Design of a Mobile App for Interspeech Conferences: Towards an Open Tool for the Spoken Language Community

Robert Schleicher, Tilo Westermann, Jinjin Li, Moritz Lawitschka, Benjamin Mateev, Ralf Reichmuth, Sebastian Möller

Quality and Usability Lab, Telekom Innovation Laboratories, TU Berlin, Germany
robert.schleicher@tu-berlin.de, kimljji@hotmail.com, moritz.lawitschka@gmail.com, benmateev@gmail.com, ralf.reichmuth@telekom.de, tilo.westermann@telekom.de, sebastian.moeller@telekom.de

Abstract

In this paper, we describe the development of a mobile phone app for supporting participants and organizers of Interspeech conferences. Based on a survey amongst future organizers and attendees, we identified the most relevant functionalities and implemented an initial set of them on two popular platforms, iOS and Android. The app is meant as an open tool to be handled under the auspices of ISCA, and can be extended with speech and language functionalities in the future. This way, we hope to turn it into a community platform which can be used for experimenting with new speech technologies on site. A first version of this app will be presented at Interspeech 2013.

Index Terms: mobile phone app, Interspeech conference, community platform.

1. Motivation

 Whereas speech and language have not been the primary interaction modalities for mobile phone apps in the past, their role is slowly increasing. This is favored by principled advantages of speech on small devices, as well as by mobile contexts where touch interaction is sometimes not viable (e.g. in parallel-task situations). Nevertheless, there is still a lack of knowledge about where and why users prefer speech over other interaction modalities, and how apps have to be designed to take full advantage of the speech communication channel.

One particular situation where speech-based interaction with mobile apps might be advantageous is during visits of conferences such as Interspeech. Although it can be observed that attendees of Interspeech conferences make use of speech and language interaction on mobile devices (e.g. with Apple’s Siri, voice dialing, translation), there is still no mobile app which would cater for the Interspeech conference in a global way. This is in contrast to other conferences in the area of human-computer interaction (HCI), which already have some history of using mobile apps for supporting attendees of the conference (e.g. CHI, UIST).

It is our aim to develop and implement an app for the Interspeech conference series, and for the spoken language community at large. We consider this to be a task which will span over several years, and which should involve a considerable part of the scientific community, so that the community as a whole can take profit of such a development. For Interspeech participants, such an app could inform about sessions and talks, about conference participants and how to contact them, and about the location of presentation rooms; it could give access to abstracts and papers for discussion during the conference; it could support travel to the conference site; and it could also facilitate the registration procedure. For Interspeech organizers, it may serve as a communication means for updating participants about latest developments; it could track participants and thus provide information about strongly and less frequented sessions, so as to guide room allocation for future organizers.

In this paper, we describe the steps we have taken so far and will give an outlook on further developments as well as research questions that arise in the context of this app.

2. Determining App Functionalities

Various conference apps already exist. We compared available offers for conferences in our and related domains like CHI 2012, UIST ’12, IEEE Transformers Committee Mobile Application etc. in order to get an overview on possible functionalities such an app should provide. This list of functionalities was extended by our own ideas.

This final list was then presented as an online survey to members of our lab to see what features users demand. Altogether 21 participants who all owned a smartphone (9 iOS, 8 Android, 2 Windows, 2 preferred not to say) and had conference experience rated the importance of 22 possible functionalities on a five-point Likert scale. Table 1 shows a summary of the results for a selection of features sorted by average rating.

Table 1. Exemplary functionality ratings ranging from +2 meaning "very important" to -2 for "unimportant".

<table>
<thead>
<tr>
<th>Feature</th>
<th>Mean rating ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indoor plan</td>
<td>1.7 ± .5</td>
</tr>
<tr>
<td>Personal program list</td>
<td>1.6 ± .8</td>
</tr>
<tr>
<td>Abstracts</td>
<td>1.5 ± .8</td>
</tr>
<tr>
<td>Web-interface</td>
<td>1.1 ± .9</td>
</tr>
<tr>
<td>Proceedings online access</td>
<td>1.0 ± 1.1</td>
</tr>
<tr>
<td>Info transportation</td>
<td>1.0 ± 1.0</td>
</tr>
<tr>
<td>Info accommodation, restaurants</td>
<td>0.3 ± 1.1</td>
</tr>
<tr>
<td>Integration Twitter</td>
<td>-1.2 ± 1.2</td>
</tr>
<tr>
<td>Integration Facebook</td>
<td>-1.4 ± 1.0</td>
</tr>
<tr>
<td>In-app SMS</td>
<td>-1.5 ± .7</td>
</tr>
</tbody>
</table>

