I'll be there for you: Quantifying Attentiveness towards Mobile Messaging

Tilman Dingler
University of Stuttgart
Institute for Visualization and Interactive Systems
tilman.dingler@vis.uni-stuttgart.de

Martin Pielot
Telefonica Research, Barcelona, Spain
martin.pielot@telefonica.com

ABSTRACT

Social norm has it that people are expected to respond to mobile phone messages quickly. We investigate how attentive people really are and how timely they actually check and triage new messages throughout the day. By collecting more than 55,000 messages from 42 mobile phone users over the course of two weeks, we were able to predict people's attentiveness through their mobile phone usage with close to 80% accuracy. We found that people were attentive to messages 12.1 hours a day, *i.e.* 84.8 hours per week, and provide statistical evidence how very short people's inattentiveness lasts: in 75% of the cases mobile phone users return to their attentive state within 5 minutes. In this paper, we present a comprehensive analysis of attentiveness throughout each hour of the day and show that intelligent notification delivery services, such as *bounded deferral*, can assume that inattentiveness will be rare and subside quickly.

Author Keywords

Attentiveness; Responsiveness; Availability; Interruptibility; Mobile Devices; Bounded Deferral

ACM Classification Keywords

H.5.m Information interfaces and presentation: misc.

BACKGROUND AND MOTIVIATION

Exchanging messages via SMS or over-the-top messengers is one of the core use cases of mobile phones. For example, in 2011 teenagers were found to exchange a median number of 60 messages per day [9].

At the same time, people tend to expect responses to their messages within minutes [3]. To not violate these expectations, people need to be *attentive* to their phone, which means to check and triage new messages quickly after their arrival to decide how to act on them [11]. Previous work [2, 11, 13] has shown that, on average, people do so within a few minutes. Therefore, people are often forced to interrupt their current activity upon arrival of each new message. However, such interruptions have negative effects, as people find it difficult to return to the activity prior to the interruption [4, 8]. Hence, previous work has proposed different notification-delivery strategies to minimize the impact of interruptions.

Horvitz *et al.* [7] proposed *bounded deferral*: if a user is predicted to be busy, alerts are being held back until a more suitable moment, but only for a maximum amount of time. In the context of mobile phones, Fischer *et al.* [6] found that *opportune moments* for delivering notifications occur right after the user has finished a task, such as writing a message. Previous work [10, 11, 12] has explored the use of mobile phone sensors and usage patterns, such as the the user's location or recency of interactions, to automatically predict such opportune moments.

However, bounded-deferral strategies may not work if there are many long phases without opportune moments. If the the algorithm delays messages for too long, it may increase response times which in turn may lead to expectations being violated. If the maximum delay is too little, messages may frequently be delivered when the user is still occupied. Hence, bounded deferral will only be an ideal strategy when users are typically attentive, and when phases of inattentiveness are brief.

To see whether this is the case, we conducted a study where we determined the attentiveness of mobile phone users throughout the day. We used phone-usage data consecutively acquired over two weeks from 42 mobile phone users to infer attentiveness for each minute of the day. The contribution of the study is three-fold:

- We confirm previous work that attentiveness can be predicted from phone usage patterns, in our case with almost 80% accuracy. In contrast to previous work, our data set is much larger and also accounts for notifications handled on other devices;
- While there is a common-sense opinion that people are 'always' next to their phones, this work yields a quantitative estimate: our participants were highly attentive to mobile messages: 12.1 hours per day (more than 75% of the time spent awake); and
- We show that times of inattentiveness are often very short. In 50 / 75% of the cases, our participants returned to a state of attentiveness within 2 / 5 minutes.

MobileHCI '15, August 25 - 28, 2015, Copenhagen, Denmark
2015 ACM. ISBN 978-1-4503-3652-9/15/08 $15.00
DOI: http://dx.doi.org/10.1145/2785830.2785840

This data supports that bounded-deferral is a viable strategy for the design of intelligent notification-delivery systems that aim to reduce the interruption of notifications.

