A Field Study on the Usability of a Nearby Search App for Finding and Exploring Places and Events

Florian Knip, Christian Bikar, Bernd Pfister, Bernd Opitz, Timo Sztyler, Michael Jess
University of Mannheim, Germany
{fknip, cbikar, bpfister, jopitz, tsztyler, mijess}@mail.uni-mannheim.de

Ansgar Scherp
Leibniz Information Center for Economics, Kiel;
Kiel University and University of Mannheim,
Germany
a.scherp@zbw.eu

ABSTRACT
Commercial apps for nearby search on mobile phones such as Qype, AroundMe, Foursquare, or Wikitude have gained huge popularity among smartphone users. Understanding the way how people use and interact with such applications is fundamental for improving the functionality and the user interface design. In our two-step field study, we developed and evaluated mobEx, a mobile app for faceted exploration of social media data on Android phones. mobEx unifies the data sources of related commercial applications in the market by retrieving information from various providers. The goal of our study was to find out, if the subjects understood the metaphor of a time-wheel as novel user interface feature for finding and exploring places and events and how they use it. In addition, mobEx offers a grid-based navigation menu and a list-based navigation menu for exploring the data. Here, we were interested in gaining some qualitative insights about which type of navigation approach the users prefer when they can choose between them. We have collected qualitative user feedback via questionnaires. We also conducted a quantitative user study, where we evaluated user-generated logging data over a period of three weeks with a group of 18 participants. Our results show that the time-wheel can serve as an intuitive way to explore time-dependent resources such as events. In addition, it seems that the grid-based navigation approach is the preferable choice when exploring large spaces of faceted data.

Keywords
field study, mobile exploration, faceted search

1. INTRODUCTION
Smartphone applications (short: apps) for nearby search of points of interests (POI) have gained huge popularity among users. Commercial apps such as Qype[1], AroundMe[2], Foursquare[3], or Wikitude[4] retrieve data from either their own databases or external social media databases and display it in form of small icons on a map. In the map view, the user's current position is displayed and helps to easily explore restaurants, bars, or other nearby POIs. Typical use cases include figuring out the address and opening hours of restaurants or museums in town while being on-the-go with the mobile phone [11]. We developed mobEx (short for "mobile exploration"), a mobile application for Android phones that retrieves social media data from different data providers such as Qype, Twitter[5], LastFM[6], Eventful[7], and Google Places[8] as well as other open web sources such as DBpedia[9] and OpenPOI[10]. The data provided by these sources is organized along different, hierarchical facets, i. e., categories such as people, locations, organizations, and events as well as subcategories of these facets [1]. The facets are obtained from the meta-data fields of the retrieved entities and encourage users to search in an exploratory manner [7]. The mobile app mobEx offers unique search features that enable the user to find nearby entities along these facets. An entity-resolution-concept merges the received entities in real-time, i. e., when a request is issued by the user [8].

The research question addresses in this paper is to understand how people use and interact with applications such as mobEx in order to improve the functionality and the user interface (UI) design. For the exploration of time-dependent entities such as events or shops with opening hours, mobEx offers the UI widget of a time-wheel [9]. It allows the users to filter time-dependent entities according to his or her preferred time. Furthermore, mobEx offers two different approaches for faceted navigation. A classical list-based approach and an alternative grid-based approach. Both nav-

MUM '14, November 25 - 28 2014, Melbourne, VIC, Australia.

ACM 978-1-4503-3304-7/14/11...$15.00.
http://dx.doi.org/10.1145/2677972.2677992

[1] `http://qype.com`
[2] `http://www.aroundmeapp.com`
[3] `http://foursquare.com`
[4] `http://wikitude.com`
[5] `http://twitter.com`
[6] `http://lastfm.com`
[7] `http://eventful.com`
[8] `https://developers.google.com/places/`
[9] `http://dbpedia.org/`
[10] `http://openpois.net/`

igation approaches have been compared in an earlier user study [10], where each subject had only seen one of the two UI options. The result of the study is that the grid-based approach requires significantly more clicks and more time, but has a higher user satisfaction. In contrast, the goal of the study reported in this paper is to provide answers to the following two research questions:

1. Do the subjects understand the metaphor of the novel time-wheel and how do they use it?
2. What type of navigation approach for faceted search do the users prefer when they can choose between a grid-based navigation menu and a list-based navigation menu?

In order to address these research questions, we designed a two-step field study [6]: In the first step, we conducted a qualitative user study with a beta version of mobEx. This qualitative study served as feedback to improve the UI design. The qualitative user study was followed by another implementation phase where we added additional features suggested by the users and at the same time improved existing features. Changes included for example modifying the size of the time-wheel or displaying the weekday when adjusting the time. After performing the changes, we conducted in the second step a quantitative study with a final version of mobEx. Here, we collected and analyzed user-generated logging data of 18 subjects over a period of three weeks. In summary, the results of the user study show that the grid-based navigation approach seems to be the preferable choice when exploring large spaces of faceted data. In addition, the time-wheel can serve as an intuitive way for exploring time-dependent entities such as events and places with opening hours. However, the usage rate of the new UI feature was not very high. Perhaps this might change if the application is used on more longer term and if more entities such as events and places with opening hours are available through the different integrated data sources.

The remainder of this paper is organized as follows: In the following section, we will introduce mobEx along with its UI features. In Section 3, we will give an overview of the two-step field study, the subjects, and procedure. In Sections 4 and 5, the results of the two steps, i. e., the qualitative user study and subsequently the quantitative user study will be presented in detail. In Section 6, we examine work that is related to our evaluation method as well as other applications that focus on faceted search, before we conclude the paper.

2. INVESTIGATED OBJECT: MOBEX

The mobile application mobEx was designed to address everyday search scenarios while being on-the-go [11]. The application encourages the user to explore the surrounding area by providing information about people, places, events, and organizations. For instance, one might be in an unknown area or city for a business trip or on vacation and looking for upcoming events during the time of the stay. Even in a familiar environment, one might look for a bakery in one's town that opens on Sundays or a grocery store that is still open 30 minutes before midnight. All these everyday scenarios have both a location-dependent and a time-dependent information component, i. e., finding entities at a certain place and/or certain time. With mobEx, we addresses these mobile search problems by retrieving data from different web sources according to the user's current location and display the retrieved entities both in a result list and as icons on a map. In addition, mobEx offers special UI features such as a time-wheel for the exploration of time-dependent entities and a faceted navigation menu for exploring the retrieved entities along different hierarchically organized categories.

