InterruptMe

Designing Intelligent Prompting Mechanism for Pervasive Applications
Veljko Pejovic, Mirco Musolesi

Presented by: J.W. Joo, M.E. Wijaya, Y.A. Winata
THE INTERRUPTME LIBRARY

• Smartphones are naturally suited for self-contained interruptibility inference:
  • sensing, learning and notification can all happen on the same device.

• Compared to inference performed on a remote server, locally executed inference algorithms can keep raw sensor data on the device
  • thus reduce data usage, energy consumption, and privacy concerns.

• the implementation of an interruption mechanism for smartphones is not straight-forward due to the inherent limitations of smartphones and peculiarities of mobile platforms.
THE INTERRUPTME LIBRARY

- **InterruptMe**
  - open-source Android library that enables identification of opportune moments for interruption.

- The library allows the overlaying application to be notified of such moments.

- **The Interruption Manager**
  - the core of the library holds a model of interruptibility and exposes a publish-subscribe API that notifies the application about opportune moments for interruption.

- To recognize these moments Interruption Manager maintains a model of interruptibility in the form of a classifier.

- **InterruptMe** periodically senses the context using a third party sensing library, and queries the classifier with the sensed data.

- The manager allows the application to provide feedback about the outcome of the interruption.
  - The feedback can include the user reaction or sentiment to the interruption.

- To recognize “very good” moments to interrupt...
THE INTERRUPTME LIBRARY

• InterruptMe
  • Designed to be a light-weight library for intelligent notification management.
  • Library consists of 926 lines of Java code.
  • The general purpose ML library it uses consists of 954 lines of code.
  • Android Nexus 4 phone accounts to 13.195 MB

• The general goal
  • To support the evolution of a common model for multiple users, and subsequent personalization of each of the model’s copies.
EXPERIMENTAL EVALUATION

• To evaluate the ability of InterruptMe
  • we embed the library within the SampleMe application.
• The library communicates with a notification manager and decides on the notification delivery timing.
• InterruptMe evolves a personalized interruptibility classifier on the mobile.
• We ran a month long trial of SampleMe with 10 subjects,
  • eight male and two female, aged between 22 and 26, all graduate students.
• The subjects were told to use their phones as usual during the experiment.
• To compare the utility of InterruptMe managed notifications with those that were received at random moments we perform an N of 1 randomized trial
• In the original SampleMe run we noticed that individual reactions to notifications differ drastically.
• The N of 1 trial helps us avoid the bias in response reactions that would skew the results if we were to employ a test method where the users are split into two groups according to the notification policy.
EXPERIMENTAL EVALUATION

• Each phone starts with a clean slate model of interruptibility that is trained every time a user reacts to a notification.
• In this trial, our goal is to identify opportune moments for interruption as defined by a user’s sentiment.
• Train the classifier with a positively labelled instance in case a user states that the moment is “very good” to interrupt.
  • otherwise, we train the classifier with a negative label.
  • In addition, non-answered notifications are labelled as non-opportune moments for interruption once a new notification comes in.
EXPERIMENTAL EVALUATION

- Set a minimum interval of 10 minutes between any two successive notifications.
- Limit the maximum number of notifications per day to 10.
- Activate random moment notifications with a probability that leads to six expected notifications per day.
- A total of 1285 notifications which 763 resulted in a completed survey.
- The distribution of notifications per user is shown in Figure 6.
- Allow three weeks for the InterruptMe classifier to be trained.
  - perform the analysis on the data coming from the last week of the experiment.

![Figure 6. Per-user distribution of answered and unanswered notifications in the N of 1 trial.](image-url)
EXPERIMENTAL EVALUATION

Responsiveness

• First examine the time users took to react to a notification
• In Figure 7 show a CDF of reaction times, for answered surveys of all the users, for both methods of notification.
• InterruptMe-based notifications result in the response time of 12 minutes on the average.
  • random notifications on the average take 22 minutes to respond to.