Obviously, the main focus for potential users is on obtaining the information that a conference attendee otherwise gets via

1 https://itunes.apple.com/de/app/chi-2012/id522222838
2 https://itunes.apple.com/de/app/uist12/id566539586
3 https://itunes.apple.com/de/app/ieee-transformers-committee/id426452584
the conference program and related documents through the app. Doubling functionalities within the app which are already available on the mobile phone anyhow like SMS services is clearly not desired. Surprising to us, there was also little interest in linkage to social media like Twitter or Facebook. This was backed up by the replies to additional questions in the survey where subjects could specify which functionalities they would explicitly prefer not to have in the app. Thus we decided to exclude functions that aim at enhancing social interactions during conferences (for examples see [1][2]) for the time being.

3. Implementation

The app is intended to be available on iOS as well as for Android. While the design guidelines for both operating systems vary slightly, this mainly affects the look and feel of the app, i.e. the graphical user interface (GUI). The interaction flow will be identical for both, and also the backend, where all the relevant information is stored, will not differ.

3.1. Graphical User Interface

Based on our results, we decided to group the functionalities in various categories which should be easily accessible to the user. A first outline is given in Figure 1.

![Figure 1: Main menu of the planned app (iOS version) offering the most important functionalities at a glance.](image)

The Agenda allows viewing the general conference program, and from there navigating to individual abstracts for all users that downloaded the app, as well as full paper texts for those who registered for the conference. Individual talks/sessions as abstracts can be added to the calendar application of the phone. People will give an overview on participants and their conference contributions. The Search functionality can be used to search for authors or keywords in the conference program. Information will provide all additional information that is relevant for conference attendees at least in the form of the accompanying conference web page. The Map tab will serve for spatial navigation throughout the conference. As indoor navigation is still a challenge for mobile phones [3], we will restrict it to static floor plan images for 2013, and extend it for future events.

3.2. Backend

Administrating Interspeech contributions and attendees is done via the START V2 ConferenceManager¹ (START) platform. Throughout the years, a certain format for naming accepted submissions and organizing them into sessions emerged, which our backend is based on. In order to serve the necessary information to our apps we implemented an API that imports the information directly from START. No additional work from the conference committee is required to provide a distinct app for their conference.

Our API follows the REST design principles and is thus using the de facto standard for web APIs. Built with the Ruby on Rails framework for web applications the API is easily extensible to support formats from other conference administration platforms. Albeit not necessary in the first stages of the app (i.e. for Interspeech 2013), the API already supports handling multiple conferences at once. This could also be useful for a future app version that may serve information for more than one conference. Authentication is handled on a conference level, hence sensitive data is strictly separated from each other.

3.3. Research Aspects

The app described so far is mostly a service for conference attendees in general with functionalities comparable to other aforementioned services, e.g. [2], and neither addresses the speech community nor research on human computer interaction in particular. However, we intend to consider both aspects in the implementation of the app:

3.3.1. Speech Technology

The open app platform can also be used as a testbed for emerging speech and language technologies, showing how they can be applied in mobile contexts. Thus, we expect to provide APIs for integrating automatic speech recognition (ASR) and text-to-speech (TTS) systems, and perhaps also translation systems. Using such APIs, experimentation with speech and language technology under realistic conditions could be easily done during Interspeech events. Researchers in multilingual speech processing could take particular profit of the multilingual composition of Interspeech attendees.

3.3.2. Data logging

As described in the beginning, it is yet not fully understood, when users use a particular modality to interact with their mobile device, partly due to the fact that this can only be examined to a limited extent in a mere laboratory context, whereas research in the wild quickly touches sensitive topics of data privacy. We intend to not only offer an app to the Interspeech attendees, but also learn something about their usage patterns in that context. In order to avoid any issue of data privacy, we completely unveil what data we are logging for our non-commercial research purposes, and put this up for discussion. The following data is supposed to be logged:

- **App session information**: time (i.e. start and end timestamps), screen flow (i.e. screens and duration per screen), location, modality.
- **Device information**: platform, model, OS version, locale, country, screen resolution, app version.

4. Conclusions & Outlook

This text served as an initial overview of what will be available as a first prototype with restricted functionalities for the Interspeech 2013. Whereas we cannot report on initial usage during the conference, we will open up some questions to be discussed during the event, in order to turn the app into a worthy tool for the community, which should be handled under the auspices of ISCA, and will be available to the entire research community under an Open Source Software paradigm.

¹ http://www.softconf.com
5. References