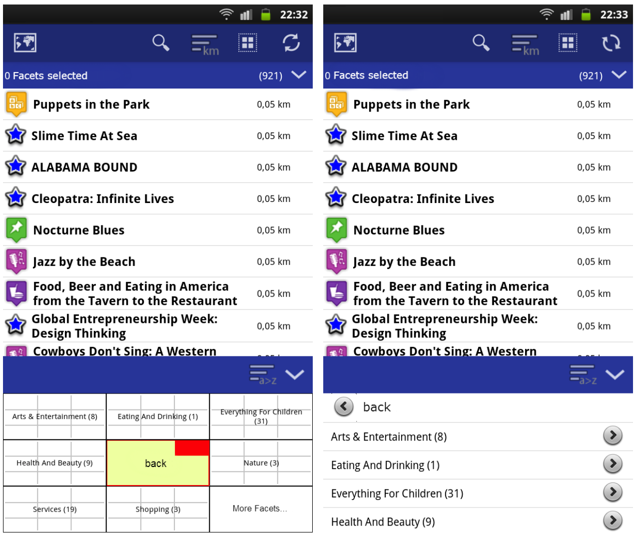
The UI of mobEx can be divided into two major parts: The first one is the facet view shown in Figure 1, where the sources can be filtered and searched according to categories or search terms. The second one is the map view depicted in Figure 2, where the web resources are visualized on the map. The following sections describe the mobEx application and its features in detail. We start with the two types of navigation menus, followed by a description of the novel time-wheel UI widget.

2.1 Navigation Type

The mobEx navigation menu is inspired by FaThumb, a search application for large data sets on mobile phones [5]. In contrast to FaThumb, which was designed for phones with a physical numerical pad, mobEx makes use of today's touchscreen technology and mobile web access. The mobEx application provides two different kinds of navigation menus that both display the same content, but offer a different layout and navigation functions. The navigation types are a grid menu and a list menu as shown in Figure 1. In both cases, the navigation menu takes the bottom third of the space on the touchscreen. The details of the list navigation and grid navigation for faceted search are described below.

When the user runs the app for the first time, the navigation type is assigned randomly with a 50% probability for each navigation type. This ensured an initially equal distribution of both navigation types amongst all installations. To inform the user about the alternative navigation, the application gives short hints at the initial start and also during the usage of mobEx via non-intruding speech bubbles. Via a simple switch button in the settings menu, the user can choose between the grid navigation menu and the list navigation menu.

Grid Navigation. The grid navigation menu (Figure 1a) visualizes the facets in form of rectangles and arranges them as 3x3 matrix with the middle one as the back button. A rectangle contains the name of the facet aside with the amount of retrieved entities in brackets. If there are more than seven facets, an additional field on the bottom right corner is shown entitled "More facets". It allows the users to reach a screen with further facets. The result list shows the entities of the selected facets, i. e., those entities that belong to the selected categories. For example, when selecting the facet "Shopping", only entities that are under this facet are displayed. If no facet is selected, all entities are shown. The result list can be filtered by activating facets via long-press on a facet. With short-press on a facet, the user can navigate deeper into the facet structure and ex-



(a) Grid navigation menu (b) List navigation menu

Figure 1: Comparison of grid vs. list navigation

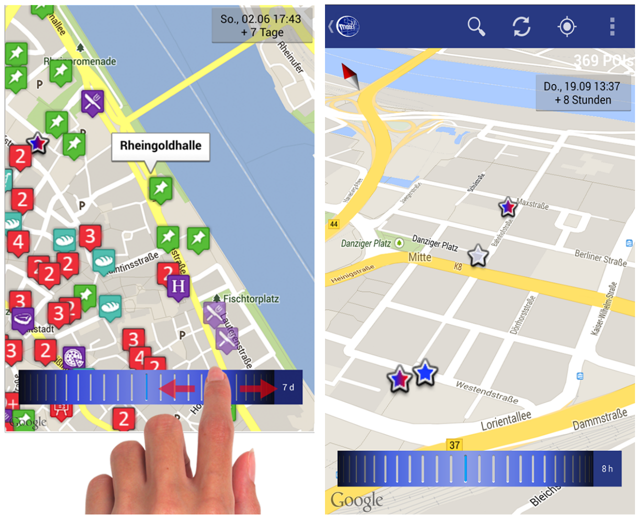
plore its sub-facets. To select facets without sub-facets, a short press is sufficient. Facets can be selected and unselected in an arbitrary combination. By default, all facets are unselected.

A characteristics of the grid navigation is the colored rectangles, which are integrated into the back button. These rectangles work like a breadcrumb trail and allow to keep track of the current navigation path within the facet structure. For example, as shown in Figure 1a, the red border indicates that the current position is one layer beneath the top layer and the red rectangle at the top right states that on the first level, the top right facet was selected.

List Navigation. In comparison to the grid menu, the list navigation menu (Figure 1b) visualizes the facets as entries in a scrollable list. The gray arrow at the right of each entry indicates that the facet has at least one more sub-facet. The back-button is at the top of the list. Depending on the screen size, the user sees a couple of facets and can scroll down the list to see the rest. In contrast to the grid menu, the list menu does not provide any information about the current position in the facet structure.

2.2 Time-wheel

The time-wheel allows the users to filter events and other time-dependent POIs such as locations with opening hours by time as shown in Figure 2a. While the users spin the time-wheel, the selected time window changes and the events on the map appear or disappear accordingly. Events and other POIs such as shops with opening hours within the chosen time frame are displayed and events/POIs outside of the time window are shown transparent as depicted in Figure 2b. The larger the time distance, the lower the opacity of the icons. Turning the time-wheel to the right shifts the time interval towards the future while turning left shifts it backwards to the present. The users can also adjust the time interval (i. e., the period in which events will be shown) via



(a) Exploring time-dependent entities using the time-wheel (b) Spinning the time-wheel changes the opacity of the icons

Figure 2: Spin the time-wheel to search for events and entities with opening hours within a specific time frame

the switch on the right edge of the time-wheel widget. Here, the user can choose intervals ranging from single hours up to several weeks. With these interval sizes, we aim at covering typical use cases such as "Looking for today's nightlife events" (interval of few hours needed) or "Planning a two week vacation in New York" (longer intervals needed). The current date and time including the current duration of the interval is displayed in the upper right corner as shown in Figure 2b.