![Figure 7. CDF of time to respond to a notification.](image)
• InterruptMe-based notifications result
  • The response time of 12 minutes on the average
  • Random notifications on the average take 22 minutes to respond to.
• The distribution also shows a non-negligible portion of surveys that are answered after a long delay.
  • Such delayed answers can lead to a false impression
• Concentrate on surveys that were answered no later than ten minutes after the corresponding notification was received
• Ten minute is also the minimum interval between consecutive interruptions in our experiment.
  • restricting the maximum response delay to this value ensure
  • all notifications are treated as independent samples of interruptibility and are not directly affecting (overwriting) each other.
EXPERIMENTAL EVALUATION

Sentiment towards interruption moments

- Trained to recognize moments that users labelled as “very good” to interrupt at.
- When considering the correctness of the classifier we have to take into account that an answer to the question about current interruptibility is given on a four-point Likert scale.
  - Thus, higher the user rated the moment, the better the classifier is.
- In the application recorded the sentiment as an integer taking its value from zero, indicating a “not at all” good, to three, indicating a “very good” moment to interrupt.
EXPERIMENTAL EVALUATION

- In Figure 8(a) show a CDF of the minimum user reported sentiment towards the interruption moment.

- InterruptMe-based answered notifications are more favorably received.
  - with 26.4% of them being marked as “very good” moment to interrupt, compared to 15.4% for randomly scheduled notifications.

- The difference diminishes as we relax the requirement for the minimum sentiment value, as seen by a higher portion of random messages received with the minimum sentiment value of “a little” interruptible or better.
Amount of Interruptibility

• Interruptibility is often associated with user’s frustration.
  • In this experiment we did not explicitly ask users about their frustration and annoyance.
• We hypothesize that the recent exposure to interruptions determines user frustration, therefore, sentiment towards interruptions in the experiment.

• In Figure 9 we plot the mean sentiment for an interruption moment versus the number of interruptions received in a two hour time period that preceded the moment.
• The figure shows that isolated notifications tend to be more favorable than the ones received after a large number of recent notifications.

Figure 9. Mean sentiment towards an interruption versus number of interruptions in the preceding two-hour period for all the notifications sent during the experiment. The sentiment towards the last received notification falls off as the user copes with an increasing number of interruptions in a limited preceding time period. A notification can be preceded by at most nine other notifications in a day. A closer investigation reveals that when preceded by nine other notifications in a two-hour window the notification often lingers for a substantial time before being answered, therefore, acting like an isolated notification.
Amount of Interruptibility

- shown in Figure 9, exhaust the interruptibility of the user, which manifests through a lower reported sentiment.
  - Note that the same “grouping” effect is unlikely to happen with context-oblivious random interruptions.

Figure 9. Mean sentiment towards an interruption versus number of interruptions in the preceding two-hour period for all the notifications sent during the experiment. The sentiment towards the last received notification falls off as the user copes with an increasing number of interruptions in a limited preceding time period. A notification can be preceded by at most nine other notifications in a day. A closer investigation reveals that when preceded by nine other notifications in a two-hour window the notification often lingers for a substantial time before being answered, therefore, acting like an isolated notification.
Amount of Interruptibility

- To evaluate the ability of InterruptMe to recognize opportune moments.
  - control for the effect of the amount of user interruptibility
  - restrict the dataset to notifications that were delivered in isolation
  - Figure 8(b) shows the CDF of user-reported interruptibility in this case.
  - Compared to Figure 8(a), the gap between InterruptMe-based and random notifications is even larger
  - The median sentiment towards InterruptMe-selected opportune moments is “somewhat” suitable for interruption, whereas it is “a little” suitable in the case of randomly selected moments.

Figure 8. CDF of minimum reported sentiment.
DISCUSSION AND LIMITATIONS

• Investigate it by observing different outcomes of an interruption, and capturing a range of contextual features.

• The complexity of the problem necessitates that certain views on the interruptibility were not considered in the study design.
  • InterruptMe was designed to recognize if an opportune moment for interruption has come.

• While we indeed get a higher overall sentiment and a faster response to mobile notifications scheduled through InterruptMe.
  • also identify the frequency of interruption as a key parameter influencing a user sentiment towards interruption.
DISCUSSION AND LIMITATIONS

• While we have no means of measuring the main task performance in non-controlled study.
  • hypothesize that the awareness of the performance degradation leads to an increased frustration and lower reported sentiment.
• In this study predominantly motivated.
  • a smoking cessation digital behavioral change intervention (dBCI) must deliver a potentially annoying content.
• In this study predominantly motivated with the case where the content of the message cannot be adjusted to the user’s interest
  • design study to measure responsiveness – a user’s feedback reaction to a notification.
CONCLUSIONS

• Implementation and evaluation of InterruptMe
  • a smartphone library that empowers an overlying application with personalized, evolving intelligent interruption models.

• The systems design point of view
  • a simple resource efficient online learner can serve as a basis for recognizing opportune moments for interruption.
  • library represents a good starting point for identifying opportune moments for interruption.