Xensible Interruptions from Your Mobile Phone

G.H. (Henri) ter Hofte
Telematica Instituut
Enschede, the Netherlands
Henri.terHofte@telin.nl

ABSTRACT
Mobile phones may interrupt in any place at any time. Using the SocioXensor research tool on people's own mobile phones, we conducted an experience sampling study to explore which context information predicts a person's availability for a phone call, and which context information people wanted to disclose to particular social relations. Like other studies, we found that a small set of context information can help initiators of phone calls to improve their ability to know when recipients are receptive to phone calls. We also found that if we restrict the information to information recipients actually want to disclose, which is only a small subset of all information, enough context information is still available for initiators of phone calls to improve their ability to know when recipients are receptive to phone calls.

Categories and Subject Descriptors
H5.2 [Information interfaces and presentation]: User interfaces; H5.3 [Information interfaces and presentation]: Group and Organization Interfaces – collaborative computing; H1.2 [Models and Principles] User/Machine Systems; I2.6 [Artificial Intelligence]: Learning

Keywords
Interruptions, Context-aware telephony, ESM, SocioXensor

1. INTRODUCTION
Mobile phones support communication "with anyone, at anytime, in anyplace". While this may sound just fine for people as an initiator of phone calls, as a recipient of phone calls, many people already found out that being communicated to by anyone at anytime in anyplace is *not* what they want.

In a study among managers of a research laboratory into availability for interruptions, Hudson et al. [7], concluded that "managers struggle with finding the balance between entertaining useful interruptions and avoiding distracting ones" and that designers of systems should focus on "making interruptions more effective rather than on decreasing them". We believe that these observations extend to interruptions from mobile phone calls and to environments other than work.

Mobile HCI'07, September 9-12, 2007, Singapore.

Two broad approaches can be distinguished to resolve these problems. One is to improve the filtering by or for the recipient of phone calls, which may range from improving the caller ID presentation to automatic profile switching based on context or filtering out calls from certain persons based on context. The major disadvantages of this approach is that improved filtering *by* the recipient still represents an interruption and automatic filtering *for* the recipient bears a risk of missing important phone calls.

The other approach, also known as context-aware telephony [10], seeks to provide an initiator of phone calls with context information about the recipient, such that the initiator can make better informed decisions if and when to make a call. Examples of context information that have been used in context-aware telephony applications include location (e.g. home/office/transit [11][12], or room number [9]), activity (e.g. calendar [12], computer use [2]), conversation (e.g. phone use [9][14]) and social context (e.g. people close by [9]).

Following context information on its path from the recipient's context to the initiator's behavior, we encounter various research questions for context-aware telephony. Below, we list them, along with references to typical studies focusing on that question.

(1) Which context information of a recipient can be captured (semi-)automatically with low cost – including low additional effort from the recipient – and high quality? [6]
(2) Which context information is predictive for a recipient's availability for an interruption like a phone call? [5]
(3) Which context information is a recipient willing to share with which potential initiators in which context? [3][8]
(4) Which context information helps to improve an initiator's *understanding* of a recipient's availability for an interruption like a phone call? [1]
(5) Which context information helps to improve an initiator's *behavior* with respect to calling recipients at better moments?

A few studies explore several research questions at the same time. In several studies (e.g. [5], [2]), researchers used devices to capture context information (1) and study to what extent it was predictive for a recipients availability for interruption (2).

In this short paper, we report on a study that seeks to answer a combination of (2) and (3), something no other study has addressed as far as we are aware. First, we describe the research question, the method used (instrument, subjects, procedure, analysis). Then, we describe the results obtained from the study. Finally, we discuss results and present conclusions.

2. Research Question
Which context information is a recipient willing to share with potential initiators of a phone call and is that information predictive for the recipient's availability for a phone call?