3. OVERVIEW OF THE USER STUDIES

We conducted a qualitative and quantitative user study for evaluating the usage of mobEx. The subjects and the overall process of the two studies are explained in this section.

Subjects. Eighteen subjects (six female) used mobEx over the entire evaluation period, i. e., both evaluation parts. All subjects were required to own an Android phone with the version 2.3.3 Gingerbread or higher. The ages ranged from 18 to 31 (M= 25.5, SD= 3.03). Seven subjects studied business informatics, five others studied other subjects. The other six subjects did an apprenticeship. Over 90% of the subjects were German natives. The subjects have been recruited by asking the family, friends, class, and flat mates of the development team. A personal reference between our subjects and the development team of mobEx was highly important in order to preserve a constant attendance during the entire evaluation time. We assumed that an individual trigger could maintain the motivation and enhance the involvement of each subject during the study more reliable compared to a group of non-personal committed subjects.

Procedure. The evaluation was planned over two steps. First, a qualitative user study was conducted followed by a quantitative user study. As shown in Figure 3, we started with the first, qualitative study at the end of April 2013 and finished this phase on May 5th, 2013. Within this time, we asked the subjects to use the mobEx application and try to understand how it works. The results of the qualitative user study were captured using a questionnaire and served as feedback at an early stage of the development process. We aggregated the feedback and prioritized the comments, before we started to implement the feedback.

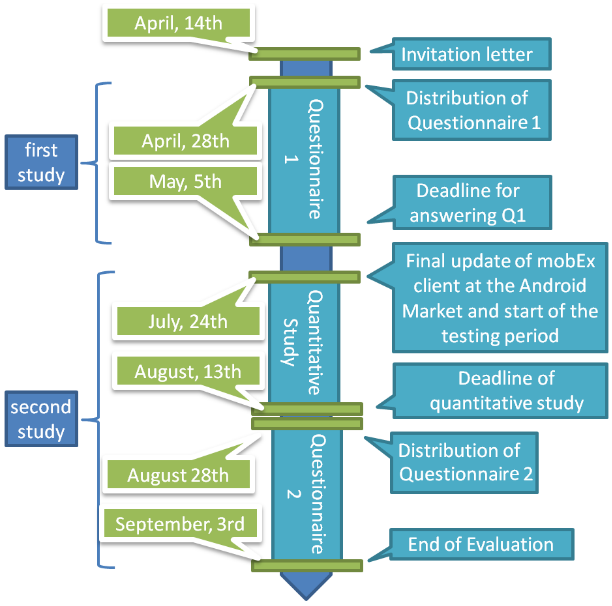


Figure 3: Overview of the time line of the two studies

Subsequently, we conducted the second step of the evaluation, the quantitative analysis. Within this part, we asked the subjects to use the app on a regular basis and to conduct tasks of their daily life. As it can be seen from Figure 3, we launched the second questionnaire, the last part of the second study, two weeks after the quantitative study was finished. Similar to the first questionnaire, this part was also completed within a period of one week and and marked the end of the entire evaluation period.

Data Set. The data set used for this study corresponds to the entities shown in the application. The client sends a query to a back-end server which is responsible for sending requests to the different data providers. By default, those providers are DBpedia, Eventful, Google Places, OpenPOI, Qype, KlickTel, Twitter, and LastFM. These providers can easily be changed in the settings menu by the user (see screenshot in Figure 4b). Before the result is send back to the client, the server uses entity-resolution technique to match similar entities from different data providers together [8]. The entities (e. g., events and places) are queried according to the user's location. Thus, the data set is different for each user except of the rare event that two users are situated at exactly the same location and at the same time.

4. STEP 1: QUALITATIVE USER STUDY

The goal of the first study was to find out if the users understood the application, what the users liked about mobEx, what they did not like, and what further improvements or changes they suggest. Here, the time-wheel was under special investigation, since it is a novel search tool which cannot be found elsewhere. However, we also addressed questions to the participants w.r.t. the two different approaches for faceted navigation.

4.1 Apparatus

The mobEx application was published on Google Play[11], which is the application distribution platform for Android. Google Play is usually pre-installed on every Android device and allows users to search, download, and install applications. The questionnaire used to capture the qualitative feedback was based on ISO-metrics [3]. It was as implemented using Google Docs[12]. Google Docs offers a function to compose questionnaires and to collect the answers in a spreadsheet.

4.2 Briefing of the Subjects

In the first study, starting from April 14th, we asked the subjects to download the first version of mobEx[13] from Google Play and use it over a period of one week to familiarize with the application. We contacted the subjects via email that briefly introduced the study and also contained a download link to mobEx. In addition to the official email, each subject was assigned to one of our team members which served as personal contact person. In an initial briefing, the contact person explained the basic idea and concept of mobEx, the overall structure of our study, and what we expected from the subject within the next couple of weeks. Furthermore, the contact person provided support in case of questions or technical problems during the evaluation.

4.3 Results of the Qualitative Questionnaire

First, we asked some general questions about mobEx: We started with whether the subjects would describe the app as innovative or not. This question was answered with more than 70% as innovative ("it is innovative" = thirteen subjects; "it is not innovative" = five subjects). The second question was about whether the subjects liked the layout of the app. Here, the subjects answered eleven times with "yes" and seven times with "no". The subjects were also asked to express their opinion about the functionality of the time-wheel and to describe the purpose of the time-wheel. The time-wheel was correctly described eleven times and wrongly described two times. The remaining five subjects did not have any particular opinion about the time-wheel.