DATA COLLECTION

To collect the input for the model, we built an apparatus for data-collection. In July of 2014 it collected (1) when a message arrived, (2) how fast the user attended to it, and (3) contextual- and phone-usage data. We used this data to train a machine-learning algorithm, where the time until the user attended to messages served as ground truth, while the contextual- and phone-usage data served as features. Our data allowed us to compute the state of those features for each moment of the data collection phase. Once trained, we applied the algorithm to the contextual- and phone-usage data for each minute of the study. Thus, we filled the gaps for all those moments where users did not receive any messages: for each minute, we obtained a prediction of how fast a user would attend to a message at this time.

Participants

The majority of the participants were recruited through a mailing list of a German university and a mailing list to announce user studies hosted by a large Spanish IT company. Hence, the sample represents German students and Spanish people with at least some basic interest in new technologies. 42 participants joined the study and left the application running for at least two weeks. Providing demographic information was voluntary. 45.2% of the participants reported to be male, 23.8% to be female, and 31% did not report their gender. The mean reported age was 28.7 years ($SD = 5.9$).

Measures

Once the probe was installed and configured, a background service registered several sensor and event listeners, and started sending all events to a data server. This included the status of the screen (on/off), data from the proximity sensor, access to the notification center, the ringer mode, the app in foreground, and the incoming notifications.

To learn about incoming messages, we used the *NotificationListenerService*, which has been recently introduced with Android 4.3. Whenever a new message creates a notification, the listener fires a *notification-posted* event, which conveys the app to which the notification belongs. In contrast to previously proposed approaches to study notification activity, such as the *AccessibilityService* (*e.g.* [13]), the *NotificationListenerService* also fires an event when a notification is removed, which includes instances where a message has been read on a different device. We created a filter that only kept notifications from messengers, and ignored all other types of notifications. Further, we filtered out notifications coming in while the corresponding app was in the foreground.

Procedure

To join the study, people had to download our app from *Google Play*. When first launched, it presented an informed consent form, which explained the background of the study, and what information would be collected. If consenting with the terms, the app helped the participant to go to the settings to grant the application special access to incoming notifications, and subsequently began logging notification-related activity. Participants received a 20 EUR Amazon gift card as compensation.

RESULTS

Over the course of two weeks we collected 55,824 messages from 42 participants. Figure 1 shows the number of messages received in total during the different hours of the day.

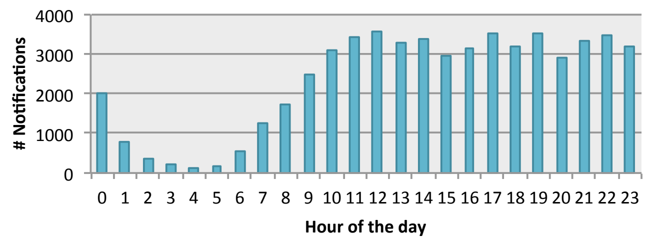

Figure 1. Total number of notifications per hour of the day. For example, '0' indicates notifications arrived between 0:00 and 1:00 o'clock.

Participants received a mean number of 66.8 (Mdn = 40) messages per day. Figure 2 illustrates the prevalence of messenger apps. With 77.7% of the messages, *WhatsApp* was by far the most popular messenger. Text messages only made up 1.8%. This could be expected, as in many European countries, WhatsApp has replaced SMS messages as primary way of exchanging messages.

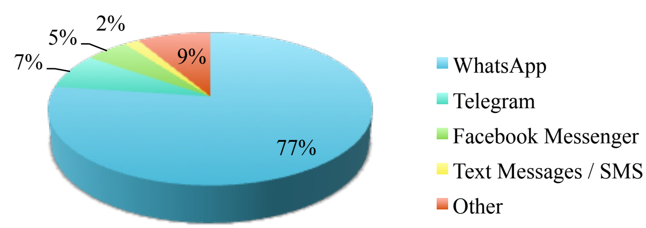

Figure 2. Messages per messenger service.

Attentiveness

Once a message has been triaged - i.e. the recipient is aware of the message and has decided on whether to act on it or not at the current time -, the receiver's response highly depends on factors that are hard to model, such as the sender-receiver relationship, or the importance and urgency of the content. Additionally, messages in itself may be interruptive even without responding to them. Hence, we focused on modeling attentiveness rather than responsiveness.