3. Method

In order to minimize problems with retrospective recall, we used the Experience Sampling Method (ESM) [4]. Unlike many other ESM studies we used the devices that most subjects used as their primary mobile phone as the ESM device. The ESM study was part of a larger study in which we also collected data for a longer time about battery usage, visible GSM cells, Wi-Fi access points and Bluetooth devices. This paper only reports on the ESM data.

3.1 Instrument

Each experience sample consisted of 6-7 questions (see Table 1). We modeled our ESM questions after [8]. Compared to that study, we asked more details about being in conversation (q4), which was found to have high predictive value in e.g. [5]. Inspired by [3], we hypothesized that people may be more inclined to share location on a coarse level, and not at a detailed level as asked in [8]. In order to keep time spent per sample low, we left out the activity-related questions from [8].

Table 1. Experience sampling questions

nr	Question	Answer options				
q1	<Someone> is going to call. What availability information would you like to show to him/her?	Red light (do not disturb)	Orange light	Green light		
q2	<Someone> is going to call. What information would you like to disclose to him/her?	where you are (coarse)	where you are (detail)	whether you are in conversation	whether you have company	none of the above
q3	Are you in conversation?	yes	no			
q4	How are you in conversation?	face to face	via fixed telephone	via mobile telephone	via Instant Messaging	otherwise
q5	Where are you now?	at home	at the office	in transit	somewhere else	
q6a	Where are you at home?	living room	kitchen	a bedroom	garden or driveway	elsewhere at home
q6b	Where are you at the office?	my own office room	colleague's office room	meeting room	company restaurant	somewhere else
q6c	How are you in transit?	car	public transport	bicycle	foot	otherwise
q6d	Where elsewhere?	shop	theatre/ cinema/ restaurant	sport/hobby place	someone's house	somewhere else
q7	With how many people are you? (incl. yourself)	1	2	3	4	5 or more

<Someone> in q1 was randomly selected from {Your partner, A family member, A friend, A colleague, Your boss, An unknown}. <Someone> in q2 was always the same as <Someone> in q1. All questions required exactly 1 answer, except for q2 and q4, which allowed for multiple answers (but "none of the above" was an exclusive option). q4 was only presented when a subject answered "yes" to q3. Which version of q6 (a/b/c/d) was presented depended on the answer of the subject to q5.

3.2 Subjects

We recruited 10 employees of a research organization, via an e-mail sent to employees that used a phone compatible with the ESM software we used. Excessive crashes caused two subjects to drop out early. Two additional subjects were recruited to use a new device. This resulted in 10 subjects (7 male), age 31-47 (average 38).

3.3 Procedure

We briefed the subjects in a group meeting about the goal of the study, the duration, part of the day, the expected average number of samples per day and the estimated time of about 20 seconds it would take to complete a sample. We asked the participants to take the device with them everywhere, also to places where they normally would not take their mobile phone, but not to places that would pose serious hazards to themselves or the device (e.g. shower, pool). We explained all questions in order to maximize a consistent interpretation. Finally, we handed out informed consent forms for the study, explained that starting the study entitled them to an incentive of €25, or a 2 GB miniSD card (market value €50), that they could terminate their participation from the study at any time without penalties or loss of benefits and how the study helpdesk could be reached during and outside office hours.

At the beginning of the study, we obtained the signed informed consent form and installed a pre-configured version of SocioXensor[1] [13], a freely available research tool, on each subject's Windows Mobile 5.0 PocketPC phone. Most subjects used the device as their primary mobile phone, 2 subjects used it as a PDA and had a separate primary mobile phone. Each subject received experience samples for 7 days, scheduled between 8 AM and 10 PM, with inter-sample intervals according to a uniform random distribution between 45 and 105 minutes. A sample was announced with an auditory/tactile alarm (cf. device settings for such alarms). If the subject would not respond, the alarm would continue to sound for 1 minute, and would time out after another minute. Subjects could respond by selecting "Ok" to start answering questions, or "Not now", if they were not willing to answer the sample at that time. Subjects could also suspend all sampling for 30 minutes, 1, 2, 4 or 8 hours via the "Suspend" menu item in the Xensor screen and selecting the suspend duration of their choice. Similarly, subjects could manually resume sampling before the device would resume automatically. Shortly after the 7 days, we collected the data, removed SocioXensor from their device and held a brief interview. In a brief survey, we inquired about device crashes, problems with questions and answers, the occurrence and reasons for non-response and use of not-now and suspend/resume.