Subsequently, we asked the subjects to rate a couple of questions on a 5-point-Likert scale. The 5-point-Likert-scale in the entire study means 1 for strongly agree (which is in all cases a positive statement) and 5 for strongly disagree (that is in all cases a negative statement). The satisfaction with the loading time ("Were you satisfied with the loading time?") was on average 2.4 with a standard deviation (SD) of 1.0. The mean of the satisfaction with the overall reactivity of the app ("Have you been satisfied with the reactiveness?") was 1.7 with a SD of 0.8. We also asked about

[11] https://play.google.com/store
[12] https://docs.google.com
[13] https://play.google.com/store/apps/details?id=de.unima.mobex.client

the usability of the time-wheel ("Was the time-wheel in the map mode helpful?"). Out of 18 subjects, this question was answered ten times with a value of 4 or worse (M= 3.4, SD= 1.2). Just one subject was totally satisfied with the usability of the time-wheel and three of our subjects rated the usability with 2. 50% of the subjects who rated the time-wheel negative (with a Likert-scale value of 4 or higher) did not fully understand the advantages of it. The question whether the subjects would recommend mobEx to friends was answered with an average of 2.7 and SD of 0.6. Thus, a little worse than the overall opinion about the app ("Generally, I like the app.") with a mean of 2.2 and SD of 0.6.

4.4 Open Feedback

In the first questionnaire, we also used open questions to obtain feedback about suggestions for further improvements. We received some constructive feedback from our subjects and analyzed it using a spread sheet, where we organized the feedback along the following topics: categorization of entities, general client issues, detail view, facet view, map view, server, and tweets. Afterwards, we aggregated the feedback how often a specific issue was mentioned. The result is the list of items shown below. The numbers behind some items indicate how often it was mentioned in the questionnaire. When there are no brackets behind an item, it means that we received this feedback exactly once. The list was discussed during a team meeting session. In this session, we prioritized each entry during and decided which comments should be addressed and which cannot be fixed during the course of our project. The (blue) colored items were not implemented either due to lack of time or because they were out of scope of our work. These are marked in the following with (b). The (green) colored items were not implemented because the subjects' feedback was too diverging. They are marked with (g). Finally, non-colored (black) issues were implemented prior to the quantitative study.

- Categorization of entities
  - Category names are useless, confusing or too abstract (10 times)
  - Breadth and depth of categories are inconsistent (g)
  - No consistent language (some elements mixed up German and English)
- General client issues
  - Loading the data takes too long (14 times)
  - Add a tutorial that explains the app
  - Offer an "offline" mode (b)
- Detail view of an entity
  - Do not show the entire URL, it is too long (2 times)
  - Dialog takes up the whole screen, it is not possible to click outside of the dialog to close it
- Facets view
  - Switching between facet view and map view is not intuitive
  - Did not find the "deselect all facets" button (better place it at the top) (4 times)
  - Use "standard icons" to show the entities' categories
  - Change the layout to reduce the thick action bar (which contains the currently selected facets) (g)
  - The "back" button in the middle of the grid navigation menu is confusing (2 times) (g)
- Map view
  - When clicking on an aggregated entity that contains multiple events and/or POIs, the map zooms in instead of showing the list of entities (2 times)
  - The time-wheel is too huge and insensitive to select different dates (9 times)
  - Better show the weekday when time is set
  - Time should be shown as dd:mm:yy hh:mm + xx minutes | hours | days | ...
- Server
  - The data received is not complete (2 times) (b)
  - Include POIs from open street maps, they are often better than Google Maps
  - Opening hours are not available for many restaurants (2 times) (b)
  - Show the menu of restaurants (b)
  - More events should be shown like cinema, cultural events, or others (b)
- Tweets
  - Show only Tweets from one's own account (b)
  - Tweets that are older than a year are not usefull (b)
  - Show the Tweets' content on the map, not only the user name

To offer an offline mode in mobEx, we would needed to change the internal data structure of mobEx and implement a persistent database at the client side. This was considered too high effort compared with the goals of our study. Furthermore, this feature was just mentioned once and thus rated with low priority. The comments regarding the server issues would have also ended up in huge effort by including, e. g., further data providers. In addition, some of the requests were out of our scope like showing the menu of the restaurants (as no structured data is available) or having opening hours for all entities (this solely depends on the data sources we integrate). Finally, the feedback on Twitter was discussed but also not implemented due to low priority.

4.5 Discussion of Results

The user suggestions were very helpful to make improvements and add features to the application. Some subjects had difficulties to understand the purpose of the time-wheel feature, which may be due to insufficient and incomplete data sources. Especially fields such as opening hours, which are essential for making full usage of the time-wheel, were often missing. So, it could happen that in the map view nothing changed when turning the time-wheel. More entities with opening hours could tremendously improve the user satisfaction in the future.

Another point of criticism—with partly contradictory user feedback—was the granularity of the facet structure. Some subjects complained that categories in the facet structure were too detailed while others suggested finer granularity. In the end, we decided not to take any reactive actions since the user feedback did not deliver a clear picture about the best level of granularity and the preferred granularity always depends on the individual user.

One particular issue, which was out of our influence and therefore hard to tackle, was the issue of "insufficient and incomplete data". It was criticized by several subjects and was mainly due to the quality of the data sources or missing entities. For example, some providers only delivered a limited amount of data entries per query rather than delivering all query results. This issue may be one of the reasons for the user dissatisfaction because it easily could happen that well-known POIs were not found. One approach to address this problem was to use several data providers to get a more complete result and at the same time use an entity resolution procedure on the server side which merged information from different data providers [8]. Yet, the usage of multiple data providers could not entirely solve the problem of incomplete data. This issue was especially mentioned during the open questions in the first qualitative study. But it holds also true for the second questionnaire, in which this question arose again.