Attending a message can be done in three ways: (1) via the notification center, which shows the sender and the first part – sometimes the full content – of the message, (2) opening the corresponding messenger application, or (3) by reading the message on another device.

In our data set, 38,180 (68.4%) messages were first attended to by being checked through the notification center. 14,134 (25.3%) of the messages were first attended to on another device, while the remaining 3,510 (6.3%) messages were first attended to by opening the corresponding messenger application. Participants attended messages within a median time of 2.08 minutes. 25% of the messages were attended to within 12.0 seconds, 75% within 12.3 minutes, and 95% within 80.0 minutes.

A Kruskal Wallis test revealed significant differences in how fast messages were attended to depending on how this was done ($\mathcal{X}^2(2) = 2505.139, p < 0.001$). Pair-wise Bonferroni-corrected Mann-Whitney tests showed that messages were attended to faster through the app ($Mdn = 0.47$ min, $p < .001$) or another device ($Mdn = 0.75$ min, $p < .001$) than through the notification center ($Mdn = 3.2$ min).

Inferring Attentiveness from Phone-Usage Patterns

Only looking at received messages would result in a sparse sample set. Since we were interested in people's attentiveness for each minute of the day we filled these gaps by predicting attentiveness through a machine-learning model, which we previously described in [11]. The model uses 16 features derived from the mobile phone, namely: the screen status (on/off) and when it last changed, the status of the proximity sensor (screen covered / not covered) and when it last changed, the time since the phone was last unlocked, the time since the last message arrived, the number of pending messages, the time since the user last opened the notification center, the hour of the day, the day of the week, and the ringer mode.

We used the median delay (2.08 min.) between arrival and attending a message for classifying attentiveness: we labeled users *attentive* when they triaged messages within these 2.08 minutes, otherwise they were labeled as *non-attentive*. We evaluated different classifiers and achieved best results with *Random Forests*: 79.29% accuracy and $\kappa = .586$. Precision and recall for being attentive were .771 and .828 respectively.

Attentiveness Throughout the Day

We used this model to computationally estimate the times that people were attentive throughout the day. Therefore, we stepwise iterated through all sensors and computed the state of each of the features for the beginning of each minute of the day, which resulted in 86,400 states per day. For each of these states, we then ran the classifier and predicted the participant's attentiveness.

On average, participants were predicted to be *attentive* to messages 50.5% ($SD = 14.6\%$) of the full 24-hours of the day. The quartiles were 40% ($1^{st}Q$), 49% (median), and 55% ($3^{rd}Q$).

Hence, for 12.1 hours per day or 84.8 hours per week, the majority of the participants attended to messages within 2 minutes after arrival. This corresponds to 75.8% of the hours typically spent awake, assuming an average of 8 hours of sleep.

Figure 3 shows the average attentiveness during the seven days of the week. There was a significant difference ($F(6, 20802) = 41.07, p < .001$) between the days of the week: a series of Bonferroni-corrected pair-wise t-tests showed that there were statistically significant differences between the weekdays and the days of the weekend (level of significance at least $p < .001$): participants were significantly more attentive to messages during the week (Mon - Fri) than during the weekend (Sat, Sun). During the week, participants were predicted to be attentive 62% to 67% of the day, whereas on the weekend these numbers dropped to 50% and 45% .

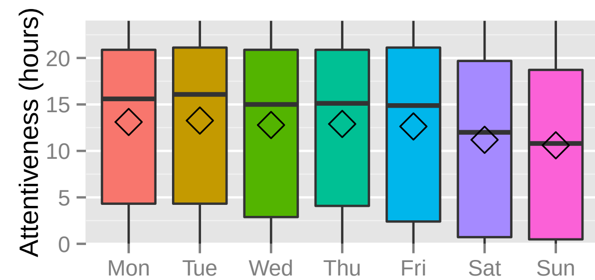

Figure 3. Average attentiveness by day. The dash indicates the median, the diamond the mean level of attentiveness. A value of *e.g.* 12 indicates that, on average, users are attentive to messages during 12 hours of the given day.

