines, Volume 3 Issue 1, Springer Netherlands, 73-96 (1993)
4. Fogarty, J., Hudson, S.E., Atkeson, C.G., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J.C., Yang, J.: Predicting Human Interruptibility with Sensors, ACM Transactions on Computer-Human Interaction, Volume: 12, Issue: 1, Publisher: ACM Press, 119-146, New York, NY, USA (2005)
5. Stern, H., Ulbrich, A., Pammer, V., Pistore, M., Bertoli, P., Lindstaedt, S.: Message Filtering Based On Mobile User Context, V. Ermolayev, J.M. Gomez-Perez, P. Haase, P. Warren: Proceedings of the Second Workshop on Context, Information and Ontologies, Vol-626, Workshop at Knowledge Engineering and Knowledge Management (EKAW) 2010, Lisbon, Portugal, October 11 (2010)
6. Chen, D., Hart, J., Vertegaal, R.: Towards a Physiological Model of User Interruptability, Proceedings of the 11th IFIP TC 13 international conference on Human-computer interaction, Volume Part II, Springer-Verlag Berlin, Heidelberg (2007)
7. Kern, N., Amtifakos, S., Schiele, B.: A Model for Human Interruptability: Experimental Evaluation and Automatic Estimation from Wearable Sensors, Proceedings of the Eight International Symposium on Wearable Computers (ISWC 04) (2004)
8. Phithakkitnukoon, S., Dantu, R.: ContextAlert: Context-Aware Alert Mode Control for a Mobile Phone. International Journal of Pervasive Computing and Communications, Vol. 6, Issue 3 (2010)
9. Rosenthal, S., Dey, A.K., Veloso, M.: Using Decision-Theoretic Experience Sampling to Build Personalized Mobile Phone Interruption Models, In Proc. International Conference on Pervasive Computing (Pervasive 2011) (2011)