5. STEP 2: QUANTITATIVE USER STUDY

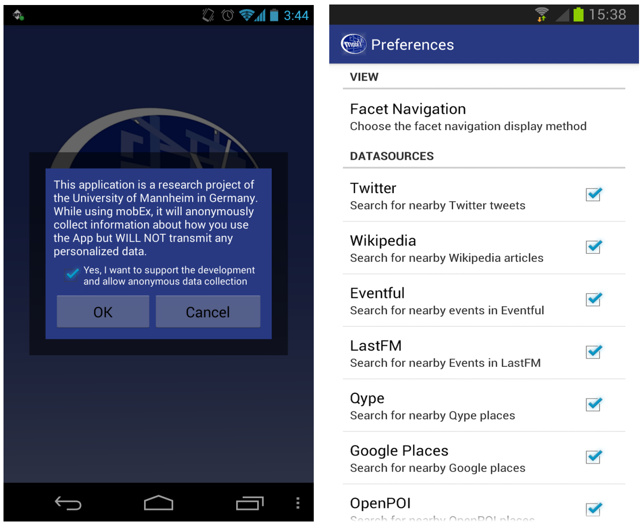
In the second step, we conducted a quantitative user study where the subjects were asked to use the app regularly, in the best case daily, for their own searches that usually arise. Within a phase of three weeks, we collected user events like starting or closing the application, using the time-wheel, or interactions with the facet navigation. Based on this collected data, we were able to analyze the user behavior. In order to verify some of the conclusions we drew from the logging data, we asked the subjects to answer a (second) questionnaire after the three week logging period.

The goal of this quantitative study was to get insights about how mobEx is used. We focused especially on how users interacted with the time-wheel and the grid navigation menu or list navigation menu. The primary questions we wanted to answer were firstly, if the time-wheel was used regularly to filter events and secondly, if list or grid navigation mode was preferred.

5.1 Logging Method

For logging the user's activity, we used Google Analytics[14]. This free service helps to analyze user traffic and can be customized to capture individual events. The data can be examined via a web application. As long as a user has not opted out logging and transferring the data to Google Analytics, every activity such as a click on a button or the interaction with the time-wheel results in an event that is

[14]http://www.google.com/analytics



(a) Users are prompted to consent to logging

(b) Default settings of data provider selections

Figure 4: Screenshot of the informed consent form and settings menu of the mobEx app

stored. The user was asked for his consent, when he or she first starts the app. A dialog screen as depicted in Figure 4a was shown, containing a short statement and a pre-selected check box to consent to the logging. A study has shown that this approach results in a good share of users who give their consent [4]. In order to accept the logging, the users just have to click on the "OK" button. If users did not accept the logging, they had to unselect the box and confirm with "OK". A click on "Cancel" closes the app.

5.2 Briefing of the Subjects

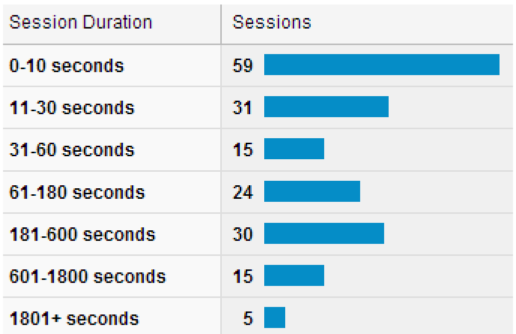
On July 21st, we started the quantitative part of our evaluation by sending a notification to our subjects. We asked them to update the app from the Android Market and place a shortcut of the app on their home screen. We motivated the subjects to generate continuous and reliable, non-artificial data and asked them to use our application instead of Google Maps during the three weeks. As introduction to the quantitative study, we provided an overview of the enhancements that were made after the first study.

During the study, the team members occasionally reminded their assigned subjects to use the application. Furthermore, we created interesting scenarios in which mobEx could be used to solve a realistic problem. We also reminded them about the study in the middle of this phase via a second motivating email and offered them our professional support. This procedure ensured a continuous data collection. During the three-week study, none of our 18 subjects dropped out.

5.3 Results of Quantitative Measurements

Common Usage Statistics By using Google Analytics as a tool to capture user activity, we can get detailed insights about the usage of mobEx during the user study. During three weeks, 18 subjects produced over 3000 single events in 179 sessions. A session is defined as active usage of the application until it is closed or after 30 seconds of inactivity.

A session lasted on average 3:57 minutes. Figure 5 shows the distribution of the different session durations. In Figure 6, the number of daily sessions during the three week period is plotted. 55% of the subjects used the application daily, 72% every second day or more. Figure 7 shows how many sessions each user has started during the study.



Session Duration	Sessions
0-10 seconds	59
11-30 seconds	31
31-60 seconds	15
61-180 seconds	24
181-600 seconds	30
601-1800 seconds	15
1801+ seconds	5

Figure 5: Distribution of Session Durations

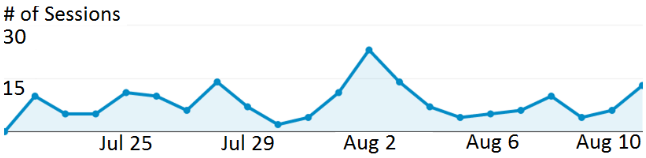


Figure 6: Daily sessions during the three week evaluation period

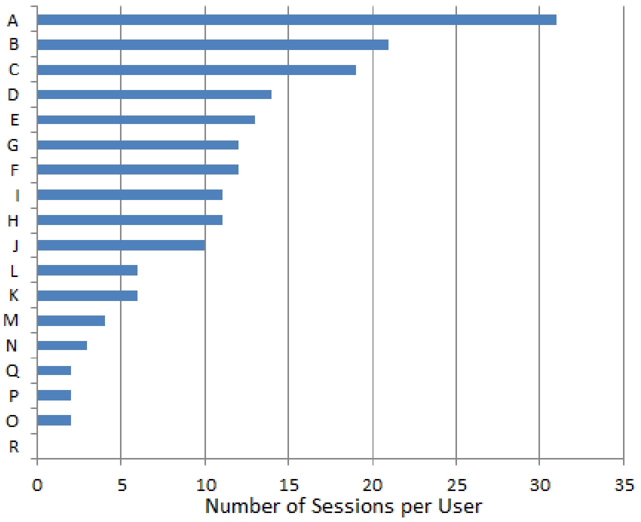


Figure 7: Sessions per user (sorted in descending order)

The users spent more time on the map screen (49 seconds) than on the screen showing the facets and the list of events (26 seconds). In terms of data providers that can be enabled or disabled using the preferences screen of the application (see Figure 4b), we observed only minor changes. DBpedia and Twitter have been disabled once and Twitter has been enabled also once (after being disabled).