3.4 Data Analysis

We used the Weka machine learning environment [15] to find models that use various attributes (context variables) to predict availability. All of our models were built using a naïve Bayes classifier and correlation-based attribute selection, a strategy that prefers subsets of attributes that are highly correlated with the phenomenon to be predicted, while having low intercorrelation. To evaluate the models, we used a standard 10-fold cross-validation approach involving 10 trials of model construction. In each trial, 90% of the samples were used to train the model and the remaining 10% were used for testing. Each sample was used for training in 9 trials and for testing in 1 trial.

4. Results

In total, the study covered 1680 hours, during 980 of which samples could have occurred (98h per subject, 14h per subject per day, 8 AM - 10 PM). We expected a sample on average every

[1] SocioXensor is short for 'Social eXperience sensor'.

(0h45+1h45)/2 = 1h15, i.e., 784 samples in total (980/1h15). We logged 785 samples with an inter-sample interval distribution nicely matching the expected uniform distribution. Of those 785 samples, 35 were logged as missing due to the device being completely off or crashed at the time the sample should have been taken. We also logged 55 device restarts, less than one per subject per day. The post-study survey and interviews revealed that most of these were caused by spurious crashes of the device. Battery samples, scheduled once per minute, reveal that 6% were missing; during 95h50 (6% of 1680h) no sampling could have taken place.

Of the 750 samples (100%), 13 (2%) were suppressed since 3 out of 10 subjects chose to suspend all sampling explicitly for a short time, 166 (22%) were not reacted to within 2 minutes and timed out, 12 (2%) were reacted to with the "Not now" option, and 7 (1%) were not completely answered (i.e. the subject started answering, but a question timed out after 2 minutes). This leaves 552 completely answered samples that we use in our analyses below; a response rate of 74%. The response rate per subject varied between 56% and 97%. On average, answering a sample took 30s: longer than we expected (samples varied 12s-237s, averages per subject varied 22s-55s).

The most frequently reported detailed location was home-living (29%, amounting to 28h45 per week, see also Figure 1), followed by office-ownoffice (25%, 24h40). The fraction of samples reported as in transit (10%) amounts to 9h45. The fraction of samples representing time during which a mobile phone is likely the primary phone to reach a person (all excluding home and office-ownoffice) adds up to 33% (32h39).

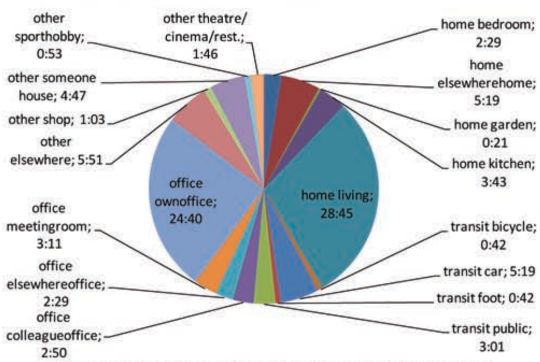

Figure 1. Fraction of locations reported in samples (and time in that location as fraction of 98h sampled/week)

4.1 Predicting Availability

In this section we first investigate which context information helps to recognize situations when a recipient is highly available for a phone call (i.e. when the subject answered "Green light" to q1), which is was the case in 51.1% of samples.

A GreenLight predictor built with Weka using a naïve Bayes classifier and correlation-based attribute selection uses the following subset of attributes in the predictor:

- Social_relation (partner, family member, friend, colleague, boss, unknown)
- InConversation (the answer to q3)
- InConversation=face to face (q4)
- LocationCoarse=at home (q5)

After 10-fold cross-validation, the model prediction for each of the 552 test cases is compared with the answer the user gave in that case. The results are presented in Table 2.