Figure 4 shows the average attentiveness during the hours of the day. The median predicted attentiveness ranges from 0% at 4:00 to a maximum of 83% at 21:00 o'clock. Here we found significant differences ($F(23, 20875) = 189.6, p < .001$) as well:

Attentiveness was highest during the evening, that is, between 18:00 and 21:00 o'clock. With a median attentiveness of at least 80%, it was significantly higher than during the rest of the day (all pair-wise comparisons at least $p < .01$, Bonferroni-corrected).

Further, we found a statistically significant difference between night time (0:00 - 8:00 o'clock) and day time (10:00 - 23:00 o'clock) (all pair-wise comparisons at least $p < .001$, Bonferroni-corrected). As expected, during nights and early mornings, median attentiveness was always below 50%, whereas during the day, median attentiveness was always above 67%.

Next, we analyzed the durations where participants were predicted not to be attentive to messages. Computing the quartiles, participants were predicted to be in an attentive state again after 1, 2, and 5 minutes, in 25%, 50%, and 75% of the cases respectively. This indicates that most of the time when entering a state of not being

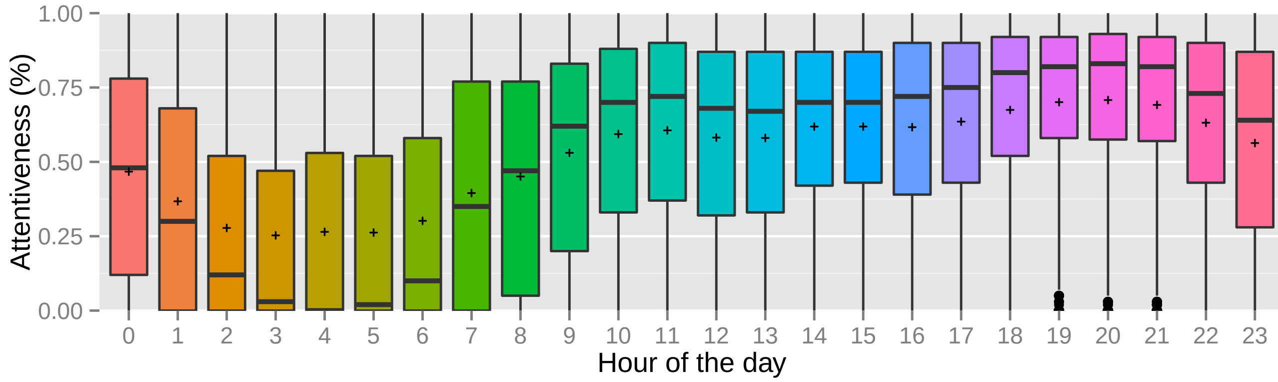

Figure 4. Average attentiveness by hour. The dash indicates the median, the cross the mean level of attentiveness. A value of .50 indicates that, on average, users would attend to messages within 2 minutes in 50% of the cases at a given hour.

attentive to messages, participants returned to a state of attentiveness after a few minutes.

DISCUSSION

In summary, our data shows that people are attentive to messages 12.1 hours of the day, attentiveness is higher during the week than on the weekend, people are more attentive during the evening, and when being inattentive, people return to attentive states within 1-5 minutes in the majority (75% quantile) of the cases.

Avrahami and Hudson [1] developed statistical models to predict users' responsiveness to incoming messages and the likelihood of receivers responding to messages within a certain time period. Although they used some of the same features, their algorithm was limited to Desktop usage not taking into account people's messaging behavior while being mobile and throughout the entire day. Other previous work has quantified the time between the arrival and acting on messages / notifications with 6 minutes (average) for replying to SMS [2], 6 minutes (median) for attending to messages [11], and 30 seconds (median) until a notifications is clicked (if it is clicked) [13]. With 2 minutes median delay until attending to messages, our work is in line with these findings, and stresses that people usually attend to messages promptly. However, previous studies only report measures of central tendency. Our study advances these findings by providing an estimation of attentiveness throughout the day - hence providing insights into *when* people are more or less attentive and *how long* phases of inattentiveness last.

Our findings of 12 hours per day exceed findings by Dey *et al.* [5], who reported in 2011 that, when phones were turned on, users kept them within arm's reach 53% and in the same room for 88% of the time. Our findings show that keeping the phone close also translates to attending to new messages promptly for large parts of the day.