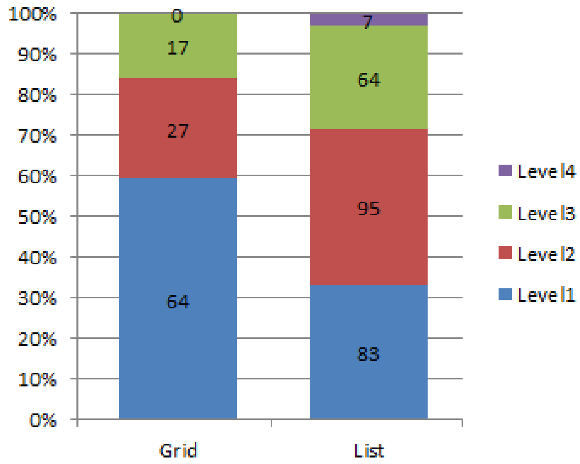


Figure 8: Distribution of how many times users navigated in a level of the facet tree using either grid or list navigation

Grid and List Navigation For investigating the acceptance of the grid and list navigation, we took two different measurements. First, the amount of sessions in which the grid or list navigation menu was used and second, how many subjects switched from grid to list or from list to grid navigation using the preferences menu. In the first case, we measured 94 sessions with list and 85 sessions with grid navigation enabled. Secondly, we observed seven times that subjects switched to grid navigation whereas four times subjects switched to list. The initial setting, grid or list, was decided by random. We counted ten subjects having list navigation enabled and eight subjects having grid navigation enabled at the beginning of the quantitative study. From these initial settings, we observed that in the end three subjects switched from list to grid, whereas none switched from grid to list.

To gain insight about how users interacted with the list or grid navigation, we took further detailed measurements. The purpose of the facet navigation is to browse through different levels of facets. With the initial screen being level zero, a click on a facet brings the user into level one. We measured how many times a user navigated into a deeper level and also how the user moved back to the upper level. For the grid navigation, we measured 64 clicks into level one, 27 clicks into level 2 and 17 clicks into level 3. In grid navigation, 10% of the "back" navigation to the upper level were done using the back button of the mobile phone and 90% using the back-button in the middle of the grid (see Figure 1a). For the list navigation, we measured 83 clicks into level 1, 95 clicks into level 2, 64 clicks into level 3 and 7 clicks into level 4. 19% used the back-button of the mobile phone and 81% the top list entry in order to move up to the prior level. Figure 8 visualizes this data.

We further evaluated the first actions users take on the facet screen by using the Behavior Flow of Google Analytics. The first action is defined as the very first interaction that the user do after entering the application. Subjects having the list navigation menu enabled switched in 37% of the cases directly to the map screen. 36% of the first actions taken were interactions with the result list, i. e., scrolling through

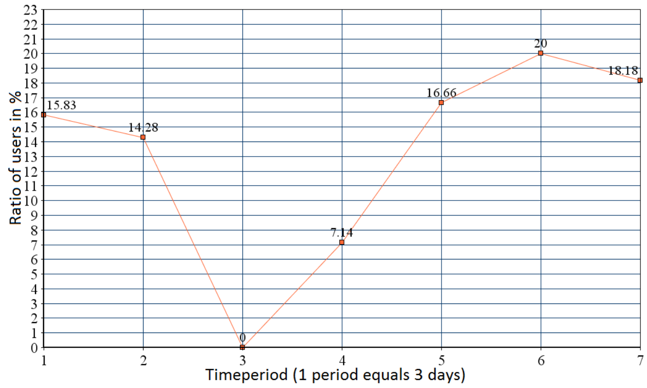


Figure 9: Ratio of users filtering the result list before switching to the map view

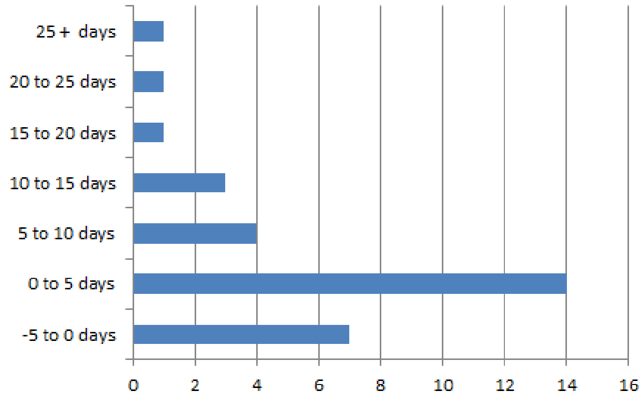


Figure 10: Distribution of the number of days users looked into the future

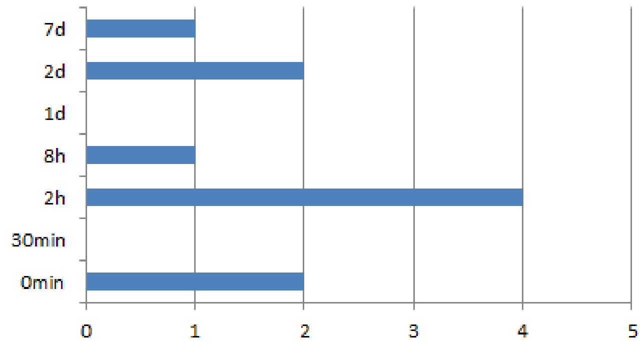


Figure 11: The different values the time interval can be changed to and the number of changes

the list of events and POIs, showing details of an event or searching for a POI. 10% can be assigned to the facet navigation (i. e., browsing or selecting facets). The corresponding values for the grid navigation menu show that in 26% of all first actions the subjects switched directly to the map screen, 30% were actions related to the result list and 15% of all actions were interactions with the grid navigation. The remaining percentages refer to interactions such as selecting the preferences menu or the help screen.

Another related measurement is to look at how many times users switched to the map view as first action in contrast to how many times users first filter the result list by selecting a facet with either grid or list navigation. To measure how this value changed within the three week period, we measured seven times. Each sample contains the sum of three consecutive days. This was necessary due to low usage of the time-wheel. The results are shown in Figure 9.