Table 2. Confusion matrix of the GreenLight predictor

	model prediction		
self-report	Green	NotGreen	TOTAL
Green	194 35.1%	88 15.9%	282 51.1%
NotGreen	111 20.1%	159 28.8%	270 48.9%
TOTAL	305 55.3%	247 44.7%	552 100%

The accuracy of this predictor is 63.9%, which corresponds to the fraction of all correctly predicted cases ((194+159)/552) in the 10-fold cross-validation. This predictor can be compared with a base predictor, a trivial predictor that always predicts "Green light", which is the case reported most often (282) and which has an accuracy of 51.1% (282/552). The accuracy of the GreenLight predictor is significantly higher ($\chi^2(1, 1104) = 18.69$, $p < .001$).

4.2 Disclosing Context to Social Relations

Similar to earlier research [8], we found (see Figure 2) that sharing patterns vary greatly across different social relations. In addition to (Share_Location_Detail), we also asked whether subjects would like to disclose their coarse location (Share_Location_Coarse), and calculated when subjects were would like to disclose either (Share_AnyLocation), which was often the case: 70% on average, varying between 94% (partner) to 9% (unknown). On average, subjects would like to disclose coarse locations 64%, whether they were in conversation 34%, their detailed location 25%, and whether they were in company 20% of the time.

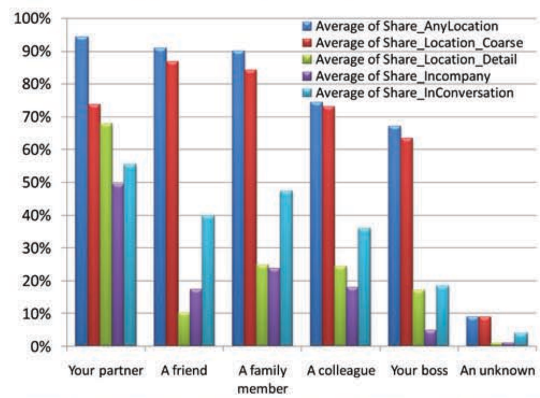

Figure 2. Average disclosure rate for different types of context information across different social relations

4.3 Predicting Availability Using Shared Data

Finally, we creating "Shared_" variants of the context factors representing information that initiators would be able to observe, by removing information recipients would not like to disclose to them. The resulting predictor uses the following attributes:

- Social_relation (partner, family member, friend, colleague, boss, unknown)
- Shared_InConversation (q2+q3)
- SharedAny_Location_Coarse=at home (q2+q5)

The accuracy of the GreenLightShared predictor (see Table 3) is 61.2% ((192+146)/552) still significantly higher than the base predictor ($\chi^2(1, 1104) = 11.54$, $p<.001$).

Table 3. Confusion matrix of the GreenLightShared predictor

	model prediction		
self-report	Green	NotGreen	TOTAL
Green	192 34.8%	90 16.3%	282 51.1%
NotGreen	124 22.5%	146 26.4%	270 48.9%
TOTAL	316 57.2%	236 42.8%	552 100%

5. Discussion and Conclusion

Our subjects indicated high availability ("Green light") for phone calls only 51.1% of their time. Initiators of a phone call that simply guess that the recipient is always available would have been right 51.1% of the time.

Initiators that use the GreenLight predictor based on a naïve Bayesian network that learns from a small subset of context information (type of social relation between recipient and initiator, whether the recipient is in conversation, in particular whether the recipient is in face to face conversation and whether the recipient is at home) would fare significantly better: 63.9% of the time ($p<.001$).

However, recipients do not want to disclose all information to everyone. In particular, our subjects indicated they wanted to share whether they were in conversation for only 35% of the time. This does not appear to be detrimental to the prediction, however: initiators that use the GreenLightShared predictor, that uses only information from cases that recipients want to share (type of social relation between recipient and initiator, shared information about being in conversation and shared information about being at home), still would fare significantly better than just guessing: 61.2% ($p<.001$).