As pointed out by Church and de Oliveira [3], people have high expectations towards the responsiveness of their conversation partners in mobile messaging. Hence, strategies to deliver notifications in opportune moments [7, 8, 10, 12], like *bounded-deferral*, which may delay the delivery of messages, can only work without violating expectations if there is a sufficient number such moments. Our findings suggest that this is the case: our participants were attentive for large parts of their awake time and phases of estimated non-attentiveness during daytime typically lasted for only 1-5 minutes. Hence, in most cases, there will be enough opportune moments sufficiently stacked together. However, whether these moments are truly *opportune moments*, or whether people just give their phones priority over other activities, such as meetings or being out with friends, remains an open question.

The main limitation of the reported results is the fact that the analysis is based on *predicted* attentiveness. This may cause our model to over-represent instances of activity, where people are typically not active, which might, *e.g.*, result in overly high predictions of attentiveness during nighttime. Further, as the model might alternate between predictions, it may cause our analysis to underestimate the duration of non-attentiveness.

CONCLUSIONS

This work provides quantitative evidence regarding mobile phone users' attentiveness to messages. Our results show that people are highly attentive to messages during 73.5% of their wake hours and that phases of inattentiveness typically only last for a few minutes.

For designers, this means that any method for intelligent notification delivery can assume a generally high level of attentiveness of mobile phone users. Hence, the concept of *bounded deferral* would work well in 75% of the cases with a bound of 5 minutes.

Future work needs to focus on gaining a better understanding of the underlying causes, such as social pressure and positive reward loops, and on how we can overcome these to create spaces where phone users can happily retreat to feeling free to think and reflect.

ACKNOWLEDGMENTS
We thank the participants of our study.

REFERENCES
1. Avrahami, D., and Hudson, S. E. Responsiveness in instant messaging: predictive models supporting inter-personal communication. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, ACM (2006), 731–740.
2. Battestini, A., Setlur, V., and Sohn, T. A large scale study of text-messaging use. In *Proc. MobileHCI '10*, ACM (2010), 229–238.
3. Church, K., and de Oliveira, R. What's up with whatsapp? comparing mobile instant messaging behaviors with traditional sms. In *Proc. MobileHCI '13*, ACM (2013).
4. Czerwinski, M., Horvitz, E., and Wilhite, S. A diary study of task switching and interruptions. In *Proc. CHI '04*, ACM (2004).
5. Dey, A. K., Wac, K., Ferreira, D., Tassini, K., Hong, J.-H., and Ramos, J. Getting closer: an empirical investigation of the proximity of user to their smart phones. In *Proc. UbiComp '11*, ACM (2011).
6. Fischer, J. E., Greenhalgh, C., and Benford, S. Investigating episodes of mobile phone activity as indicators of opportune moments to deliver notifications. In *Proc. MobileHCI '11*, ACM (2011).
7. Horvitz, E., Apacible, J., and Subramani, M. Balancing awareness and interruption: investigation of notification deferral policies. In *Proc. UM '05*, Springer-Verlag (2005).
8. Iqbal, S. T., and Bailey, B. P. Effects of intelligent notification management on users and their tasks. In *Proc. CHI '08*, ACM (2008).
9. Lenhart, A. Teens, smartphones & texting. Tech. rep., PewResearch Internet Project, 2012.
10. Pejovic, V., and Musolesi, M. Interruptme: Designing intelligent prompting mechanisms for pervasive applications. In *Proc. UbiComp '14*, ACM (2014).
11. Pielot, M., Kwak, H., de Oliveira, R., and Oliver, N. Didn't you see my message? predicting attentiveness in mobile instant messaging. In *Proc. CHI '14* (2014).
12. Poppinga, B., Heuten, W., and Boll, S. Sensor-based identification of opportune moments for triggering notifications. *Pervasive Computing 13*, 1 (2014), 22–29.
13. Sahami Shirazi, A., Henze, N., Dingler, T., Pielot, M., Weber, D., and Schmidt, A. Large-scale assessment of mobile notifications. In *Proc. CHI '14*, ACM (2014).