Time-wheel The time-wheel has been used in 25 of 179 sessions. In those 25 sessions, we measured on average 3.4 interactions per session with the time-wheel (e. g., scrolling or changing the time interval). After switching to the map screen on which the time-wheel is situated, 7% of the subjects used the time-wheel as first action. Other possible actions were a click on the buttons of the action bar, a click on an event on the map, or a change of the time interval. 65% of all users who used the time-wheel opened the detail screen for an event as subsequent action.

We measured the difference between the date that was set by the users operating the time-wheel and the date when the users left the map screen. The results summarized over different time spans are shown in Figure 10. Note that no data has been logged if the user did not use the time-wheel. Thus, Figure 10 only contains data from users who changed the date using the time-wheel.

The time interval has been changed only 10 times, so most of the time the users were working with the standard value of eight hours. The distribution of time intervals that the users have set can be seen in Figure 11. In the UI, the time interval is changed by a click on a button that simply alternates through a set of predefined time interval values. We only logged a change in the time interval if the value has been set and there is no further click on the button for four seconds.

The usage of the time-wheel has not strongly increased or decreased within the three weeks of our evaluation, when we normalize it with regard to the overall usage of the application. We measured 0.47 time-wheel interactions per session in the first week, 0.37 interactions per session in the second week, and 0.5 interactions per session in the third week.

5.4 Results of the Questionnaire

After the complete evaluation process, we identified that more than 61% of the subjects liked the overall application (“Generally, I like the app”) with a mean of 2.6 and a SD of 0.8, compared to approximately 70% at the beginning of the evaluation. At the end of our evaluation, eleven subjects are likely to recommend mobEx to friends (“Would you recommend the app to family and friends?”). Comparing this to the results of the first questionnaire, only five subjects made this recommendation. Furthermore, two out of three subjects are going to use the application in the future (“Are you going to use it further on?”). Sixteen of our subjects answered that the impression of the application (“Did the application confirm your first impression?”) did not get worse over the evaluation process. Eight users assessed even a better impression in the second questionnaire. Two subjects, who expressed a negative opinion of our app in both questionnaires explicitly complained about the incomplete data of mobEx.

The completeness of the data ("The data provided is in my opinion complete.") as well as the data quality ("In my opinion, the data is of good quality.") was rated on average with 2.4 and 2.2 (each with a SD= 1.0) on the 5-point-Likert scale. Please note, a value of 1 on the 5-point-Likert-scale refers to a strong agreement (i. e., a positive statement) and a value of 5 to a strong disagreement (i. e., a negative statement). 55% of the subjects rated with a score of 2 or better regarding the question whether some entries are missing ("I did not have the feeling that some events or places were missing.").

Nine subjects preferred the grid navigation menu in the facet screen whereas six subjects preferred the list navigation menu. The remaining three subjects did not have a preference for list or grid ("Which layout do you like more?"). The sorting of the facets in the grid menu or list menu was rated with a mean of 2.2 and a SD of 0.6 ("The sorting of the facets in the grid/list menu was very intuitive."). Thirteen subjects rated the layout of the user interface with a value of 2 or higher at the end of our evaluation ("I did like the layout."). This number increased compared to the beginning (eleven subjects liked the app at the beginning). Just one subject down rated the layout in the second questionnaire compared to the first rating. In the second questionnaire, the subjects rated the loading time of the data with an average value of 2.1 and SD of 0.8 ("The loading time of the data was appropriate."). This result did not change to the beginning of the evaluation. The software stability of the app was rated with a mean of 1.9 and SD of 1 ("The app works very stable."). Finally, we also asked how often the subjects did not find an entity they searched for ("How often did you end up without success?"), which was rated on average with 2.7 and a SD of 1.1.

5.5 Discussion of Results

Over the three weeks of logging user data, we could observe a balanced usage of mobEx. The subjects used the application on a regular basis. Thus, it produced enough data to create measurements about the behavior of the users. Regarding the navigation type, we observed that more subjects were using the grid navigation menu at the end of the study. Having the list navigation menu enabled, more users were using the back-button of the mobile phone in order to go up to a parent facet. This result is expected, as in some cases the back button in the list view is not visible as it is the first entry in the list whereas the back button in the grid view can be reached immediately. Our data further shows that grid users most of the times stop at the first facet level when navigating through the facets while list users tend to move frequently between level two and three, see Figure 8. Moreover grid users were more likely to filter the results using the grid navigation before switching to the map, whereas list users were more likely to switch directly to the map instead of interacting with the list navigation. Thus, the grid navigation menu attracted more attention from the subjects in order to use it for filtering the result list.

The measurements for the time-wheel show relatively low usage numbers throughout the evaluation. In only 14% of all sessions, the time-wheel has been used. 7% of the subjects used the time-wheel as first action when entering the map screen. Thus, the number of logged events is quite low and representative conclusions are difficult to draw. The time aspect has not been accepted or the benefits of changing the date and time have not been clear. The time has been changed mostly to only a few days into the future. Changes of the date to weeks or months in the future were very rare, even though our dataset usually contains events that take place in up to one year ahead. One reason is likely the lack of appropriate data with time-dependent information. It is necessary to add more time-dependent events and POIs and intensify the focus on the temporal dimension in the app. This would increase the usefulness of the time-wheel.

In our field study, we faced the problem that we were not able to tell if a user has found the event that he intended to find while using mobEx, i. e., if the search was successful. Thus, the logged data would be more helpful if we could have observed a set of actions together with the information whether the search was successful or not. This would be possible by either conducting a controlled study with given tasks or by implementing an extra button in the app that can be pressed when the search was successful.

6. RELATED WORK

The idea of a grid-based faceted navigation on mobile phones was initially presented in 2006 by Karlson et al. [5]. The authors describe a novel approach for searching large data sets via a 3x3 grid containing facets under which the data is classified. In a user study with 17 subjects, they gained insights about the performance of the grid-based approach in comparison to a text-based one. The results showed that for tasks where only certain characteristics of the data is known, facet navigation is faster. If some specific aspects of the data, like a describing name is known, text-based search performs better. Another finding of the study was that although users like the faceted navigation, they were frustrated when data was not classified as expected. Another approach for faceted browsing is FacetZoom [2]. FacetZoom is a UI widget that scales well in terms of different screen sizes and allows to browse through different levels of structured data. This approach is different from the grid navigation used in mobEx in a way that every hierarchical level is presented as horizontal bar that is subdivided in as many cells as nodes are available on that level. The bar can be scrolled horizontally and nodes can be selected with a click in order to navigate to the next level. The authors also conducted a user study that showed that different screen sizes had no effect on the performance when solving different tasks.