Although these improvements are *statistically* significant, a 10% increase in accuracy for estimating a person's availability for a phone call may not be a significant gain in *practice*. Further research will have to establish whether predictors for individual recipients are more accurate and whether context-aware telephony systems based inspired by these results lead to improved *understanding* when a recipient is available and, ultimately, to improved calling *behavior*.

6. Acknowledgments

We thank Raymond Otte, Arjan Peddemors, Ingrid Mulder and Peter Ebben for their help and support. Also, we thank all subjects for their participation in this study and the MobileHCI 2007 reviewers for their comments and suggestions. The projects FRUX and AWARENESS in the Dutch Research Programme Freeband Communication supported this research under contract BSIK 03025.

7. References

[1] Avrahami, D., Fogarty, J. & Hudson, S.E. Biases in human estimation of interruptibility: Effects and implications for practice. In *Proceedings of CHI 2007*. ACM: New York, 2007, 50-60.

[2] Begole J., Matsakis N. E. & Tang J. C. Lilsys: Sensing unavailability. In *Proceedings of CSCW2004*, ACM: New York, 2004, 511-514.

[3] Consolvo, S. Smith, I.E., Matthews, T. LaMarca, A., Tabert, J. & Powledge, P. Location disclosure to social relations: Why, when, & what people want to share. In *Proceedings of CHI 2005*. ACM: New York, 2005, 81-90.

[4] Csikszentmihalyi, M. & Larson, R. Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease, 175* (1987), 526-536.

[5] Fogarty, J., Hudson, S.E, Atkeson, C.G., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J.C., and Yang, J. Predicting Human Interruptibility with Sensors. *ACM TOCHI, 12*,1 (2005), 119-146

[6] Ho, J., & Intille, S.S. Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In *Proceedings of CHI 2005*. ACM: New York, 2005, 909 - 918.

[7] Hudson, J.M., Christensen, J., Kellogg, W.A. and Erickson, T. "I'd be overwhelmed, but it's just one more thing to do": Availability and Interruption in Research Management. In *Proceedings of CHI 2002*, ACM: New York, 2002, 97-104.

[8] Khalil, A. & Connelly, K. Context-aware telephony: Privacy preferences and sharing patterns. In *Proceedings of CSCW2006*, ACM: New York, 2006, 469-478.

[9] Raento, M., Oulasvirta, A., Petit, R., & Toivonen, H. ContextPhone: A prototyping platform for context-aware mobile applications. *Pervasive computing, 4*, 2 (April-June 2005), 51-59.

[10] Schilit, B.N., Hilbert, D.M., & Trevor, J. (2002). Context-aware communication. *IEEE Wireless Communications, 9*, 5 (2002), 46-54.

[11] Schmidt, A., Stuhr, T. and Hans Gellersen. Context-Phonebook - Extending Mobile Phone Applications with Context. In *Proceedings of MobileHCI 2001*, IHM-HCI: Lille, France, *2001*.

[12] Ter Hofte, G.H., Otte, R.A.A., Kruse, H.C.J., & Snijders, M. Context-aware communication with Live Contacts. In *Conference Supplement of CSCW2004*, 2004.

[13] Ter Hofte, G.H., Otte, R.A.A., Peddemors, A., Mulder, I. What's Your Lab Doing in My Pocket? Supporting Mobile Field Studies with SocioXensor. In *Conference Supplement of CSCW2006*. ACM: New York, 2006, 109-110.

[14] Tang, J.C., Yankelovich, N., Begole, J., Van Kleek, M., Li, F., & Bhalodia, J. ConNexus to awarenex: extending awareness to mobile users. In *Proceedings of CHI'01*, ACM: New York, 2001, 221-228.

[15] Witten, I.H. & Frank. E. *Data mining: Practical Machine Learning Tools and Techniques, Second edition*. Morgan Kaufman: San Francisco, CA, 2005.