The procedure of our quantitative field study was based on the experience of Sohn et. al [11]. The authors describe a two-week diary study with 20 subjects about mobile information needs. During the experiment period, the subjects were asked to keep track about their information needs when being on the go with a mobile phone. Similar to our study, the subjects were briefed regularly and allowed to use their own cell phones. The study helped us in designing the briefing sessions as well as the rewarding system for the focus group. Since our research goal was also related to mobile information needs, we also decided to conduct a field study in order to have a high external validity and to find out how the subjects use mobEx in their 'natural habitat'. Another field study that was conducted with users in their 'natural habitat' was done by Niels Henze [4]. They describe how they used a mobile app as a tool to conduct a number of HCI studies with a large user group. In comparison to most other studies which are conducted in a highly controlled area, this work provides - with over 400,000 installations - a very high external validity. Since we also conducted a field study with subjects in a non-laboratory environment, our process of publishing the mobEx app was very similar. We also deployed our application via the Android Market in order to reach the subjects and enable them to easily use, rate, or update mobEx from home or while they are on the go. This enabled us to deploy new versions of the app on a regular basis during the development phase of mobEx without the need to explicitly notify the users about each update. However, during the three weeks period of our quantitative study, we did not modify the application and kept it stable except from fixing minor errors. We did not need to push any critical updates during our test period.

Schneider et al. [10] compared a grid-based and list-based approach for faceted navigation on a mobile phone in a between group experiment with 64 subjects. In this experiment, subjects had to execute a set of tasks with both navigation approaches. The results of this study showed that the grid-based menu requires significantly more clicks, yet subjects need less time for completing the tasks. Further, the results indicated that the additional amount of clicks has no negative influence on the user satisfaction. Consequently, the authors suggest that the grid-based approach is the better choice, if efficiency is not the primary goal. However, it has to be kept in mind that this study was conducted in a laboratory environment while our study was done in a natural environment. Our subjects did not complete predefined tasks. Instead they were briefed to use the app under daily life conditions and to fulfill personal information needs.

7. CONCLUSION

Unlike existing applications, mobEx provides access to multiple databases and matches the data from different providers to offer a unified view on the entities [8]. Users are enabled to explore these entities with unique search features such as a time-wheel and facet navigation. The qualitative study revealed that the purpose of the time-wheel was not well understood at the beginning. We also observed that the usage of the time-wheel during the quantitative study was not very high. The low usage may be either because the users still found it uncommon to use the time-while or due to a lack of more time-dependent entities such as events and places with opening hours. Enriching the application with more time dependent events could make the time-wheel more practical and increase its usage in the future. Regarding the navigation type, we can conclude from the findings of the quantitative study and qualitative study that the grid navigation seems to be the better option when exploring large data spaces. However, it has to be kept in mind that the list navigation is faster and requires less clicks [10]. Thus, the choice which approach is the better one seems to be whether one wants to adopt a novel and thus grid-based approach for faceted search or adopt the list-based approach that has shown to be more efficient in selecting and exploring facets.

Acknowledgement. We thank the Telegate AG and in particular Katja Beer and Florian Rang for their support in conducting this research.

References

[1] Peter Morville & Jeffery Callender. *Search Patterns.* O'Reilly, 2010. ISBN 978-0-596-80227-1.

[2] Raimund Dachselt, Mathias Frisch, and Markus Weiland. Facetzoom: a continuous multi-scale widget for navigating hierarchical metadata. In *Conference on Human Factors in Computing Systems*, pages 1353–1356. ACM, 2008.

[3] G. Gediga, K.-C. Hamborg, and I. Düntsch. The isometrics usability inventory: An operationalization of iso 9241-10 supporting summative and formative evaluation of software systems. *Behaviour & Information Technology*, 18(3):151–164, 1999.

[4] Niels Henze. Hit it!: an apparatus for upscaling mobile hci studies. In *Extended Abstracts on Human Factors in Computing Systems*, pages 1333–1338. ACM, 2012.

[5] Amy K. Karlson, George G. Robertson, Daniel C. Robbins, Mary P. Czerwinski, and Greg R. Smith. Fathumb: a facet-based interface for mobile search. In *Conference on Human Factors in Computing Systems*, pages 711–720. ACM, 2006.

[6] F. Knip, C. Bikar, B. Pfister, B. Opitz, T. Sztyler, M. Jess, and A. Scherp. Investigating the usability of a mobile app for finding and exploring places and events. In *Workshop on Semantic Ambient Media Experiences; Helsinki, Finland.* Ambient Media Association, 2014.

[7] Gary Marchionini. Exploratory search: from finding to understanding. *Commun. ACM*, 49(4):41–46, April 2006.

[8] B. Opitz, T. Sztyler, M. Jess, F. Knip, C. Bikar, B. Pfister, and A. Scherp. An approach for incremental entity resolution at the example of social media data. In *AI Mashup Challenge; Crete, Greece.* CEUR-WS, 2014. http://ceur-ws.org/Vol-1200/.

[9] Daniel Schmeiß, Ansgar Scherp, and Steffen Staab. Integrated mobile visualization and interaction of events and pois. In *Int. Conf on Multimedia; Firenze, Italy*, pages 1567–1570. ACM, 2010.

[10] Mark Schneider, Ansgar Scherp, and Jochen Hunz. A comparative user study of faceted search in large data hierarchies on mobile devices. In *Mobile and Ubiquitous Multimedia; Luleå, Sweden*, page 28, 2013.

[11] Timothy Sohn, Kevin A. Li, William G. Griswold, and James D. Hollan. A diary study of mobile information needs. In *Conference on Human Factors in Computing Systems*, pages 433–442. ACM, 2008.